nature genetics

ARTICLES

https://doi.org/10.1038/s41588-019-0556-y

Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression

Jamie E. Craig[1,40], Xikun Han[2,3,40*], Ayub Qassim[1,40], Mark Hassall[1], Jessica N. Cooke Bailey[4], Tyler G. Kinzy[4], Anthony P. Khawaja[5], Jiyuan An[2], Henry Marshall[1], Puya Gharahkhani[2], Robert P. Igo Jr.[4], Stuart L. Graham[6], Paul R. Healey[7,8], Jue-Sheng Ong[2], Tiger Zhou[1], Owen Siggs[1], Matthew H. Law[2], Emmanuelle Souzeau[1], Bronwyn Ridge[1], Pirro G. Hysi[9], Kathryn P. Burdon[10], Richard A. Mills[1], John Landers[1], Jonathan B. Ruddle[11], Ashish Agar[12], Anna Galanopoulos[13], Andrew J. R. White[7,8], Colin E. Willoughby[14,15], Nicholas H. Andrew[1], Stephen Best[16], Andrea L. Vincent[17], Ivan Goldberg[18], Graham Radford-Smith[2], Nicholas G. Martin[2], Grant W. Montgomery[19], Veronique Vitart[20], Rene Hoehn[21,22], Robert Wojciechowski[23,24], Jost B. Jonas[25], Tin Aung[26], Louis R. Pasquale[27], Angela Jane Cree[28], Sobha Sivaprasad[29], Neeru A. Vallabh[30,31], NEIGHBORHOOD consortium[32], UK Biobank Eye and Vision Consortium[32], Ananth C. Viswanathan[5], Francesca Pasutto[33], Jonathan L. Haines[4], Caroline C. W. Klaver[34], Cornelia M. van Duijn[35], Robert J. Casson[36], Paul J. Foster[5], Peng Tee Khaw[5], Christopher J. Hammond[9], David A. Mackey[10,37], Paul Mitchell[38], Andrew J. Lotery[28], Janey L. Wiggs[39], Alex W. Hewitt[10,40] and Stuart MacGregor[2,40]

Glaucoma, a disease characterized by progressive optic nerve degeneration, can be prevented through timely diagnosis and treatment. We characterize optic nerve photographs of 67,040 UK Biobank participants and use a multitrait genetic model to identify risk loci for glaucoma. A glaucoma polygenic risk score (PRS) enables effective risk stratification in unselected glaucoma cases and modifies penetrance of the *MYOC* variant encoding p.Gln368Ter, the most common glaucoma-associated myocilin variant. In the unselected glaucoma population, individuals in the top PRS decile reach an absolute risk for glaucoma 10 years earlier than the bottom decile and are at 15-fold increased risk of developing advanced glaucoma (top 10% versus remaining 90%, odds ratio = 4.20). The PRS predicts glaucoma progression in prospectively monitored, early manifest glaucoma cases ($P = 0.004$) and surgical intervention in advanced disease ($P = 3.6 \times 10^{-6}$). This glaucoma PRS will facilitate the development of a personalized approach for earlier treatment of high-risk individuals, with less intensive monitoring and treatment being possible for lower-risk groups.

Glaucoma refers to a group of ocular conditions united by a clinically characteristic optic neuropathy associated with, but not dependent on, elevated intraocular pressure[1]. It is the leading cause of irreversible blindness worldwide and is predicted to affect 76 million by 2020 (ref. [2,3]). There is no single definitive biomarker for glaucoma, and diagnosis involves assessing clinical features, with characterization of the optic nerve head carrying the strongest evidential weight. Primary open-angle glaucoma (POAG) is the most prevalent subtype of glaucoma in people of European and African ancestry[2,4]. POAG is asymptomatic in the early stages; currently approximately half of all cases in the community are undiagnosed even in developed countries[5]. Early detection is paramount since existing treatments cannot restore vision that has been lost, and late presentation is a major risk factor for blindness[6]. Thus, better strategies to identify high-risk individuals are urgently needed[7]; more refined approaches can capitalize on the fact that POAG is one of the most heritable of all common human diseases[8–10]. The lack of a currently cost-effective screening strategy for glaucoma[7], coupled with very high heritability, make glaucoma an ideal candidate disease for the development and application of a PRS to facilitate risk stratification.

Overlap of features shared by healthy optic nerves with those in the early stages of glaucoma makes it a difficult disease to diagnose early, necessitating costly ongoing monitoring of patients for progressive optic nerve degeneration[1]. Once a glaucoma diagnosis is established, rates of progression vary widely between individuals, and considerable time can elapse before surveillance techniques adequately differentiate slow from more rapidly progressing cases[1]. Progressive vision loss from glaucoma can be slowed, or in some cases halted, by timely intervention to reduce intraocular pressure

A full list of affiliations appears at the end of the paper.

using medical therapy, laser trabeculoplasty or incisional surgery[1]. The ability to predict progression is currently crude, with delays in treatment escalation for high-risk individuals an important and inevitable consequence, as well as substantial cost and morbidity associated with the overtreatment of lower-risk cases.

The chronicity, heritability, clinical heterogeneity and treatability of POAG make it an ideal candidate for genetic risk profiling[11,12]. In this study, we evaluated the optic nerve head in 67,040 UK Biobank (UKB) participants, enabling the largest genome-wide association study (GWAS) of optic nerve morphology to date, using the vertical cup/disc ratio (VCDR) as an endophenotype for glaucoma. Then, we incorporated additional genetic data from a second well-established glaucoma endophenotype, intraocular pressure (IOP), and combined this with glaucoma disease status, using a recently developed multitrait analysis of GWAS (MTAG)[13] approach to first identify new risk loci for glaucoma and then generate a comprehensive glaucoma PRS. We examined the impact of newly implicated glaucoma genes in independent case-control cohorts from Australia, the United Kingdom and the United States, and then evaluated the utility of the PRS for predicting glaucoma risk and important clinical outcomes in well-characterized cases across a range of disease severities.

Results

Study design. Our overall study design is illustrated in Extended Data Fig. 1a. We first conducted a GWAS on glaucoma (7,947 cases and 119,318 controls) and on the key endophenotypes for glaucoma: VCDR (including new data on 67,040 UKB participants and International Glaucoma Genetics Consortium (IGGC), $n=23,899$); and IOP (including data on 103,914 UKB participants and GWAS summary statistics from the IGGC, $n=29,578$; Supplementary Table 1). These data were then combined using MTAG[13] to identify new glaucoma risk loci and construct a PRS. The clinical significance of the PRS was investigated in advanced glaucoma cases in two populations and a separate prospectively monitored clinical cohort with early manifest glaucoma. The predictive ability of the PRS was also explored in other datasets; however, to ensure our results generalize to further cohorts, we selected mutually exclusive samples for inclusion in the discovery and testing datasets to ensure no sample overlap. When required, we rederived the PRS to ensure no sample overlap (Extended Data Fig. 1b–d and Supplementary Note).

Discovery of new optic nerve morphology loci. The GWAS of VCDR (adjusted for vertical disc diameter) identified 76 statistically independent, genome-wide significant SNPs (66 loci), of which 49 (43 loci) had not previously been associated with VCDR (Supplementary Figs. 1 and 2 and Supplementary Table 2). Using linkage disequilibrium score regression, we found no evidence for genomic inflation (intercept = 1.04, s.e.m. = 0.01; Supplementary Fig. 3). The genetic correlation between VCDR (adjusted for vertical disc diameter) and glaucoma in the UKB was 0.50 (s.e.m. = 0.05); the correlation in effect size estimates at the 76 SNPs was 0.60 ($P=9.0\times10^{-9}$; Supplementary Fig. 4). We further combined the UKB VCDR (adjusted for vertical disc diameter) GWAS and the IGGC VCDR GWAS summary statistics using MTAG, and identified 107 independent genome-wide significant SNPs (across 90 loci; Supplementary Table 3) for VCDR (adjusted for vertical disc diameter). As reported previously, the genetic correlation between IOP and glaucoma was high (0.71)[14], but as expected the genetic correlation between VCDR (adjusted for vertical disc diameter) and IOP was substantially lower (0.22, s.e.m. = 0.03).

