Heritability of glaucoma and glaucoma-related endophenotypes: Systematic review and meta-analysis

Nigus Gebremedhin Asefa, MPH\textsuperscript{a} \textsuperscript{*} \textsuperscript{†}
n.g.asefa@umcg.nl, niguurayu2003@gmail.com

Anna Neustaeter, MSc.\textsuperscript{b}, \textsuperscript{†}
Nomdo M. Jansonius, MD, PhD

Harold Snieder, PhD\textsuperscript{a}

\textsuperscript{a}Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

\textsuperscript{b}Department of Ophthalmology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands

\textsuperscript{*}Corresponding author: Nigus Gebremedhin Asefa, Department of Epidemiology, University of Groningen, University Medical Center Groningen, P.O. Box 30.001, the Netherlands.

\textsuperscript{†}These authors contributed equally.

**Funding:** This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 661883. Additional funding has been provided by the Rotterdamse Stichting Blindenbelangen under grant number B20150036. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Conflict of interest:** None.
Abstract

We have systematically extracted all available heritability ($h^2$) estimates of glaucoma and related endophenotypes from the literature and summarized the evidence by meta-analysis. Glaucoma endophenotypes were classified into 10 clusters: intraocular pressure, anterior chamber size, central corneal thickness, cup-to-disc ratio, disc size, cup size, corneal hysteresis, retinal nerve fiber layer thickness, cup shape, and peripapillary atrophy. Random-effects meta-analyses were performed for each cluster. For clusters with $n \geq 10$ $h^2$ estimates, we also performed subgroup and meta-regression analyses. The literature search yielded 53 studies.

The $h^2$ of primary open-angle glaucoma ranged from 0.17 to 0.81, and was 0.65 for primary angle-closure glaucoma in a single study. The pooled endophenotype $h^2$ estimates were intraocular pressure, 0.43 (0.38–0.48); anterior chamber size, 0.67 (0.60–0.74); central corneal thickness, 0.81 (0.73–0.87); cup-to-disc ratio, 0.56 (0.44–0.68); disc size, 0.61 (0.37–0.81); cup size, 0.58 (0.35–0.78); corneal hysteresis, 0.40 (0.29–0.51); retinal nerve fiber layer thickness, 0.73 (0.42–0.91); cup shape, 0.62 (0.22–0.90); and peripapillary atrophy, 0.73 (0.70–0.75). We identified mean age, ethnicity, and study design as major sources of heterogeneity. Our results confirm the strong influence of genetic factors on glaucoma and its endophenotypes. These pooled $h^2$ estimates provide the most accurate assessment to date of the total genetic variation that can ultimately be explained by gene-finding studies.

Keywords: heritability; glaucoma; endophenotype; gene; twin study; family study
1 Background

Glaucoma, the second leading cause of irreversible blindness in the world, is asymptomatic, insidious, and is expected to affect over 100 million people by the year 2040. Glaucoma is a generalized term for a group of progressive neurodegenerative optic nerve disorders characterized by the loss of retinal ganglion cells, thinning of the retinal nerve fiber layer thickness, and visual field loss.

Several different types of glaucoma exist, with three major adult-onset variants. The first is primary open-angle glaucoma (POAG), where the anterior chamber angle remains open, but insufficient outflow of aqueous humor through the trabecular meshwork results in increased intraocular pressure (IOP), leading to optic nerve damage and visual field loss. The second variant is normal-tension glaucoma, where optic nerve damage occurs while IOP remains within normal limits. Finally, in primary angle-closure glaucoma (PACG), the increase in IOP is related to a progressive shallowing of the anterior chamber angle. These forms of glaucoma have complex inheritance where multiple genetic variants and environmental factors, in addition to their interactions, contribute to disease risk.

The total variance in a phenotype ($V_P$), such as glaucoma, can be subdivided into the genotypic variance ($V_G$) and remaining environmental variance ($V_E$). The $V_G$ can be further partitioned into 1) additive genetic variance ($A$), the sum of individual allele effects, and 2) non-additive genetic variances, the interactions between alleles at the same locus (dominance, $D$) or on different loci (epistasis). Similarly, environmental variance ($V_E$) can be broken down into 1) common environmental influences shared by family members ($C$), and 2) unique environmental influences ($E$) resulting from differences among family members as well as measurement error. The $V_P$ attributed to the additive variance ($V_A$) of the $V_G$ is termed (narrow-sense) heritability ($h^2$). Heritability estimation is an essential prerequisite for further gene-finding studies, such as genome-wide association studies (GWAS). Twin and family study designs are excellent approaches to quantify the genetic as well as shared and unique environmental factors to complex traits such as glaucoma. Traditionally, $h^2$ was estimated using correlation coefficients of relative pairs in twin and family studies. In the 1980s, a dedicated software program was developed for more powerful approaches for estimating $h^2$ via variance component or structural equation modeling, typically maximum likelihood estimation. Recent advances in genetic data analyses have enabled researchers to quantify the $h^2$ of complex diseases by taking the variance explained by all common genetic variants. This variance is called single-nucleotide polymorphism–based or chip-based $h^2$. 