Discovery of previously unknown glaucoma loci via multivariate analysis. Given the high correlation between glaucoma and its endophenotypes, we then conducted a multivariate GWAS (with 8,002,429 SNPs after quality control) to identify 114 statistically independent SNPs (107 loci, $P<5\times10^{-8}$) associated with glaucoma; this includes all previously published glaucoma loci as well as 49 previously unknown loci (Fig. 1, Supplementary Figs. 5 and 6 and Supplementary Table 4). At the more stringent multiple testing threshold ($P<1\times10^{-8}$) suggested by a simulation study[15], 95 loci reached significance, 39 of which were previously unknown (Supplementary Table 4); 27 of the 49 top SNPs at these loci were not associated individually with any of the individual input traits at the genome-wide significance level ($P=5\times10^{-8}$) and only reached this threshold for glaucoma due to the MTAG method leveraging the strong correlation between the input traits. Then, we attempted to replicate the 49 previously unknown SNPs in two independent glaucoma cohorts (Australian and New Zealand Registry of Advanced Glaucoma (ANZRAG) and National Eye Institute Glaucoma Human Genetics Collaboration Heritable Overall Operational Database (NEIGHBORHOOD)). Given the much smaller effective sample size of these replication cohorts (versus the discovery datasets from the MTAG analysis), we did not expect all of the SNPs to be strongly associated; rather, if they were genuine associations, we would expect the odds ratios (ORs) to be highly concordant, with some of the smaller ORs being individually nonsignificant. The concordance between the discovery cohort and our replication cohorts log ORs was excellent (correlation 0.88, $P=1.6\times10^{-36}$), indicating that our multivariate model was successful in identifying genuine glaucoma risk loci (Fig. 2 and Supplementary Fig. 7). Of the 49 previously unknown SNPs, 9 were replicated after Bonferroni correction ($P<0.05/49=0.001$, one-sided test; bold text in Supplementary Table 4), 26 were associated at a nominal significance level ($P<0.05$, one-sided test; italic text in Supplementary Table 4) and 46 (94%) were in the expected direction. While the concordance between multivariate and glaucoma replication sample log ORs was high, only 9 of the 49 loci were significant for glaucoma after correction for multiple comparisons; further studies are required to replicate the remaining 40 loci for glaucoma.

We conducted a genome-wide gene-based association analysis and a gene set enrichment analysis to assess which predefined biological pathways were enriched in our multitrait glaucoma GWAS; we found 196 genes and 14 gene sets, respectively, that were significant after Bonferroni correction (Supplementary Tables 5 and 6). The most significant pathways were also previously implicated (that is, extracellular matrix, collagen and circulatory system development)[14,16]. Further studies are warranted to investigate the role of these pathways in the risk of glaucoma.

Optimizing the prediction of glaucoma risk by combining correlated traits. We derived our PRS based on the MTAG of GWAS data from glaucoma and its endophenotypes. As well as increasing the number of SNPs that reach genome-wide significance (mean chi-squared statistic increased from 1.12 to 1.30, implying our effective sample size was 2.59 times larger than if we had used UKB glaucoma cases and controls alone), our multivariate model improved the power of risk prediction by reducing the error in the estimate of the effect size for every SNP (assuming the MTAG homogeneity assumption is true; see Discussion)[13]. We first tested the discriminatory power of the MTAG-derived PRS in the ANZRAG cohort of advanced glaucoma. We found that SNPs with MTAG P values ≤ 0.001 (corresponding to 2,673 uncorrelated SNPs after linkage disequilibrium clumping at $r^2=0.1$ and a P value threshold of 0.001) had the highest Nagelkerke R^2 (13.2%) and area under the curve (AUC, 0.68, 95% confidence interval (CI) = 0.67–0.70; Supplementary Table 7). The MTAG PRS has better prediction ability than any of the input traits alone (Supplementary Table 8). Based on this, we set the P value threshold at 0.001 for all the remaining prediction target sets (Progression Risk Of Glaucoma: REelevant SNPs with Significant Association (PROGRESSA), Blue Mountains Eye Study (BMES), UKB).

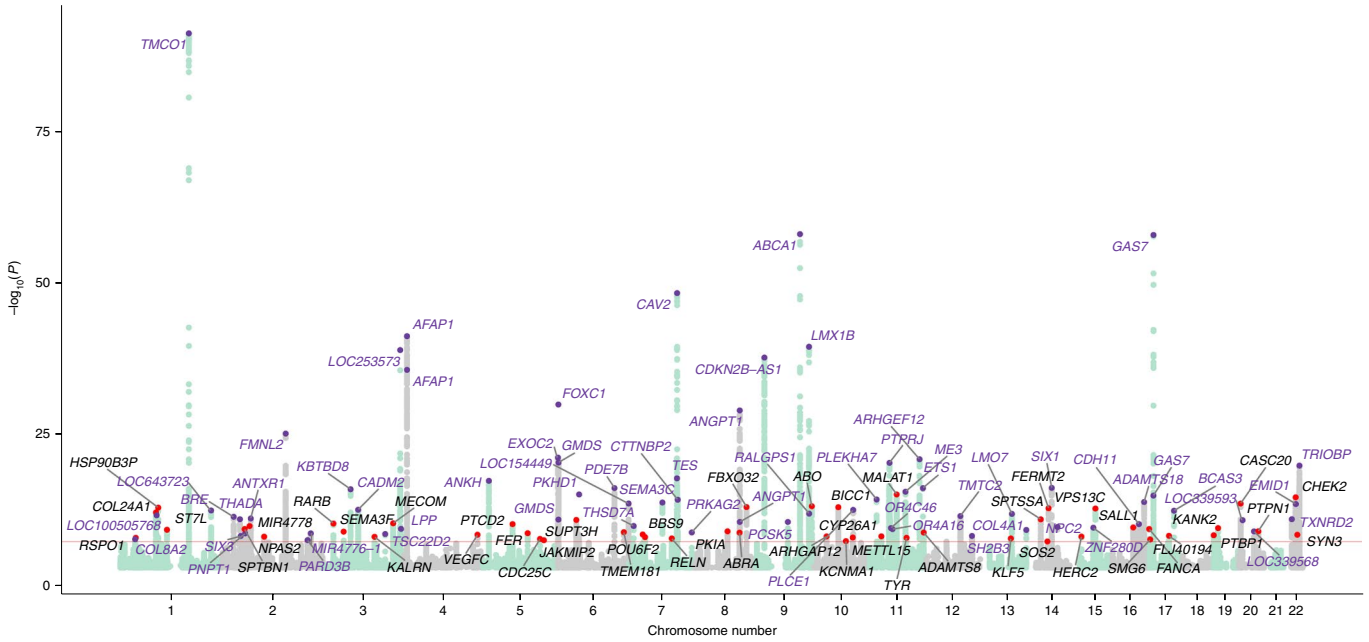


Fig. 1 | Manhattan plot displaying glaucoma-specific *P* values from the MTAG analysis. The samples used in the multitrait analysis are presented in Extended Data Fig. 1a. Previously unknown SNPs are highlighted with red dots, with the nearest gene names in black text. Known SNPs are highlighted with purple dots, with the nearest gene names in purple text. The red line is the genome-wide significance level at 5×10^{-8}.

The MTAG-derived PRS was effective at separating advanced glaucoma individuals in terms of risk, with a clear dose–response over deciles (Fig. 3a and Supplementary Fig. 8). In the ANZRAG cohort, individuals in the top decile of the PRS had a 14.9-fold higher risk (95% CI = 10.7–20.9) relative to the bottom decile, with even better discrimination for the more common high-tension glaucoma (OR = 21.5, 95% CI = 12.5–37.0) than normal-tension glaucoma (Supplementary Fig. 9). We replicated the dose–response of the PRS in a smaller UK advanced glaucoma dataset (Southampton and Liverpool); the top versus bottom PRS decile had an OR = 11.6 (95% CI = 6.0–25.3) and again had better discrimination for high-tension glaucoma (OR = 12.9, 95% CI = 6.2–31.3). While comparing the top and bottom deciles shows the dose–response across deciles, one can also consider the risk in the high-PRS individuals versus all others; when this is done in the ANZRAG cohort, the OR is 4.2 and 8.5 in the top 10 and 1%, respectively, of individuals versus all remaining individuals (Supplementary Table 9).

Glaucoma risk score performance in individuals carrying high-penetrance variants. Previous studies indicated that PRS modifies the penetrance of rare *BRCA1* and *BRCA2* mutation carriers for breast, ovarian and prostate cancers[17,18]. Although the MTAG-derived PRS only contains common variants, given that it indexes general glaucoma risk, we hypothesized that it could stratify individuals carrying known high-penetrance glaucoma variants. Pathogenic *MYOC* (myocilin) gene variants account for 2–4% of POAG cases in most populations, the most common disease-causing variant being that encoding p.Gln368Ter (rs74315329)[19]. Penetrance is age-related and is lower in population-based than family-based studies[19,20]. We speculated that this difference in penetrance could be due to enrichment of common glaucoma-associated variants in families modifying age-related penetrance. Within the UKB, we identified 965 *MYOC* p.Gln368Ter carriers based on imputation (Supplementary Note)[21]. Figure 3c shows the cumulative risk of glaucoma in p.Gln368Ter carriers, stratified by PRS tertiles. For p.Gln368Ter carriers in the lowest tertile PRS, glaucoma risk remained very low (2%) up to age 60. In contrast, the highest tertile PRS group had substantially increased risk of early diagnosis, reaching a sixfold increase in absolute risk of glaucoma by age 60, relative to the lowest PRS tertile (considering whole-age range, hazard ratio = 3.4, 95% CI = 1.7–6.6). This supports the utility of PRS in optimizing risk stratification and prediction, and early screening for patients carrying high-penetrance *MYOC* variants in the presence of high PRS scores.

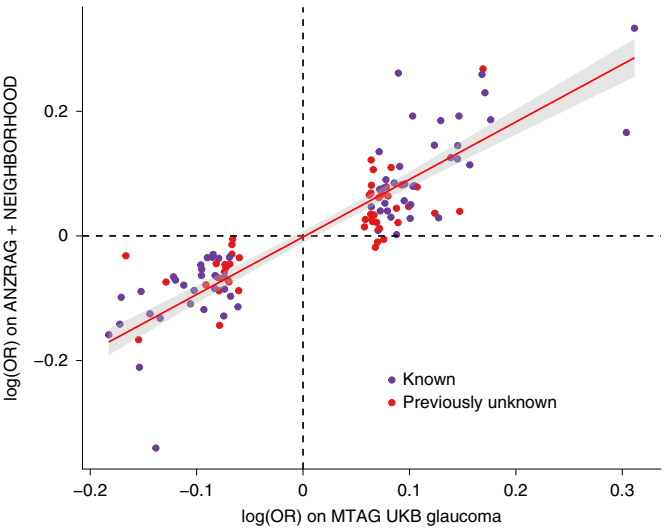


Fig. 2 | Comparison of the effect sizes (log ORs) for 114 genome-wide significant independent SNPs identified from the glaucoma multiple trait analysis of GWAS in the UKB versus those in independent glaucoma cohorts (meta-analysis of ANZRAG and NEIGHBORHOOD). The Pearson correlation coefficient is 0.88 ($P = 1.6\times10^{-36}$). The red line is the best fit line, with the 95% confidence interval region in gray. Previously unknown glaucoma SNPs are highlighted in red; known SNPs are shown in purple.

Potential for glaucoma risk score in screening in the general population. We considered a general population screening scenario using the UKB (PRS was rederived to ensure no sample overlap;

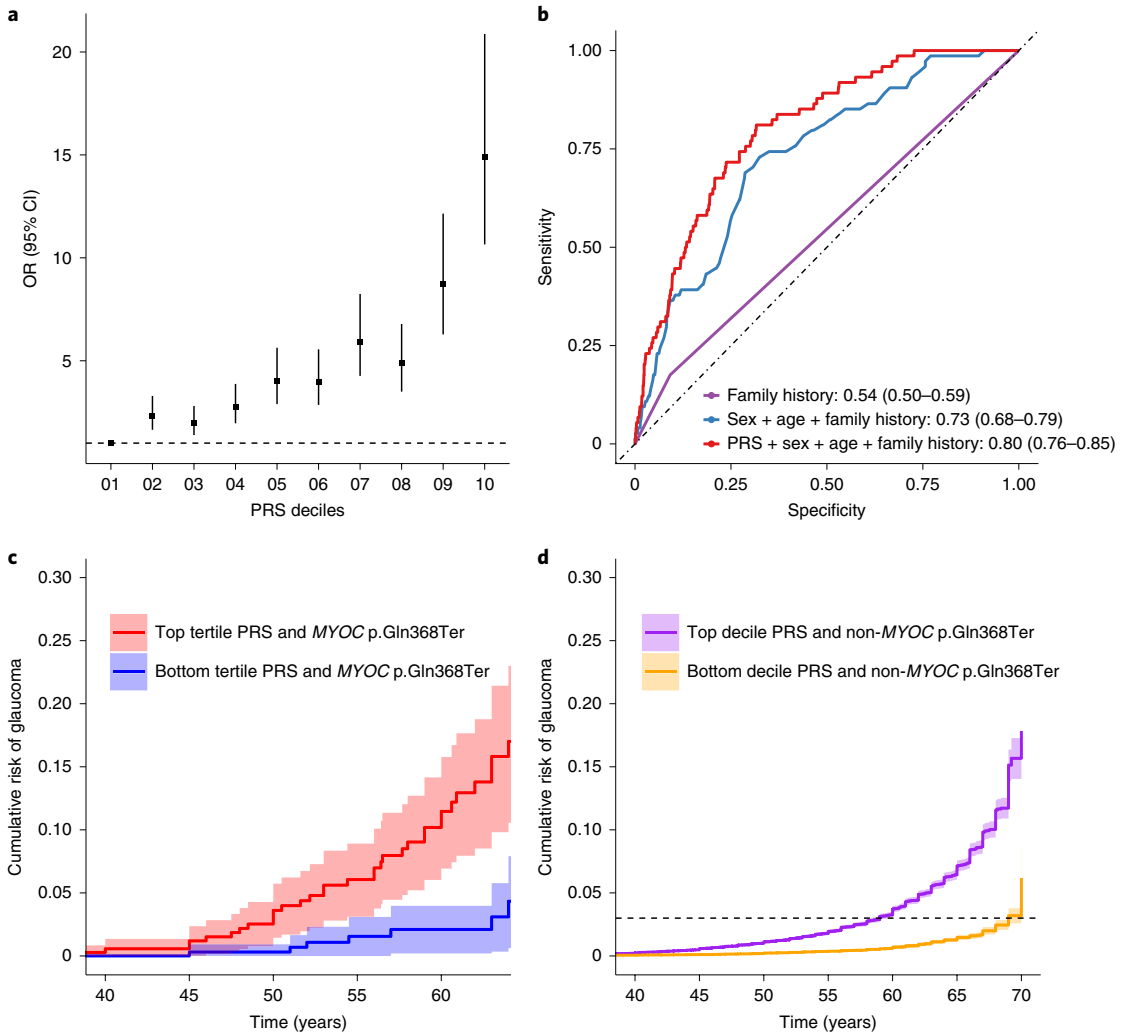


Fig. 3 | Multitrait analysis of GWAS PRS prediction. a, OR of developing advanced glaucoma in the ANZRAG cohort (with 1,734 advanced glaucoma cases and 2,938 controls) for each PRS decile. The square dots are the OR values (adjusted for sex and the first four principal components) and the error bars are the 95% CI. The dashed line is the reference at the bottom PRS decile (OR = 1). **b**, AUCs of PRS in the BMES cohort. The MTAG-derived PRS provided additional predictive ability on top of traditional risk factors (age, sex, and self-reported family history; DeLong test $P = 0.002$). The AUC is based on a logistic regression model with the coefficients for age, sex, family history and PRS estimated from the BMES data (Supplementary Table 10). **c**, Cumulative risk of glaucoma in UKB *MYOC* p.Gln368Ter carriers stratified by the PRS (adjusted for sex and first six genetic principal components). The cumulative risk of tertiles (with 95% CIs) of PRS are displayed given the relatively small number of *MYOC* p.Gln368Ter carriers ($n = 965$). **d**, Cumulative risk of glaucoma for people in the top and bottom decile (with 95% CIs) of PRS of the UKB participants who do not have the *MYOC* p.Gln368Ter variant (adjusted for sex and first six genetic principal components). The dashed line is the reference line of cumulative risk at 3%.

Extended Data Fig. 1d), where we excluded the 965 *MYOC* p.Gln368Ter carriers. Over the 40–69-year-old age range for individuals sampled in the UKB, glaucoma prevalence increases from 0.1% at age 40, reaching 3% (95% CI = 2.9–3.1) by age 64. The MTAG-derived PRS stratifies UKB participants very effectively; for those in the top PRS decile, 3% prevalence (prevalence in general population) is reached by age 59, while it takes an additional 10 years for this disease prevalence to be reached for people in the bottom PRS decile. Alternatively, prevalence can be well stratified by PRS deciles (Fig. 3d).

To benchmark the performance of the MTAG-derived PRS with traditional risk factors, we computed the AUC in datasets for which this was possible: BMES, UKB glaucoma (broad glaucoma definition) and UKB POAG (International Statistical Classification of Diseases and Related Health Problems, 10th revision (ICD-10) definition; Fig. 3b and Supplementary Table 11 (PRS was rederived to ensure no sample overlap), and Extended Data Fig. 1). In the BMES cohort, our PRS provided additional predictive ability beyond that imparted by traditional risk factors (age, sex and self-reported family history), with a significant change in the AUC (from 0.73 to 0.80, $P = 0.002$; Fig. 3b). Clear improvement in prediction using this PRS is also observed in people of South Asian ancestry (Supplementary Table 11), although we were underpowered to explore this further across other groups.