Heritability estimates obtained by these different approaches vary considerably and can include statistical, methodological, or clinical heterogeneity. Heritability estimates are generally higher in twin than in family studies, and estimates for single-nucleotide polymorphism–based \( h^2 \) typically range from one-third to one-half of the total \( h^2 \) derived from twin or family studies.\(^{80,91,116}\) Furthermore, \( h^2 \) estimates in complex disorders can vary due to factors such as age and sex.\(^{83}\) In meta-analyses, meta-regression and subgroup analyses can address these sources of heterogeneity.\(^{8,49}\)

In addition to a heritable component, age is also a major risk factor for glaucoma. A recent meta-analysis reported the odds of developing POAG and PACG among Asian populations to be 1.50 and 2.18 per decade of age, respectively.\(^{20}\) Positive family history and ethnicity are two additional major risk factors for glaucoma development.\(^{2,58,102,109,126}\) There is a 20% lifetime risk for developing glaucoma in siblings of patients with POAG compared with approximately 2% in the general population.\(^{102,126}\) In terms of ethnicity, POAG affects about 2% of Dutch-Caucasians and Chinese and 6-7% of Afro-Caribbeans.\(^{28,45,64}\)

Glaucoma-related endophenotypes are also a consideration for glaucoma risk.\(^{39,42,60,65,107}\) They are a mixture of heterogeneous, continuous, heritable traits, where population-specific deviations from normality guide physicians in glaucoma diagnosis.\(^{7,22,35,39,53,85,113}\) Relevant endophenotypes for glaucoma are IOP, anterior chamber size (ACS), central corneal thickness (CCT), and optic disc excavation depicted by the cup-to-disc ratio (CDR). Reported \( h^2 \) estimates for glaucoma-related endophenotypes vary between and within studies.\(^{29,60,65}\)

A body of literature exists for the inheritance of glaucoma; however, as glaucoma is a heterogeneous condition, there is a need for a quantitative summation of the findings. A 2010 publication reviewed ocular trait heritabilities; however, the review was not systematic and did not focus on glaucoma and its endophenotypes.\(^{91}\) Therefore, we aimed to collate all available \( h^2 \) estimates of glaucoma subtypes and related endophenotypes by systematically selecting and quantitatively combining \( h^2 \) estimates from all relevant twin, family, and GWASs. More specifically, we aimed to answer the following research questions: 1) How much of the variance in glaucoma and related endophenotypes is due to genetic factors? 2) What is the proportion of variance accounted for by additive (A) and/or dominant (D) genetic influences, as well as common (C) and/or unique environment (E) factors? 3) Does \( h^2 \) vary between different age, sex, ethnic groups, study design, or data analysis method?

2 Methods

This systematic review and meta-analysis is in accordance with the Preferred Reporting Items for Systematic review and Meta-Analysis (PRISMA) 2009 checklist (Supplementary Table
The inception of this systematic review was on March 2017 and is registered with the International Prospective Register of Systematic Reviews (PROSPERO) under registration number CRD42017064504. The protocol for this systematic review and meta-analysis has previously been published.5

2.1 Data collection and risk of bias

The National Institute of Health Quality Assessment tool for Observational Cohort and Cross-Sectional Studies was used to assess the quality of the primary articles (Supplementary Table S2). Two reviewers independently assessed the methodological quality of each article using the parameters shown in Figure 1. This includes whether the research question/objective was clearly stated and whether the method of data analysis and outcome measures were clearly defined. The quality score per article was based on a total of seven questions and categorized into three classes: high (≥6), medium (4-5), and low (≤3).

To ensure all relevant data were collected per study, a data extraction form was created, containing 1) author's name; 2) title and abstract; 3) year of publication; 4) sample size; 5) type of glaucoma; 6) type of endophenotype; 7) study design; 8) data analysis method; 9) $h^2$ estimates; and 10) relevant sociodemographic variables (Supplementary Table S3).

2.2 Data synthesis and statistical analysis

Heritability estimates were weighted based on within-study sample size.27,66 Assuming that $h^2$ estimates differed between studies/population, we used a random-effects model. As glaucoma endophenotypes are pathophysiologically different, they were classified into 10 clusters: IOP, ACS, CCT, CDR, disc size, cup size, corneal hysteresis, retinal nerve fiber layer thickness, cup shape, and peripapillary atrophy. A separate meta-analysis was performed for each cluster using Meta and Metafor packages in R (version 3.3.3).114

Sex-specific $h^2$ estimates and $h^2$ estimates of different traits within the same study population were treated as independent if they were originally reported as such. Each $h^2$ estimate was transformed using the logit function.9 For ease of interpretation, the final pooled $h^2$ and 95% confidence intervals (CIs) were back-transformed. Heterogeneity between studies was quantified with Cochran's $Q$ test and the $I^2$ statistic.17,23,49 For endophenotype clusters that included 10 or more studies, publication bias was evaluated using an Egger's test and visualized with a funnel plot.32 Full details of data analyses are provided in Supplementary Material S4. We also performed trim and fill analyses for clusters showing publication bias.31
2.3 Outcome measure
The outcome measures of this systematic review and meta-analysis were estimates of $h^2$, reported as a percentage or proportion.