A previous study examined the cost-effectiveness requirements for glaucoma screening and highlighted the key 50–60 age bracket[7]. In the BMES data (Extended Data Fig. 1b), screening only those with a top decile PRS identified 40% of all early-onset cases in the 50–60 age bracket (40% of the 10 cases, $P = 0.013$). Such individuals represent a set of individuals who would probably benefit from referral for immediate clinical assessment with skilled clinical examination, retinal imaging and visual fields. We replicated this result in the UKB POAG cohort (ICD-10 cases in Extended Data Fig. 1c; the top 10% PRS screening finds 29% of 24 cases aged

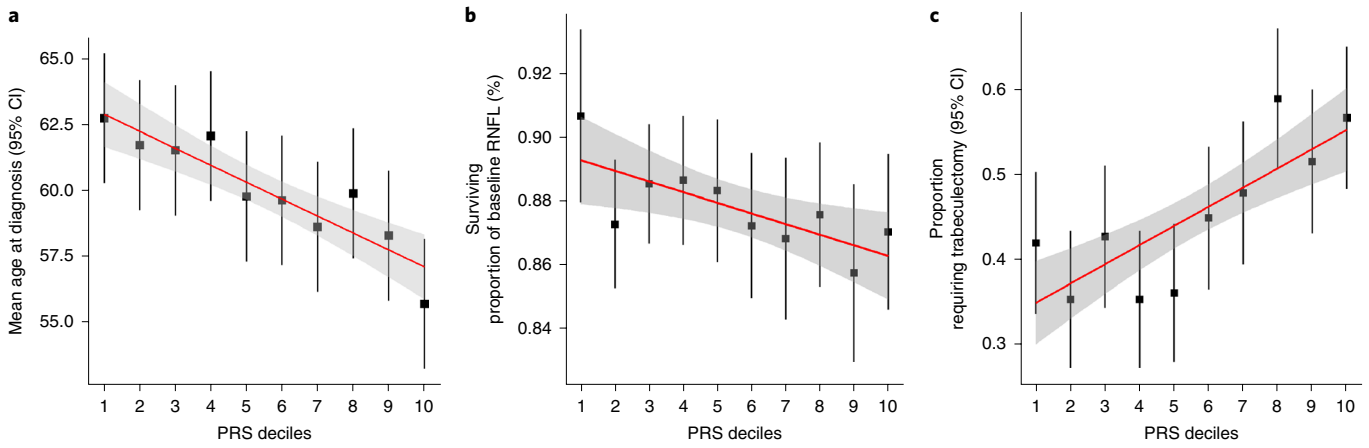


Fig. 4 | Clinical implications of the glaucoma PRS. a, Mean age at diagnosis (years) for each decile of the PRS in the ANZRAG cohort (linear regression $P=1.8\times10^{-5}$). A total of 1,336 cases had accurate age at diagnosis information. We calculated the mean age at diagnosis for each decile of PRS, adjusted for sex and the first four principal components in a linear regression model. The black squares are the regression-based mean age at diagnosis, with the error bars for the 95% CIs. The red line is the line of best fit, with the 95% CIs in gray. **b**, Proportion of preserved baseline retinal nerve fiber layer for PROGRESSA participants with early manifest glaucoma plotted against PRS decile ($n=388$; linear regression $P=0.004$). The black squares are the retinal nerve fiber layer proportions, with the error bars showing the 95% CIs. The remaining retinal nerve fiber layer proportion is calculated for the most affected quadrant of the most affected eye of each patient, as determined on optical coherence tomography scans at baseline and latest follow-up scan. **c**, Proportion of patients requiring trabeculectomy in either eye in the ANZRAG POAG cohort (linear regression $P=3.6\times10^{-6}$). There were 1,360 cases with records of surgical treatment status. The black squares represent the observed average proportion of cases in each decile of PRS who required trabeculectomy, with 95% CI bars. The line of best fit is shown in red, with the 95% CI shaded in gray.

50–60, $P=0.0075$). In this way, PRS-based screening would satisfy the cost-effectiveness requirements outlined by Burr et al.[7], identify a meaningful proportion of cases and capture those cases most at risk of severe disease.

Clinical implications of the glaucoma risk score. We evaluated the predictive power of the PRS in advanced glaucoma; in 1,336 ANZRAG advanced POAG cases with accurate age at diagnosis information available (Supplementary Table 12), the PRS was significantly associated with age at diagnosis of POAG ($P=1.8\times10^{-5}$). Individuals in the top 10% of the PRS distribution were, on average, diagnosed at ages 7 years younger than people in the bottom 10% (Fig. 4a). We also found that ANZRAG individuals with a higher PRS had more family members affected by glaucoma ($P=3.5\times10^{-9}$), with the highest decile having twice as many members affected (Supplementary Fig. 10).

Retinal nerve fiber layer (RNFL) thinning is a major structural change evident in early-stage glaucoma[22]. In the early manifestation of glaucoma (PROGRESSA) cohort, the PRS predicted both the proportion lost and rate of loss of peripapillary RNFL. Given that glaucomatous loss of retinal ganglion cells generally progresses unequally between eyes, with some quadrants of the retina damaged more rapidly than others, we analyzed the most affected quadrant of the most affected eye in individuals with early manifest glaucoma and >2 years of longitudinal optical coherence tomography data. The PRS was significantly associated with the proportion of RNFL lost from baseline to most recent review, even after adjustment for known risk factors: age, IOP and RNFL thickness at presentation ($P=0.004$; Fig. 4b and Supplementary Table 13). Expressed in terms of rate of loss, each decile change in PRS was associated with an accelerated progression rate of 0.05 µm year^{-1}, which was twice the rate of thinning per mmHg (approximately 1 decile change for IOP) of baseline IOP (0.022 µm year^{-1}).

Incisional surgery for glaucoma (trabeculectomy) is highly effective at reducing IOP, but has significant complications that can adversely impact vision[1]. Trabeculectomy is performed either when IOP cannot be controlled with medical or laser therapy, or when there is progressive visual field loss despite well-controlled IOP. Patients with a high PRS were more likely to have undergone surgery for glaucoma (Fig. 4c and Supplementary Fig. 11). In the ANZRAG cohort of POAG cases, a higher PRS was associated with requiring trabeculectomy, even after adjustment for maximum recorded IOP and age ($P=3.6\times10^{-6}$); the OR of requiring trabeculectomy in either eye for people in the top PRS decile was 1.78 (95% CI = 1.07–3.00) compared to the bottom decile. We observed a very similar trend in our UK replication (Southampton/Liverpool) samples (Supplementary Fig. 11).

Discussion

Through a large-scale multivariate GWAS we identified previously unknown genes for glaucoma, the leading cause of irreversible blindness worldwide[2]. Despite a smaller replication cohort, many of these previously unknown hits were replicated, and all but three SNPs showed a consistent direction of effect. We then expanded this analysis to derive a PRS and interrogated its utility across a wide spectrum of clinically relevant glaucoma outcomes.

From the multivariate GWAS, we identified 49 previously unknown loci associated with glaucoma (9 of which were replicated after correction for multiple comparisons in independent glaucoma case-control cohorts; 26 were replicated with a $P<0.05$). Most of the loci replicated at $P<0.001$ are at genes previously associated with glaucoma risk factors (myopia, central corneal thickness, IOP, VCDR). Specifically, *RSPO1* is associated with ocular axial length[23]. *BICC1* is associated with myopia and corneal astigmatism[24–26]. *POU6F2* modulates corneal thickness and increases glaucoma risk in animal experiments[27]. *FBXO32*, *PTPN1* and *VPS13C* are associated with IOP[14,28,29], while *CASC20* was identified in our VCDR (adjusted for vertical disc diameter) GWAS. These findings show that our multivariate GWAS improves the power to identify previously unknown glaucoma genes and advance our understanding of the causes of glaucoma risk.

The MTAG-derived PRS was validated in independent samples, confirming its high predictive ability. Individuals in the top PRS decile were at 15-fold increased risk of advanced glaucoma and at

21.5-fold increased risk of advanced high-tension glaucoma, relative to the bottom decile, which represents a substantial improvement on previously reported genetic profiling strategies, where, based on SNPs that were genome-wide significantly associated with IOP and SNPs previously associated with VCDR and glaucoma, top decile individuals had a 5.6-fold increased risk[14]. This new glaucoma PRS also outperforms those derived from other well-studied conditions; for example, our OR comparing the top 1% PRS individuals versus the remaining individuals was 8.5, which is higher than that seen in a recent study that surveyed coronary artery, atrial fibrillation, type 2 diabetes, inflammatory bowel disease and breast cancer[30]. The etiology of complex diseases depends on both environmental and genetic factors; thus, PRS alone will never achieve the very high predictive power (for example, AUC > 0.99) required for accurate population screening[31]. Our glaucoma PRS will be primarily useful for stratifying individuals into risk groups; for example, in the BMES data, screening the top decile of the PRS in individuals between 50 and 60 years old identifies 40% of cases. Moreover, as argued by Khera et al.[30], individuals with a high PRS for glaucoma are probably at a similar risk to individuals carrying rare 'high-penetrance' *MYOC* mutations[20]. Finally, PRS performance for glaucoma is particularly noteworthy given the clinical implications of identifying at-risk individuals and the prevention of irreversible blindness with readily available treatment proven to be effective at preventing visual loss.

While current treatments are effective in preventing or reducing POAG progression[12], many patients are not diagnosed before irreversible damage to visual function has already occurred. Earlier diagnosis of glaucoma can reduce glaucoma blindness and our work demonstrates that people with a higher PRS require earlier clinical assessment. In the UKB, individuals in the top PRS decile reach an equivalent absolute risk for glaucoma 10 years earlier than people in the bottom decile. In advanced glaucoma cases, individuals in the top decile were diagnosed at ages 7 years earlier than those in the bottom decile. Similarly, the MTAG-derived PRS was associated with significantly earlier disease onset in UKB *MYOC* p.Gln368Ter carriers who are at high disease risk. The MTAG-derived PRS can also identify people with early manifest glaucoma who are at higher probability of disease progression, as well as the probability of requiring surgical intervention, which is highly effective at reducing IOP but carries substantial treatment morbidity, meaning it should always be targeted specifically at those at higher risk of disease progression and blindness.

A concern with the MTAG method is the homogeneous assumption, which could be violated for some SNPs that have no effect on one trait but are non-null for other traits (that is, it is possible that a small number of the variants may be more specific for IOP or VCDR rather than glaucoma). The homogeneity assumption has been studied in detail by Turley et al.[13] We have evaluated the possible inflation using the maximum false discovery rate (FDR) as recommended by Turley et al.[13]. The baseline maximum FDR for the MTAG glaucoma-specific input GWAS summary statistics is 0.049, and the maximum FDR for the MTAG glaucoma-specific output summary statistics is 0.03. Since these are similar, there is no evidence of inflation due to violation of the homogeneity assumption. As recommended by the MTAG authors, we also performed replication analysis to assess the credibility of previously unknown SNPs in two independent datasets (an Australasian cohort of advanced glaucoma (ANZRAG) and a consortium of cohorts from the United States (NEIGHBORHOOD)); this analysis shows there is very good concordance between the MTAG-based effect sizes and those from the glaucoma cohorts. Furthermore, using MTAG output instead of the individual input traits improves the predictions in independent cohorts (Supplementary Table 8), providing additional evidence that we are not merely identifying IOP- or VCDR-specific loci that have no effect on glaucoma. Further research needs to be undertaken to investigate the biological mechanisms of these previously unknown genes on glaucoma risk.

A limitation of this work is that in our 7,947 UKB glaucoma cases, only a small proportion had documented disease subtype; however, since the proportion of UK glaucoma cases that have POAG is high (87% in a recent study by Chan et al.[4]), this probably would not have a large influence on our results. A further limitation is that it is not yet clear how applicable our findings are to other populations. We showed that the PRS improved prediction accuracy over and above traditional risk factors in homogeneous groups (as defined by genetic principal components) of either European or South Asian ancestry. The performance of the PRS in other populations should be tested to investigate the generalizability of our findings. The performance of the PRS in aiding clinical decision-making and guiding earlier treatment could be evaluated prospectively in a longitudinal intervention study, with participants randomized to have their PRS provided or withheld from their treating specialist.

In summary, we have applied a multivariate approach using weighted data on glaucoma, and the endophenotypes IOP and VCDR, to identify previously unknown glaucoma loci, and develop a PRS. This PRS was shown to be predictive of: (1) increased risk of advanced glaucoma; (2) glaucoma status significantly beyond traditional risk factors; (3) earlier age of glaucoma diagnosis; (4) high levels of absolute risk in persons carrying high-penetrance glaucoma variants; (5) increased probability of disease progression in early-stage disease; and (6) increased probability of incisional glaucoma surgery in advanced disease. This glaucoma PRS has good predictive power across a range of clinical cohorts and its application will facilitate the rational allocation of resources through clinical screening and timely treatment in high-risk patients, with reduced clinical monitoring costs in lower-risk groups.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41588-019-0556-y.

Received: 24 May 2019; Accepted: 21 November 2019;
Published online: 20 January 2020

References

1. Weinreb, R. N. & Khaw, P. T. Primary open-angle glaucoma. *Lancet* **363**, 1711–1720 (2004).
2. Tham, Y.-C. et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology* **121**, 2081–2090 (2014).
3. Quigley, H. A. & Broman, A. T. The number of people with glaucoma worldwide in 2010 and 2020. *Br. J. Ophthalmol.* **90**, 262–267 (2006).
4. Chan, M. P. Y. et al. Glaucoma and intraocular pressure in EPIC-Norfolk Eye Study: cross sectional study. *BMJ* **358**, j3889 (2017).
5. Mitchell, P., Smith, W., Attebo, K. & Healey, P. R. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology* **103**, 1661–1669 (1996).
6. Fraser, S., Bunce, C. & Wormald, R. Risk factors for late presentation in chronic glaucoma. *Invest. Ophthalmol. Vis. Sci.* **40**, 2251–2257 (1999).
7. Burr, J. M. et al. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol. Assess.* **11**, https://doi.org/10.3310/hta11410 (2007).
8. Wang, K., Gaitsch, H., Poon, H., Cox, N. J. & Rzhetsky, A. Classification of common human diseases derived from shared genetic and environmental determinants. *Nat. Genet.* **49**, 1319–1325 (2017).
9. Sanfilippo, P. G., Hewitt, A. W., Hammond, C. J. & Mackey, D. A. The heritability of ocular traits. *Surv. Ophthalmol.* **55**, 561–583 (2010).
10. Choquet, H. et al. A multiethnic genome-wide association study of primary open-angle glaucoma identifies novel risk loci. *Nat. Commun.* **9**, 2278 (2018).
11. Leske, M. C., Heijl, A., Hyman, L., Bengtsson, B. & Komaroff, E. Factors for progression and glaucoma treatment: the Early Manifest Glaucoma Trial. *Curr. Opin. Ophthalmol.* **15**, 102–106 (2004).
12. Garway-Heath, D. F. et al. Latanoprost for open-angle glaucoma (UKGTS): a randomised, multicentre, placebo-controlled trial. *Lancet* **385**, 1295–1304 (2015).

13. Turley, P. et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat. Genet.* **50**, 229–237 (2018).
14. MacGregor, S. et al. Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma. *Nat. Genet.* **50**, 1067–1071 (2018).
15. Wu, Y., Zheng, Z., Visscher, P. M. & Yang, J. Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biol.* **18**, 86 (2017).
16. Huang, L. et al. Genome-wide analysis identified 17 new loci influencing intraocular pressure in Chinese population. *Sci. China Life Sci.* **62**, 153–164 (2019).
17. Kuchenbaecker, K. B. et al. Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in *BRCA1* and *BRCA2* mutation carriers. *J. Natl. Cancer Inst.* **109**, djw302 (2017).
18. Lecarpentier, J. et al. Prediction of breast and prostate cancer risks in male *BRCA1* and *BRCA2* mutation carriers using polygenic risk scores. *J. Clin. Oncol.* **35**, 2240–2250 (2017).
19. Hewitt, A. W., Mackey, D. A. & Craig, J. E. Myocilin allele-specific glaucoma phenotype database. *Hum. Mutat.* **29**, 207–211 (2008).
20. Han, X. et al. Myocilin gene Gln368Ter variant penetrance and association with glaucoma in population-based and registry-based studies. *JAMA Ophthalmol.* **137**, 28–35 (2019).
21. Gharahkhani, P. et al. Accurate imputation-based screening of Gln368Ter myocilin variant in primary open-angle glaucoma. *Invest. Ophthalmol. Vis. Sci.* **56**, 5087–5093 (2015).
22. Na, J. H. et al. Detection of glaucoma progression by assessment of segmented macular thickness data obtained using spectral domain optical coherence tomography. *Invest. Ophthalmol. Vis. Sci.* **53**, 3817–3826 (2012).
23. Cheng, C.-Y. et al. Nine loci for ocular axial length identified through genome-wide association studies, including shared loci with refractive error. *Am. J. Hum. Genet.* **93**, 264–277 (2013).
24. Pickrell, J. K. et al. Detection and interpretation of shared genetic influences on 42 human traits. *Nat. Genet.* **48**, 709–717 (2016).
25. Verhoeven, V. J. M. et al. Genome-wide meta-analyses of multiancestry cohorts identify multiple new susceptibility loci for refractive error and myopia. *Nat. Genet.* **45**, 314–318 (2013).
26. Lopes, M. C. et al. Identification of a candidate gene for astigmatism. *Invest. Ophthalmol. Vis. Sci.* **54**, 1260–1267 (2013).
27. King, R. et al. Genomic locus modulating corneal thickness in the mouse identifies *POU6F2* as a potential risk of developing glaucoma. *PLoS Genet.* **14**, e1007145 (2018).
28. Khawaja, A. P. et al. Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. *Nat. Genet.* **50**, 778–782 (2018).
29. Gao, X. R., Huang, H., Nannini, D. R., Fan, F. & Kim, H. Genome-wide association analyses identify new loci influencing intraocular pressure. *Hum. Mol. Genet.* **27**, 2205–2213 (2018).
30. Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
31. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348 (2013).