2.4 Subgroup analysis
Subgroup analyses were conducted for endophenotype clusters with $n \geq 10$ $h^2$ estimates, to assess factors that could explain variation in $h^2$. These analyses were implemented based on 1) mean age; 2) ethnicity; 3) study design; 4) data analysis method; 5) the number of variables controlled for confounding; and 6) methodological quality score.

A meta-analysis of 50 population-based studies suggested that after the age of 50, glaucoma (POAG and PACG) prevalence increases exponentially. Thus, mean age was stratified into two categories, <50 and $\geq$50 years. When studies reported median age, the mean age was re-estimated using the formula proposed by Hozo and coworkers. When studies did not report mean age or used a range of ages and reported a single $h^2$ estimate, we used the calculated midpoint. We categorized ethnicity into four subgroups: White European ancestry, East Asians, Black African ancestry, as well as “Mixed” which was classified when an $h^2$ estimate was reported for a mix of two or more ethnic groups. The study design was divided into twin, family, and GWAS designs. Data analysis methods were subgrouped into four types: threshold liability model, which accounts for the trait prevalence, correlational model, where correlation coefficients were used, linear regression model, where estimates were derived from the slope of parent-offspring regression, and the variance components method, where the maximum likelihood approach was used. For IOP, $h^2$ estimates were further subgrouped by the device type used as some equipment accounts for CCT (which is highly heritable) when measuring IOP while other devices do not.

2.5 Meta-regression analyses
For endophenotype clusters with $n \geq 10$ $h^2$ estimates, a univariable meta-regression was performed to explore potential within-cluster sources of heterogeneity. The relationships between reported $h^2$ and 1) mean age (in both continuous and categorical forms); 2) ethnicity; 3) study design; 4) method of data analysis; 5) number of variables controlled for confounding; and 6) methodological quality score were assessed. For variables that did not meet linear regression assumptions, we used a nonparametric estimator in R. For all tests, the significance threshold was set to $P<0.05$.

2.6 Sensitivity analysis
A Baujat plot was used to identify the top three studies contributing to the heterogeneity within each cluster. Subsequently, sensitivity analyses were carried out by excluding these
studies. Additional sensitivity analyses were conducted by excluding studies that 1) included children (≤12 years of age); 2) used a GWAS design; or 3) estimated $h^2$ from Black African ethnicity.

3 Results

3.1 Search results

The total number of articles retrieved and the final number of eligible studies are summarized in Figure 2. The literature search of four electronic databases yielded 1,170 articles, and the reference list search resulted in five more articles. The agreement between the two authors was 0.75 (95% CI: 0.68-0.82). After abstract review, the number of studies suitable for full review was 70. According to the selection criteria, 53 full-text primary articles were included in this meta-analysis; the characteristics of these studies are summarized in Supplementary Table S5. In total, 127 $h^2$ estimates were obtained, approximately two estimates per study (Supplementary Table S5).

More than half of the $h^2$ estimates, 76 of 127 (60%), were derived by maximum likelihood variance component methods, and 11 of 127 (9%) were estimated using the threshold liability model (Supplementary Table S5). Approximately half of the $h^2$ estimates, 62 of 127 (49%), were from family studies, and 46 of 127 (36%) were from twin studies.

3.2 Quality assessment and risk of bias

The majority of included articles, 49 of 53 (92%), specified the study population; however, no study described a power analysis, and only seven of 53 (13%) reported a random sampling of study participants from a defined source population.21,26,54,60,72,104,112 One study did not describe participants' age,62 two studies did not define an inclusion criteria clearly,58,93 and 11 studies did not control for sufficient (i.e., at least two) confounding variables. The majority of the included studies, 40 of 53 (75%), were rated as medium quality, and 5 of 53 (9%) were rated as high quality.

3.3 Age and sex of participants

Within the included studies, mean age ranged from 10.8 to 75 years. The majority of $h^2$ estimates, 102 of 127 (80%), were pooled for both sexes, while 25 of 127 (20%) estimated sex-specific estimates (n = 15 for females, n = 10 for males). Due to the low number of estimates, the effect of sex on $h^2$ could not be computed. Sample size varied widely; the largest study included 481,657 individuals with 128,989 families,121 while the smallest study included 94 individuals using a family study.39 The total number of individuals was 687,234 when correcting for two or more $h^2$ estimates per study. Approximately two-thirds, 87 of 127 (69%), of the $h^2$ estimates were
carried out among White European ancestry, followed by East Asians (24 of 127; 19%) (Supplementary Table S5).

3.4 Glaucoma heritability

There was a paucity of articles reporting \( h^2 \) for glaucoma (\( n = 6 \)), two for “unspecified” glaucoma,\(^ {41,121} \) three studies for POAG,\(^ {22,26,106} \) and one for PACG.\(^ {122} \) In the UK Biobank, Ge and coworkers\(^ {41,121} \) used self-reported glaucoma, whereas in a large US study, the authors used insurance claim data based on the International Classification of Diseases (ICD-9) diagnostic codes.\(^ {41,121} \) Charlesworth and coworkers\(^ {22} \) defined POAG as “an optic neuropathy that exhibits optic nerve head excavation with thinning of the neuroretinal rim, often with Drance-type nerve fiber layer hemorrhages; notching; pitting; significant focal loss; or general loss of retinal fiber layer (generally measured by an enlarged vertical CDR \( \geq 0.7 \)); as well as visual field defects consistent with the disc changes and with common descriptions of glaucomatous field loss.” In contrast, in the study by Cuellar-Partida and coworkers, neuroretinal rim and nerve fiber layer were not part of the inclusion criteria and individuals with CDR >0.95 were included.\(^ {26} \) Finally, Teikari used the registry of right for free medication, which covers all Finnish citizens, including 30,000 patients receiving free medication for glaucoma. All diagnosis codes are reviewed by a board of the Social Insurance Institution. Hospital records of all patients in this twin study were received and carefully checked by the author for mistakes in diagnostic coding the diagnosis.\(^ {106} \)

The study by Teikari\(^ {106} \) used data from 29 monozygotic and 79 dizygotic twin pairs to obtain a POAG estimate \( h^2 \) of 0.13. The author erroneously calculated \( h^2 \) as twice the difference of the pairwise concordance ratio between monozygotic and monozygotic twins. When re-analyzed with the threshold liability model, a more appropriate approach which takes disease prevalence into account,\(^ {73} \) the \( h^2 \) (95% CI) was 0.71 (0.46–0.86), with AE as the best-fitting variance component model. Moreover, Charlesworth and coworkers\(^ {22} \) assessed POAG among 648 family members and obtained an \( h^2 \) estimate of 0.81 (0.64–0.98). Cuellar-Partida and coworkers published a GWAS of sex-specific POAG \( h^2 \), 0.52 (0.36–0.68) for females and 0.66 (0.47–0.85) for males.\(^ {26} \) Wang and coworkers\(^ {122} \) calculated PACG \( h^2 \) in a family-based study of 100 patients and 552 first-degree relatives, with an estimate of 0.65. Finally, the \( h^2 \) estimates of “unspecified” glaucoma from a large US and UK family data were 0.70 (0.67–0.83)\(^ {121} \) and 0.26 (0.15–0.37),\(^ {41} \) respectively.

The aforementioned studies defined glaucoma in different ways and used different approaches to estimate \( h^2 \). Due to these issues, data could not be pooled, and a meta-analysis of glaucoma \( h^2 \) could not be performed. Consequently, subgroup and sensitivity analyses could not be performed for glaucoma \( h^2 \).
3.5 Endophenotypes

The most commonly reported endophenotypes were IOP (n = 41), ACS (n = 26), CCT (n = 15), and CDR (n = 11) (Supplementary Table S5). Conversely, there was only one $h^2$ estimate reported for the following endophenotypes: iris thickness, ocular pulse amplitude, pulsatility of blood flow, and ganglion cell complex thickness.

3.5.1 Pooled $h^2$ estimates

Using the CCT cluster as an example, the pooled $h^2$ is depicted in a forest plot (Fig. 3). The pooled $h^2$ estimates of the 10 endophenotype clusters are shown in Figure 4. CCT studies reported the highest pooled $h^2$ estimate at 0.81 (0.73-0.87) and the lowest for corneal hysteresis at 0.40 (0.29-0.51) (Fig. 4).

The degree of heterogeneity ($I^2$) between studies was high, >75%, for nine of the 10 endophenotype clusters, and was 58% for peripapillary atrophy. Funnel plots contained asymmetry for both IOP and ACS (Supplementary Figure S6). The Egger's test results indicated the presence of publication bias for IOP and ACS, $P < 0.001$ and $P = 0.001$, respectively (Supplementary Figure S6). After adjusting for publication bias, the pooled $h^2$ estimates remained essentially unchanged, 0.43 (0.38–0.48) before correction to 0.44 (0.39–0.49) after correction for IOP, and 0.67 (0.60–0.74) to 0.69 (0.62–0.76) for ACS.

With the $h^2$ estimates derived from variance components modeling, the AE model was the best fit for most estimates, 34 of 42 (81%), followed by the ACE model, 5 of 42 (12%). Figure 5 shows the proportion of variance explained by A, D, C, and E effect per endophenotype cluster. For IOP, the pooled variance explained by A ($h^2$) was 0.58 (0.49–0.66), and within the ACS cluster, D explained 0.03 (0.01-0.09) of the variance. Similarly, C explained 0.04 and 0.01 of the variances in CDR and cup size, respectively (Fig. 5).