[1]Department of Ophthalmology, Flinders University, Flinders Medical Centre, Bedford Park, South Australia, Australia. [2]Queensland Institute of Medical Research Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [3]School of Medicine, University of Queensland, Brisbane, Queensland, Australia. [4]Department of Population and Quantitative Health Sciences, Institute for Computational Biology, Case Western Reserve University School of Medicine, Cleveland, OH, USA. [5]National Institute for Health Research Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK. [6]Faculty of Medicine and Health Sciences, Macquarie University, Sydney, New South Wales, Australia. [7]Centre for Vision Research, Westmead Institute for Medical Research, University of Sydney, Sydney, New South Wales, Australia. [8]Clinical Ophthalmology & Eye Health, Westmead Clinical School, University of Sydney, Sydney, New South Wales, Australia. [9]Department of Ophthalmology, King's College London, St. Thomas' Hospital, London, UK. [10]Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia. [11]Centre for Eye Research Australia, University of Melbourne, Melbourne, Victoria, Australia. [12]Department of Ophthalmology, Prince of Wales Hospital Randwick, Sydney, New South Wales, Australia. [13]South Australian Institute of Ophthalmology, Royal Adelaide Hospital, Adelaide, South Australia, Australia. [14]Biomedical Sciences Research Institute, Ulster University, Coleraine, UK. [15]Royal Victoria Hospital, Belfast Health and Social Care Trust, Belfast, UK. [16]Eye Department, Greenlane Clinical Centre, Auckland District Health Board, Auckland, New Zealand. [17]Department of Ophthalmology, University of Auckland, Auckland, New Zealand. [18]Discipline of Ophthalmology, Faculty of Medicine and Health, University of Sydney, Sydney Eye Hospital, Sydney, New South Wales, Australia. [19]Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland, Australia. [20]Medical Research Council Human Genetics Unit, Medical Research Council Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK. [21]Department of Ophthalmology, University Hospital Bern, Inselspital, University of Bern, Bern, Switzerland. [22]Department of Ophthalmology, University Medical Center Mainz, Mainz, Germany. [23]Department of Epidemiology and Medicine, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. [24]Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA. [25]Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University of Heidelberg, Mannheim, Germany. [26]Singapore Eye Research Institute, Singapore National Eye Centre, Singapore, Singapore. [27]Department of Ophthalmology, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [28]Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK. [29]National Institute for Health Research Moorfields Biomedical Research Centre, London, UK. [30]Department of Eye and Vision Science, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK. [31]St Paul's Eye Unit, Royal Liverpool University Hospital, Liverpool, UK. [32]A list of members and affiliations appears in the Supplementary Note. [33]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany. [34]Department of Ophthalmology, Erasmus Medical Center, Rotterdam, the Netherlands. [35]Department of Epidemiology, Erasmus Medical Center, Rotterdam, the Netherlands. [36]South Australian Institute of Ophthalmology, University of Adelaide, Adelaide, South Australia, Australia. [37]Lions Eye Institute, Centre for Ophthalmology and Vision Sciences, University of Western Australia, Perth, Western Australia, Australia. [38]Department of Ophthalmology and Westmead Institute for Medical Research, University of Sydney, Sydney, New South Wales, Australia. [39]Department of Ophthalmology, Harvard Medical School, Massachusetts Eye and Ear Infirmary, Boston, MA, USA. [40]These authors contributed equally: Xikun Han, Jamie E. Craig, Ayub Qassim, Alex W. Hewitt, Stuart MacGregor. *e-mail: Xikun.Han@qimrberghofer.edu.au

Methods

Study design and overview. Our overall study design is illustrated in Extended Data Fig. 1. We first conducted a GWAS on glaucoma and on the key endophenotypes for glaucoma: VCDR and IOP. These data were then combined using MTAG[13], a method for combining multiple genetically correlated traits to maximize power to identify previously unknown loci and improve genetic risk prediction. Specifically, our MTAG analysis outputs glaucoma-specific effect size estimates and *P* values for SNPs across the genome. Newly associated loci ($P < 5 \times 10^{-8}$) were validated in two independent cohorts with well-characterized POAG. We created a PRS based on the MTAG GWAS summary statistics. The clinical significance of the PRS was investigated in advanced glaucoma cases in two populations and a separate prospectively monitored clinical cohort with early manifest glaucoma. The predictive ability of the PRS was also explored in other datasets; however, to ensure our results generalize to further cohorts, we selected mutually exclusive samples for inclusion in the discovery and testing datasets to ensure no sample overlap. When required, we rederived the PRS to avoid any sample overlap (Extended Data Fig. 1). Study procedures were performed in accordance with the World Medical Association Declaration of Helsinki ethical principles for medical research (2013 version).

Study populations. Detailed information of individual studies, phenotypic definitions and genetic quality control procedures are provided in the Supplementary Note.

The UKB is a population-based study of half a million people living in the United Kingdom[32]. We measured VCDR and vertical disc diameter in all patients with gradable retinal images (67,040 participants following exclusions, detailed in the Supplementary Note) and undertook a GWAS to identify SNPs influencing optic nerve head morphology. Vertical disc diameter adjustment of the VCDR was used to account for optic cup and disc size covariation[33,34]. To improve the power in the multitrait analysis, we combined the VCDR data with data on corneal-compensated IOP (103,914 participants) and glaucoma (7,947 cases, 119,318 controls) in the MTAG analysis[14]. We also used publicly available VCDR and IOP GWAS summary results for individuals of European ancestry from the IGGC ($n_{VCDR} = 23{,}899$, $n_{IOP} = 29{,}578$)[35].

The ANZRAG consists of 3,071 POAG cases of European ancestry, who were compared to 6,750 controls[36,37]. For sub-analyses restricted to advanced POAG, there were 1,734 advanced POAG cases and 2,938 controls; of these cases, 1,336 participants had accurate age at diagnosis information available. Replication of the ANZRAG findings was performed using 332 advanced glaucoma cases from Southampton and Liverpool in the United Kingdom; for case-control analysis, cases were matched to 3,000 randomly selected European ancestry individuals from the QSkin Sun and Health Study[38]. The NEIGHBORHOOD GWAS results were generated through meta-analyzing summary data from 8 independent datasets (3,853 POAG cases and 33,480 controls) of European ancestry from the United States[39].

The BMES is a population-based cohort study investigating the etiology of common ocular diseases among suburban residents aged 49 years or older in Australia[5]. Data from 74 POAG cases and 1,721 controls of European ancestry with genotype information were included.

The PROGRESSA study is a prospective longitudinal study of the clinical and genetic risk factors, and course of early-stage glaucoma ($n = 388$). Patients with confirmed early manifest POAG on sequential automated perimetry testing were consecutively recruited from ophthalmology clinics in South Australia (detailed criteria found in the Supplementary Note). Individuals underwent 6-monthly evaluation of IOP, optic disc assessment, RNFL analysis by optical coherence tomography and achromatic Humphrey visual field perimetry. Longitudinal data were used from all visits since baseline presentation; participants were followed for 1–8 years. The change in RNFL was measured between the baseline optical coherence tomography and the most recent scan in the most affected quadrant of the most affected eye. Treating clinicians and graders were unaware of the patient's genetic risk for glaucoma or any PRS data.

POAG in the ANZRAG, NEIGHBORHOOD, BMES and PROGRESSA cohorts was defined as outlined previously[40], and in accordance with the consensus statement from the World Glaucoma Association[41]. IOP was not used in the clinical case definition of POAG[41].

Statistical analysis. Detailed information on the statistical analysis is provided in the Supplementary Note.