The most frequently reported endophenotype was IOP, and the overall pooled $h^2$ was 0.43 (0.38–0.48). Subgroup analyses suggest heterogeneity within age, ethnicity, study design, and method of data analysis. Pooled $h^2$ estimates for twin studies were higher, 0.63 (0.56–0.70), than those for both family studies and GWASs, 0.39 (0.34–0.44) and 0.26 (0.18–0.36), respectively, $P < 0.001$ (Table 1). In the subgroup analyses, the highest estimates were obtained from the variance component methods, 0.50 (0.43–0.56; $P < 0.001$). No significant differences in $h^2$ were identified between device type: 0.43 (0).

There were 26 $h^2$ estimates in the ACS cluster, with an overall $h^2$ estimate of 0.67 (0.60–0.74). When comparing age subgroups, older people (≥50 years) had an estimate of 0.53 (0.39-0.68) compared with 0.73 (0.66-0.79; $P = 0.015$) for those aged <50 years (Table 1). Ethnicity
was a source of heterogeneity for ACS; the highest $h^2$ was found among East Asians at 0.82 (0.75–0.87) (Table 1).

In the ACS cluster, twin-based $h^2$ estimates were significantly higher, 0.79 (0.72–0.84), than 0.56 (0.44–0.67; $P < 0.001$) in family studies. The $h^2$ estimate from variance component estimation was 0.69 (0.59–0.77), which is significantly higher than that from the correlation coefficient method, 0.48 (0.32–0.64), $P = 0.041$. There was no significant difference in $h^2$ estimates between low- and medium-quality studies (Table 1).

In total, there were 15 $h^2$ estimates for the CCT cluster, with an overall $h^2$ of 0.81 (0.73–0.87). Subgroup analyses indicated that CCT $h^2$ estimates differ as a function of mean age, study design, and the number of variables controlled for confounding. Heritability for CCT was higher for younger participants (aged <50 years), at 0.86 (0.81–0.90), than for those aged ≥50 years, 0.70 (0.59–0.79; $P = 0.006$). Again, twin studies reported higher $h^2$ estimates, 0.89 (0.85–0.92; $P < 0.001$), than corresponding family studies, 0.67 (0.56–0.77; $P < 0.001$) (Table 1).

The CDR endophenotype cluster, consisting of both vertical and horizontal CDRs, as well as rim area-to-disc area ratio, contained a total of 11 $h^2$ estimates. The pooled $h^2$ estimate was 0.56 (0.44–0.68). Subsequent subgroup analyses showed no significant sources of heterogeneity (Table 1). Similar to IOP, CDR measurements are also affected by the type of device used; however, it was not possible to model this effect; all CDR estimates were based on some form of fundus photography, mostly simultaneous or non-simultaneous stereo photography.

### 3.5.2 Meta-regression analysis

The meta-regression further explored previously identified sources of heterogeneity within the four major endophenotype clusters (IOP, ACS, CCT, and CDR). Age, <50 compared to ≥50 years, was significantly associated with $h^2$ estimates in both IOP and CCT. The $h^2$ of CCT in people younger than 50 years was on average 0.14 (0.01–0.27) higher than in people aged ≥50 years (Table 2). Modeling age as a continuous variable also supported the inverse relationship between $h^2$ and age, yielding a significant decrease in $h^2$ with age for IOP and CCT. The decrease in $h^2$ of IOP for each unit increase in age (years) was 0.007 (0.006–0.009) (Table 2; Fig. 6A). Moreover, the age-based meta-regression was extended to visualize the relationship between the combined $h^2$ of all ten endophenotype clusters, seen in Figure 6B. For CCT, the $h^2$ estimates were higher by 0.16 (0.03–0.30) for ≤2 confounding variables than those when controlling for more than two variables (Table 2).

Ethnicity was a significant predictor of $h^2$ for ACS; the variation explained by genetic factors in East Asians was 0.18 (0.11-0.26), which is higher than that in White Europeans (Table 2). Heritability was 0.16 to 0.22 higher in twin studies than in family, or the combination of
family and GWA studies. These results indicate that study design is a significant source of heterogeneity for IOP, ACS, and CCT (Table 2, Fig. 3). The method of data analysis was only significant for IOP; $h^2$ estimates obtained from correlation coefficients were 0.21 (0.05–0.36) less than those from variance component methods.