For the VCDR (adjusted for vertical disc diameter) and IOP GWAS in the UKB, we used linear mixed models (BOLT-LMM software v.2.3) to account for cryptic relatedness and population stratification, adjusting for sex, age and the first ten principal components[42]. We meta-analyzed the UKB IOP GWAS results with those from the IGGC using the inverse variance-weighted method (METAL software 2011-03-25 release)[43]. For the UKB glaucoma GWAS, we removed relatives (pi-hat > 0.2 calculated using identity by descent based on autosomal markers) and used the PLINK software (v.1.90beta) for the association analysis[44].

Then, we conducted a multitrait GWAS using the MTAG v.1.0.7 software to combine the European ancestry GWAS summary statistics from the UKB glaucoma, UKB VCDR (adjusted for vertical disc diameter), IGGC VCDR and IOP meta-analysis (Extended Data Fig. 1)[13]. MTAG performs joint analysis of GWAS summary results from related traits to improve statistical power to identify new genes and maximize the predictive ability of our PRS[13]. In MTAG, GWAS summary results from related traits are used to construct the variance–covariance matrix of their SNP effects and estimation error; MTAG improves the accuracy of effect estimates by incorporating information from other genetic correlated traits. The MTAG method explicitly models sample overlap in the input studies and provides valid estimates even when sample overlap is present[13]. To benchmark the increase in effective sample size relative to just using UKB glaucoma, we calculated $(\overline{\chi 2}_{MTAG} - 1)/(\overline{\chi 2}_{GWAS} - 1)$, where $\overline{\chi 2}_{MTAG}$ and $\overline{\chi 2}_{GWAS}$ are the mean chi-squared statistics from MTAG and the UKB glaucoma analyses, respectively[13].

We used a stepwise model selection procedure in the GCTA-COJO software (v.1.26) to identify independent, genome-wide significant SNPs[45]. Gene-based and pathway analysis were conducted in MAGMA v.1.06, as implemented in FUMA v.1.3.1 (refs. [46,47]).

Prediction was based on the estimated glaucoma ORs from the MTAG analysis. To derive a PRS, we considered a range of *P* value thresholds (5×10^{-8}, 1×10^{-5}, 0.001, 0.05, 1) with linkage disequilibrium clumping $r^2 = 0.1$ for inclusion of SNPs in the prediction model, applying each to our first prediction cohort (advanced glaucoma from the ANZRAG cohort). To avoid falsely inflating prediction accuracy, we applied the threshold with the greatest predictive value in ANZRAG ($P \leq 0.001$) for the subsequent predictions into other target sets (rather than repeatedly taking the best *P* value threshold for each of the datasets). We tested the LDpred (v.0.9.09)[48] approach for PRS construction, although the predictions were no better than those from the thresholding approach described earlier. There was no sample overlap between any of the training and target datasets (Extended Data Fig. 1).

Bivariate linkage disequilibrium score regression was used to estimate the genetic correlation between pairs of traits[49]. The pROC package (v.1.14.0) was used to calculate the AUC[50]. Analyses were performed with the R software (v.3.4.1)[51].

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The UKB data are available through the UK Biobank Access Management System https://www.ukbiobank.ac.uk/. The GWAS summary statistics from the glaucoma MTAG analysis is available for research use at https://xikunhan.github.io/site/publication/. We will return the derived data fields following UKB policy; in due course, they will be available through the UK Biobank Access Management System.

References

32. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
33. Bengtsson, B. The variation and covariation of cup and disc diameters. *Acta Ophthalmol. (Copenh.)* **54**, 804–818 (1976).
34. Han, X. et al. Genome-wide association analysis of 95,549 individuals identifies novel loci and genes influencing optic disc morphology. *Hum. Mol. Genet.* https://doi.org/10.1093/hmg/ddz193 (2019).
35. Springelkamp, H. et al. New insights into the genetics of primary open-angle glaucoma based on meta-analyses of intraocular pressure and optic disc characteristics. *Hum. Mol. Genet.* **26**, 438–453 (2017).
36. Souzeau, E. et al. Australian and New Zealand Registry of Advanced Glaucoma: methodology and recruitment. *Clin. Exp. Ophthalmol.* **40**, 569–575 (2012).
37. Gharahkhani, P. et al. Common variants near *ABCA1*, *AFAP1* and *GMDS* confer risk of primary open-angle glaucoma. *Nat. Genet.* **46**, 1120–1125 (2014).
38. Olsen, C. M. et al. Cohort profile: the QSkin Sun and Health Study. *Int. J. Epidemiol.* **41**, 929–929i (2012).
39. Wiggs, J. L. et al. The NEIGHBOR consortium primary open-angle glaucoma genome-wide association study: rationale, study design, and clinical variables. *J. Glaucoma* **22**, 517–525 (2013).
40. Kwon, Y. H., Fingert, J. H., Kuehn, M. H. & Alward, W. L. M. Primary open-angle glaucoma. *N. Engl. J. Med.* **360**, 1113–1124 (2009).
41. Weinreb, R. N., Garway-Heath, D. F., Leung, C., Medeiros, F. A. & Liebmann, J. *Diagnosis of Primary Open Angle Glaucoma: WGA consensus series—10* (Kugler Publications, 2017).
42. Loh, P.-R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).
43. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
44. Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
45. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).

46. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput. Biol.* **11**, e1004219 (2015).
47. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
48. Vilhjálmsson, B. J. et al. Modeling linkage disequilibrium increases accuracy of polygenic risk scores. *Am. J. Hum. Genet.* **97**, 576–592 (2015).
49. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
50. Robin, X. et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* **12**, 77 (2011).
51. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2017).

Acknowledgements

This work was conducted using the UK Biobank Resource (application no. 25331) and publicly available data from the IGGC. The UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and Northwest Regional Development Agency. It also had funding from the Welsh Assembly Government, British Heart Foundation and Diabetes UK. The eye and vision dataset has been developed with additional funding from the National Institute for Health Research Biomedical Research Centre at Moorfields Eye Hospital and the UCL Institute of Ophthalmology, Fight for Sight charity, Moorfields Eye Charity, Macular Society, International Glaucoma Association and Alcon Research Institute. This work was also supported by grants from the National Health and Medical Research Council (NHMRC) of Australia (nos. 1107098, 1116360, 1116495, 1023911, 1150144, 1147571), the Ophthalmic Research Institute of Australia, the BrightFocus Foundation, the UK and Eire Glaucoma Society and charitable funds from the Royal Liverpool University Hospital. S.M., J.E.C., K.P.B., D.A.M. and A.W.H. are supported by NHMRC Fellowships (APP1154543, APP1154824, APP1059954, APP1154513, APP1103329). S.M. was supported by an Australian Research Council Future Fellowship (FT130101902). L.R.P. is supported by National Institutes of Health grant no. R01 EY015473. X.H. is supported by the University of Queensland Research Training Scholarship and Queensland Institute of Medical Research Berghofer PhD Top Up Scholarship. We thank D. Whiteman, R. Neale and C. Olson for providing access to the QSkin samples for use as controls as part of NHMRC grant no. 1063061. We thank S. Wood, J. Pearson and S. Gordon from the Queensland Institute of Medical Research Berghofer Research Institute for their support. The NEIGHBORHOOD consortium is supported by National Institutes of Health grant nos. P30 EY014104, R01 EY015473 and R01 EY022305.

Author contributions

S.M., J.E.C., A.W.H., X.H., P.G., J.L.W. and D.A.M. designed the study and obtained the funding. X.H., A.Q., M.H., J.N.C.B., T.G.K., A.P.K., P.G.H., J.A., H.M., P.G., R.P.I., J.-S.O., T.Z., O.S., M.H.L. and S.M. analyzed the data. J.E.C., X.H., A.Q., M.H., A.P.K., H.M., R.P.I., S.L.G., P.R.H., O.S., E.S., B.S., P.G.H., K.P.B., R.A.M., J.L., J.B.R., A.A., A.G., A.J.R.W., C.E.W., N.A., S.B., A.L.V., I.G., G.R.-S., N.G.M., G.W.M., V.V., R.H., R.W., J.B.J., T.A., L.R.P., A.J.C., S.S., N.A.V., A.C.V., F.P., J.L.H., C.C.W.K., C.M.v.D., R.J.C., P.J.F., P.T.K., C.J.H., D.A.M., P.M., A.J.L., J.L.W., A.W.H. and S.M. contributed to data collection and contributed to genotyping. X.H., J.E.C., A.Q., A.W.H. and S.M. wrote the first draft of the paper. All authors contributed to the final version of the paper.

Competing interests

D.A.M. is consultant/advisor to Allergan, Inc. J.E.C., A.W.H. and S.M. are listed as coinventors on a patent application for the use of genetic risk scores to determine risk and guide treatment.