3.5.3 Sensitivity analysis

Baujat plot was used to visualize each study's contribution to the overall heterogeneity. The three most heterogeneous studies per cluster were as follows: for IOP, $h^2 = 0.14, 0.30, 0.30$ and 0.29 $^{21}$; for ACS, $h^2 = 0.08, 0.04, 0.926^{112}$; for CCT, $h^2 = 0.95, 0.47, 0.91^{130}$; and for CDR, $h^2 = 0.79, 0.31, 0.98$ and 0.43 $^{21}$ When compared to the initial model, the pooled $h^2$ estimates of all four clusters did not change significantly after removal of the most heterogeneous studies. We observed discrepancies between pre- and post-sensitivity and post-sensitivity analyses were observed in CCT; age, (in both categorical and continuous forms) which was a source of heterogeneity, failed to show a significant effect in the post-sensitivity analysis (Table 3).

Finally, when compared to the full model, no significant differences were identified after performing sensitivity analyses, when excluding GWASs (CDR, n = 1), studies with Black African ethnicity (IOP, n = 1; CDR, n = 1), or studies that included children (aged ≤12 years) (IOP, n = 1; ACS, n = 6; CCT, n = 1; CDR, n = 1).

4 Discussion

Our findings indicate that glaucoma is heritable and the $h^2$ ranges from 0.17 to 0.81 for POAG. Similarly, glaucoma-related endophenotypes demonstrate moderate to high $h^2$, ranging from 0.40 in corneal hysteresis to 0.81 in CCT. For studies that used variance components models, the $AE$ model was generally the best-fitting model, suggesting high proportions of variance attributable to additive genetic and unique environmental effects.

Through subgroup and meta-regression analyses, factors related to within-study setup, including mean age (higher $h^2$ in younger people), ethnicity (higher $h^2$ of ACS in East Asians), study design (higher $h^2$ in twin-based studies), as well as method of data analyses (higher $h^2$ for IOP obtained from the variance component) were major sources of heterogeneity. Heritability estimates for the direct glaucoma phenotype were limited; three studies for POAG, one for PACG, and two for “unspecified” glaucoma. As such, exploration of possible sources of heterogeneity was not possible. Teikari estimated POAG $h^2$ at 0.13 $^{106}$; however, the main reason for this low estimate is that the statistical approach was not adequate for the type of data collected. We re-estimated the $h^2$ as 0.71 (0.46–0.86) via the threshold liability model, $^{34}$ which is much closer to the estimate obtained by Charlesworth and coworkers for POAG (0.81). $^{22}$
Although the 95% CI overlapped, the (chip based) $h^2$ of POAG for males (0.66) was higher than that in females (0.52). Interestingly, previous meta-analysis studies suggest higher POAG prevalence in men than in women, even after adjusting for variables such as age. The effect of gender diminishes at high age; evidence from the Rotterdam cohort suggested that there were no sex differences in glaucomatous visual field loss at age $\geq$55 years. It is important to note that $h^2$ estimates do not automatically transfer into disease risk, as the environment can be changed or manipulated. Whether there is an actual effect of sex on the $h^2$ of glaucoma has yet to be determined in large-scale studies.

The most reported cluster was IOP, with a pooled estimate of 0.43 (0.38–0.48). The large number of studies for this endophenotype is logical, as IOP is currently the only modifiable risk factor associated with glaucoma. A recent GWAS has found more than 100 genetic loci implicated with glaucoma risk. However, the variance of IOP explained by these single-nucleotide polymorphisms was only 9% in the UK Biobank and 17% in the EPIC-Norfolk cohort. IOP is affected by different factors including time and month of examination, type of tonometer, and the eye being measured. In other studies, IOP was significantly associated with diabetes status, body mass index, refractive error, cholesterol, pulse, smoking, and diastolic blood pressure in Caucasian participants. The low pooled $h^2$ could also be partly due to the fundamental variability in IOP measurements, as the environmental ($E$) component encompasses measurement errors; so the greater the measurement error, the more $h^2$ could be underestimated.

ACS was the second most studied endophenotype, with a pooled $h^2$ estimate of 0.67 (0.60–0.74). This endophenotype has a well-established connection with PACG. Nongpiur and coworkers, in their two-stage GWAS, reported the ABCC5 gene is significantly associated with ACS and mediates the risk of PACG among Asian descent population. The variation of PACG explained by the ABCC5 gene was 0.35%.

Among all endophenotype clusters, CCT was the most heritable at 0.81 (0.73–0.87), (Fig. 4). Our findings are higher than a previously reported pooled estimate ($h^2 = 0.71$), for all human traits classified under the ophthalmological domain. One reason for this could be the large number of twin-based studies (53%) for CCT, which typically yield higher estimates than family studies. GWASs have identified several loci associated with CCT in European, Australian, and Asian populations; however, the variation explained by these loci was low (0.8% in Chinese to 8.3% in European populations). Additional gene-finding studies are needed to uncover the origin of the high $h^2$ of CCT.