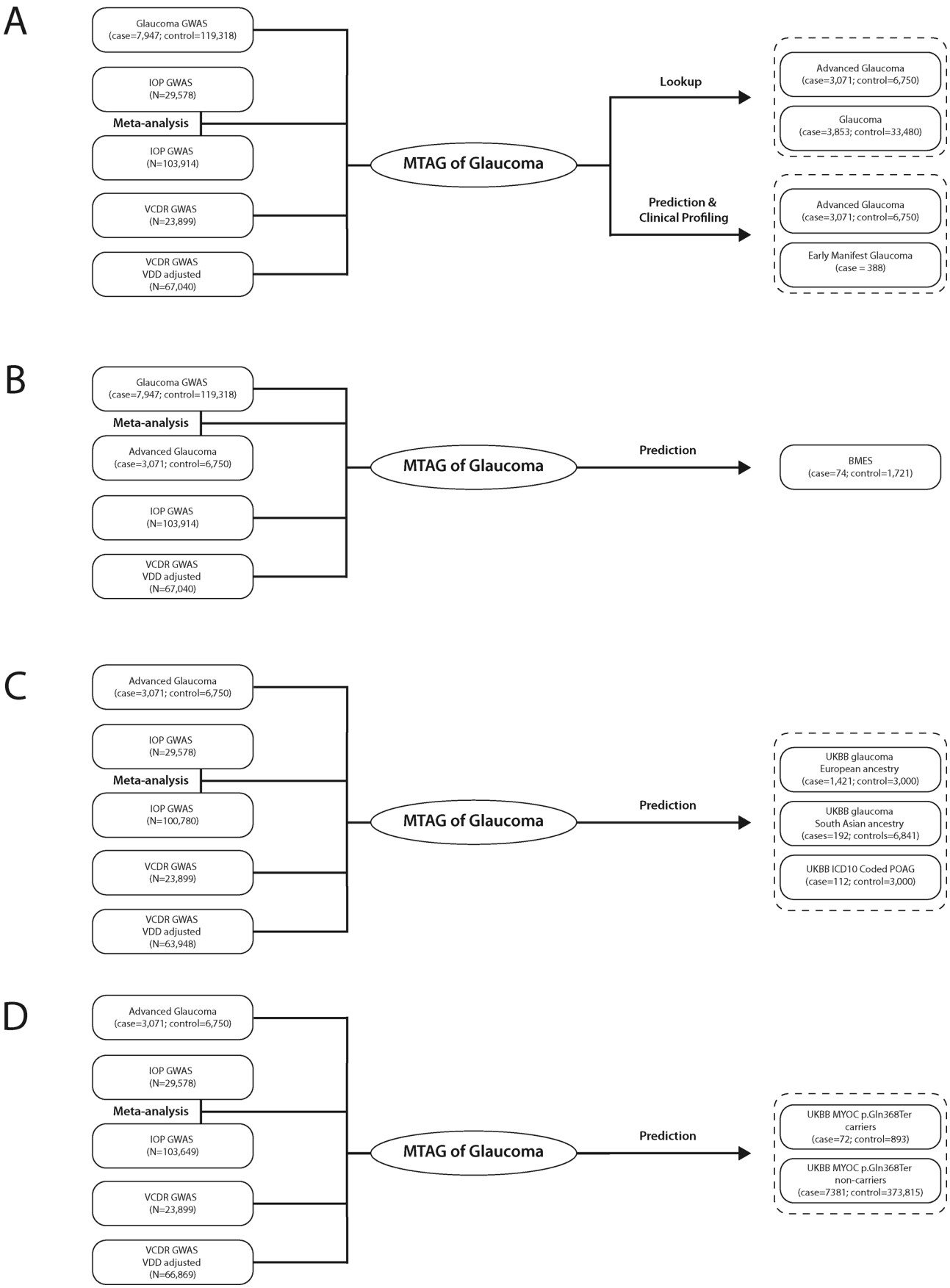
Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41588-019-0556-y.

Supplementary information is available for this paper at https://doi.org/10.1038/s41588-019-0556-y.

Correspondence and requests for materials should be addressed to X.H.

Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Study design. We applied the multi-trait analysis of GWAS (MTAG) algorithm to datasets of European descent (unless otherwise specified). **a**, We applied MTAG to four datasets (glaucoma case-control GWAS from the UKBB; GWAS meta-analysis of intraocular pressure (IOP) from the International Glaucoma Genetics Consortium (IGGC) and the UKBB; Vertical cup-disc ratio (VCDR) GWAS data that was either adjusted for vertical disc diameter (VDD) in the UKBB dataset; or not adjusted for VDD in the IGGC). Novel variants identified through this analysis were then confirmed in two independent data sets: an Australasian cohort of advanced glaucoma (ANZRAG) and a consortium of cohorts from the United States (NEIGHBORHOOD). The clinical significance of the PRS derived from the MTAG analysis was validated in independent samples: first, in advanced glaucoma cases (ANZRAG and samples from Southampton/Liverpool in the UK), and second, in a prospectively monitored clinical cohort with early manifest glaucoma (PROGRESSA). **b**, Prediction in BMES, where we removed the IGGC VCDR and IGGC IOP GWAS from the training datasets, given that they contain BMES data. **c**, Prediction in the UKBB glaucoma and ICD-10 POAG cases. Here we removed all glaucoma cases and 3,000 controls with IOP/VCDR measurements as well as their relatives from UKBB VCDR/IOP GWAS. We also evaluated the performance of PRS in non-European ancestry (192 cases and 6,841 controls of South Asian ancestry in UKBB). **d**, Cumulative risk of glaucoma in UKBB. For the analysis of *MYOC* p.Gln368Ter carriers ($n = 965$; cases = 72; controls = 893), participants were stratified into tertiles of PRS. We also examined cumulative risk of glaucoma in the general population (*that is* in *MYOC* p.Gln368Ter non-carriers, $n = 381{,}196$; cases = 7,381; controls = 373,815) stratifying by deciles of the PRS. The discovery and testing datasets were designed to derive the PRS with no sample overlap (Supplementary Note).

nature research

Corresponding author(s): Xikun Han

Last updated by author(s): Nov 14, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BOLT-LMM software (version 2.3): https://data.broadinstitute.org/alkesgroup/BOLT-LMM/ GCTA software (version 1.26): http://cnsgenomics.com/software/gcta/ LOCUSZOOM (version 1.4): http://locuszoom.sph.umich.edu/ LDlink: https://analysistools.nci.nih.gov/LDlink/ LD score regression software (version 1.0.0): https://github.com/bulik/ldsc MAGMA (v1.06): https://ctg.cncr.nl/software/magma METAL software (2011-03-25 release): http://csg.sph.umich.edu/abecasis/Metal/ MTAG: Multi-Trait Analysis of GWAS (version 1.0.7) https://github.com/omeed-maghzian/mtag PLINK software (version 1.90 beta): http://www.cog-genomics.org/plink2 R (version 3.4.1): https://cran.r-project.org/ More details are presented in the method section and Supplementary Note.
Data analysis	Please see above.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data used in this work were obtained from the UK Biobank (UKBB), the Australian & New Zealand Registry of Advanced Glaucoma (ANZRAG), the Blue Mountains Eye Study (BMES), the The Progression Risk Of Glaucoma: RElevant SNPs with Significant Association (PROGRESSA) study, and the National Eye Institute Glaucoma Human Genetics Collaboration Heritable Overall Operational Database (NEIGHBORHOOD).
More details are presented in the method section, Supplementary Note, and Supplementary Table 1.
UK Biobank data are available through the UK Biobank Access Management System https://www.ukbiobank.ac.uk/.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We assembled the largest possible sample size to maximize the number of novel loci. This approach was effective because we identified many novel loci for glaucoma and replicated them in independent datasets.
Data exclusions	Some samples were excluded based on genetic ancestry to ensure homogeneity. This is described in the methods.
Replication	We assembled multiple independent sample sets for glaucoma in the project. We replicated the novel loci for glaucoma and prediction findings in independent datasets.
Randomization	Samples were from collected from observational studies and were not randomized.
Blinding	Genotyping and quality control for the genetic data was conducted without knowledge of the phenotypes.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	The detailed information of population characteristics about all study cohorts can be found in Supplementary Tables 1 and Supplementary Note.
Recruitment	The detailed information of recruitment for each of the study cohorts can be found in Supplementary Note.
Ethics oversight	All participants provided informed written consent, and all study procedures were performed in accordance with the World

Ethics oversight	Medical Association Declaration of Helsinki ethical principles for medical research. More information about these study cohorts can be found in Supplementary Note.

Note that full information on the approval of the study protocol must also be provided in the manuscript.