The endophenotypes cup and disc size are physiologically correlated, and the cup-to-disc ratio is used for clinical assessment of optic nerve damage due to glaucoma. Increased CDRs
indicate glaucomatous damage, that is, loss of retinal ganglion cells.\textsuperscript{105} However, the reduction of retinal ganglion cells could also arise from the nonpathological aging process.\textsuperscript{43} CDR showed moderate $h^2$ (0.56 (0.44–0.68)), with substantial variability, which may be due to the wide biological variation of cup size in the general population.\textsuperscript{40} GWASs have investigated the genetic component of CDR, and several associated loci have been found so far; two in Latinos,\textsuperscript{79} two in European,\textsuperscript{85} and 10 in European and Asian mixed populations.\textsuperscript{85,97,98}

Meta-regression analysis further identified the influence of genetic and environmental factors at different ages; people under 50 years of age had 0.14–0.15 higher $h^2$ estimates than people aged $\geq$ 50 years in the four endophenotype clusters. For each year increase in age, $h^2$ decreased 0.004–0.006, indicating glaucoma-related endophenotypes are increasingly influenced by environmental factors later in life. These findings reflect work done in the UK Biobank, where IOP and corneal hysteresis had decreasing heritability over age,\textsuperscript{41} an average decrease of 0.004 per year. Our findings are similar to studies addressing body mass index; $h^2$ for body mass index was higher in children.\textsuperscript{33,77} A meta-analysis of endurance-related phenotypes also reported decreasing $h^2$ with increasing age.\textsuperscript{78} This relationship could be explained through increased cumulative environmental exposure with advancing age, including smoking, drinking, chemical exposure, and diet.\textsuperscript{33,77} McCartney and coworkers proposed that as twins grow, their independent life experiences increase the unique environmental variances, and thus decreases genetic variances.\textsuperscript{75} However, other studies reported constant $h^2$ over time: the $h^2$ of blood pressure,\textsuperscript{51} grip strength, and well-being\textsuperscript{36} remained the same across different age groups. Alternatively, meta-analyses of social behaviors, including, intelligence quotient, anxiety symptoms, social attitudes, and gambling behavior, reported increasing heritability throughout life,\textsuperscript{11,128} which may be explained by the decrease of environmental factors later in life. These contradicting findings suggest that $h^2$-age relationship is determined by the type of phenotype, environmental factors, and interactions between them.

The $h^2$ of ACS was higher in East Asians than that in the White European population (Table 2). Visscher and coworkers\textsuperscript{116} suggested the $h^2$ in phenotypes may differ between populations. Genetic influence depends on allele frequencies and effect sizes, as well as environmental factors, which vary across different populations. This was illustrated by ACS $h^2$ that was higher in the East Asian population in this study. Shallow anterior chamber depth is a major risk factor for PACG,\textsuperscript{6,94} and several studies reported that the anterior chamber depth is more shallow in East Asians than in Europeans or Africans.\textsuperscript{24,38,46} A recent GWAS identified genes that may increase the risk of having PACG by affecting the anterior chamber dimensions of the eye.\textsuperscript{81}
Meta-regressions of IOP, ACS, and CCT clusters suggest study design related differences in $h^2$; twin-based estimates were, in general, higher than estimates from family or GWAS-based studies (Table 2). Our finding reflects previous body mass index and personality meta-analysis studies, reporting significantly higher $h^2$ estimates in twin-based studies than family-based studies.\textsuperscript{77,120} An explanation for this could be the age differences within family-based studies, as opposed to twin-based studies. Longitudinal twin studies for other complex traits, such as C-reactive protein\textsuperscript{92} and blood pressure,\textsuperscript{44} report that partly different (sets of) genes influence these traits at different ages. Based on family studies, this age difference would reduce the correlations between pairs of relatives with different ages and their corresponding $h^2$ estimates.

Some authors have argued the validity of contested assumptions of twin studies, that is, over-reliance of the twin method on the equal environment assumption, which states that both monozygotic and monozygotic twins share the environment to the same extent.\textsuperscript{82} Previous psychiatric literature suggested that different treatment effects from parents might partly explain twin correlations.\textsuperscript{88} Alternatively, researchers have compared $h^2$ estimates of cardiac conduction traits via the classical twin modelling and genome-wide genetic relatedness and found no significant difference.\textsuperscript{80} Additional studies are required to determine the validity of equal environment assumptions in ocular traits.

For IOP, $h^2$ estimates derived from variance component data analyses were significantly higher than those from correlation coefficients. This result is in agreement with a previous meta-analysis by Sanfilippo and coworkers.\textsuperscript{91} Within a simulation study of animal body mass $h^2$, where parent-offspring regression was compared to the maximum likelihood variance approach, results indicate the maximum likelihood variance approach is, in general, more precise in estimating trait $h^2$.\textsuperscript{115} The variance component estimation method uses full pedigree information, which allows explicit modeling of environmental effects shared by individuals. This variance component technique is more efficient for estimating $h^2$ in populations with many relationships and unbalanced family sizes.\textsuperscript{1,116} Results of sensitivity analyses were similar to the original full-model analyses, with overlapping CIs. This similarity indicates that our results are robust and not affected by exceedingly low or high $h^2$ estimates.

The heritable aspect of glaucoma and its related endophenotypes has already led to successful gene-finding studies.\textsuperscript{124} Future multi-trait GWASs that combine different endophenotypes will likely improve power and give rise to the identification of more genetic loci associated with glaucoma.\textsuperscript{111} The identified genetic markers can be used to quantify the genetic risk of an individual, via developing genetic risk scores for glaucoma.
Glaucoma often affects more than one member of a family. The high heritability, as found in the present study, confirms this clinical observation and reinforces the importance of inviting seemingly unaffected relatives of glaucoma patients as part of case finding. Case finding remains the cornerstone of glaucoma detection as long as population-based glaucoma screening cannot be conducted cost-effectively. Screening could be made more efficient by focusing on high-risk groups. However, currently known (nongenetic) risk factors are not strong enough to allow for screening limited to high-risk groups, as too many cases would be overlooked. Similarly, the currently known genes (from GWAS and other techniques) do not explain the observed heritability completely, and as such, screening based on genes is currently considered premature as well. Knowledge on heritability drives further gene finding (aiming to annihilate the discrepancy between heritability and variance explained by known genes), ultimately aiming for efficient screening and personalized health care based on genetic profiles.

5 Strengths and limitations

A unique feature and major strength of this current meta-analysis is its comprehensive and systematic approach in the summation of \( h^2 \) estimates for both glaucoma and related endophenotypes. The inclusion of \( h^2 \) estimates for ten endophenotype clusters for glaucoma is a novel approach; however, \( h^2 \) estimates in most of the included studies varied widely in both \( h^2 \) estimate and sample size, which led to significant heterogeneity. Although publication bias was no longer statistically significant in both IOP and ACS endophenotype clusters after the trim and fill analysis (Supplementary Figure S6), some asymmetry remained. Egger and coworkers suggested that, in addition to publication bias, there may be other sources of funnel plot asymmetry, including poor methodological design, true heterogeneity between studies, and sampling variation.

Our analysis showed that the \( AE \) model was generally the best-fitting model, suggesting that the largest proportions of variance are attributed to additive genetic and unique environmental effects. However, by design, family and twin studies do not allow identification of the specific unique environmental factors affecting the total phenotypic variance.

Some endophenotypes included in this study are associated with other eye diseases and not specific to glaucoma. For example, IOP and CCT are correlated to corneal arcus, ACS to myopia, and CDR to age-related macular degeneration. Further genetic studies are needed to address the question of how much of the genetic risk of each endophenotype translates to the overall genetic risk of glaucoma.

Most of the studies were from White Caucasians and East Asian populations. Due to power considerations, we used univariate meta-regression analysis; hence, the results should be interpreted within the context of these limitations.
6 Conclusion
The pooled \( h^2 \) estimates of glaucoma and glaucoma-related endophenotypes in this systematic review and meta-analysis provide the most accurate estimates to date of the total genetic variation that can ultimately be explained by gene-finding studies.

7 Search strategy and inclusion criteria

7.1 Method of literature search
Using both Medical Subject Heading (MeSH) terms and free-text words, we conducted an initial limited search in MEDLINE. Then, with identified keywords and index terms, a systematic search was performed in MEDLINE, EMBASE, Web of Science, and ScienceDirect (Table 4). The search was restricted to articles written in the English language or articles that included English abstracts. Authors were requested to send the full text when only the abstract could be obtained. All retrieved articles were exported to RefWorks (version 3.0; Refworks, Bethesda, MD) and duplicates were manually removed.

We developed a database for screening relevant articles according to predefined selection criteria. Two reviewers (N.G.A. and A.N.) independently evaluated and then compared selected abstracts for eligibility. The percentage of agreement for article inclusion was calculated via Cohen's kappa. Any disagreement about article eligibility was resolved through reassessment, followed by discussion, and third-party consultation when required. Finally, a supplementary Google scholar search was conducted to retrieve potential relevant articles from reference lists.

7.2 Inclusion criteria
We included original peer-reviewed studies up until December 31, 2017. Studies were included if 1) the participants were twins, families, or unrelated individuals (population); 2) authors reported \( h^2 \) of glaucoma or glaucoma-related endophenotypes or where \( h^2 \) could be re-estimated from intraclass correlations or linear coefficients (outcome); and 3) an observational or cross-sectional study design was used (study design).

We excluded studies if 1) whole articles were not in English and 2) if the genetic variance was estimated only from significant single-nucleotide polymorphisms or genetic loci within a GWAS. The number of articles obtained, reviewed, and excluded is summarized using the PRISMA flow chart (Fig. 2).

8 References


92. A.A. Sas, A. Vaez, Y. Jamshidi, et al., Genetic and environmental influences on stability and change in baseline levels of C-reactive protein: A longitudinal twin study, Atherosclerosis 265, 2017, 172–178.


128. Y.-H. Xuan, S. Li, R. Tao, et al., Genetic and environmental Influences on gambling: A meta-analysis of twin studies, Front Psychol 8, 2017, 2121.

**Other cited material**