



OPEN

## Causal factors in primary open angle glaucoma: a phenome-wide Mendelian randomisation study

Thomas H. Julian<sup>1,2</sup>, Zain Girach<sup>3</sup>, Eleanor Sanderson<sup>4</sup>, Hui Guo<sup>5</sup>, Jonathan Yu<sup>2</sup>, Johnathan Cooper-Knock<sup>6</sup>, Graeme C. Black<sup>1,7</sup> & Panagiotis I. Sergouniotis<sup>1,2,7,8</sup>✉

Primary open angle glaucoma (POAG) is a chronic, adult-onset optic neuropathy associated with characteristic optic disc and/or visual field changes. With a view to identifying modifiable risk factors for this common neurodegenerative condition, we performed a 'phenome-wide' univariable Mendelian randomisation (MR) study that involved analysing the relationship between 9661 traits and POAG. Utilised analytical approaches included weighted mode based estimation, the weighted median method, the MR Egger method and the inverse variance weighted (IVW) approach. Eleven traits related to POAG risk were identified including: serum levels of the angiotensin-converting enzyme 2 (ACE2) (OR [odds ratio] = 1.11, IVW  $p = 2.34E-06$ ) and the cadherin 5 protein (OR = 1.06, IVW  $p = 1.31E-06$ ); intraocular pressure (OR = 2.46–3.79, IVW  $p = 8.94E-44$ – $3.00E-27$ ); diabetes (OR = 5.17, beta = 1.64, IVW  $p = 9.68E-04$ ); and waist circumference (OR = 0.79, IVW  $p = 1.66E-05$ ). Future research focussing on the effects of adiposity, cadherin 5 and angiotensin-converting enzyme 2 on POAG development and progression is expected to provide key insights that might inform the provision of lifestyle modification advice and/or the development of novel therapies.

Glaucoma is a heterogeneous group of disorders characterised by progressive retinal ganglion cell degeneration and distinctive optic disc changes<sup>1,2</sup>. A key clinical consequence of glaucoma is progressive visual field loss that can affect central vision and lead to irreversible blindness<sup>3</sup>. The most common glaucoma subtype is primary open angle glaucoma (POAG), a chronic optic neuropathy associated with open iridocorneal angles and an apparent absence of unequivocal secondary causes<sup>4</sup>. In 2020, the global adult population of POAG sufferers was estimated to stand at 53 million; this is projected to increase to around 80 million affected individuals by 2040<sup>5</sup>.

Major risk factors for POAG include increasing age and raised intraocular pressure (IOP)<sup>6</sup>. Several genetic contributors have also been described including 127 statistically independent single nucleotide variants (SNVs) associated with glaucoma at a genome-wide significant level, indicating potential contributions of several genes to POAG pathogenesis (e.g. *SVEP1*, *RERE*, *VCAM1*, *ZNF638*, *CLIC5*, *SLC2A12*, *YAP1*, *MXRA5*, *SMAD6* and *MYO7A*)<sup>7,8</sup>.

Whilst several strategies exist to reduce glaucoma progression, including surgical, laser and topical drug-based approaches, these largely target a single underlying mechanism: raised IOP. Notably, a proportion of individuals with POAG have normal IOP, while a proportion of individuals with high IOP never develop glaucoma. It is therefore widely accepted that additional pathophysiological processes are contributing to the development of POAG. Identifying such IOP-independent mechanisms has the potential to inform complementary therapeutic approaches for this common cause of visual impairment. Further, accruing evidence on the impact of lifestyle factors and interventions on POAG progression is expected to give affected individuals the opportunity to personally impact their prognosis through behavioural change. It is noted that reversing the optic nerve damage that

<sup>1</sup>Division of Evolution, Infection and Genomics, School of Biological Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. <sup>2</sup>Manchester Royal Eye Hospital, Manchester University NHS Foundation Trust, Manchester, UK. <sup>3</sup>Sheffield Medical School, University of Sheffield, Sheffield, UK. <sup>4</sup>MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. <sup>5</sup>Centre for Biostatistics, Division of Population Health, Health Services Research and Primary Care, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. <sup>6</sup>Department of Neuroscience, Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK. <sup>7</sup>Manchester Centre for Genomic Medicine, Saint Mary's Hospital, Manchester University NHS Foundation Trust, Manchester, UK. <sup>8</sup>European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Wellcome Genome Campus, Cambridge, UK. ✉email: panagiotis.sergouniotis@manchester.ac.uk

defines glaucoma is presently not possible and, as such, identifying novel strategies to prevent the development and/or to modify the progression of the disease remains of great interest.

Mendelian randomization (MR) is a statistical approach that uses genetic variation to study the causal relationships between *exposures* (such as body mass index) and medically-relevant *outcomes* (such as risk of a specific disease)<sup>9,10</sup>. MR offers advantages over conventional epidemiological studies due to the ability of this method to obtain causal effect estimates that are not biased by confounding and reverse causation (which occurs when the outcome precedes and causes the exposure instead of the other way around). In MR, genetic variants in the form of SNVs, are exploited as instrumental variables which can be utilised to study the causal relationship between an exposure variable and an outcome of interest. MR principles are based on Mendel's laws of segregation and independent assortment, which state that offspring inherit SNVs randomly from their parents and randomly with respect to most other locations in the genome. As a result, genetic variants that are associated with an exposure of interest can be used to proxy the part of the exposure variable that is independent of possible confounding influences (e.g. from the environment and other traits). Further, because genetic variation is necessarily 'upstream' of disease and environmental exposures in terms of mechanism, the risk of reverse causation is dramatically reduced<sup>9</sup>.

To date, the use of MR approaches in the context of POAG has been relatively limited. Despite this, these methods have been successfully implemented to explore the relationship between POAG and a small number of traits including brain alterations<sup>11</sup>, plasma lipid levels<sup>12</sup>, fatty acids<sup>13</sup>, type 2 diabetes<sup>14</sup>, HbA1c levels<sup>14</sup>, fasting glucose, coffee consumption<sup>15</sup> and blood pressure<sup>16</sup>.

In this study, we discuss the findings of a systematic, broad ('phenome-wide') univariable MR analysis and report a set of traits that show evidence of a causal effect on POAG. Our method points to a number of previously unreported, potential causal factors and provided further evidence on the role of known glaucoma risk factors.

## Results

Overall, 9,661 traits were eligible for analysis. Of these, 11 were identified to have statistically robust relationships with POAG in a univariable analysis (Table 1; Supplementary table 1.1). It is noted that 264 traits were related to POAG in a false discovery rate (FDR) corrected inverse variance weighted (IVW) and leave one out analysis; however, there was no agreement about the statistical significance of these signals among all MR methods used (suggesting weaker evidence of a causal effect; Supplementary table 1.2). The results for all traits considered are presented in Supplementary table 1.3.

Serving as a positive control, IOP was identified to represent a causal factor for POAG. Both right (OR = 2.46, IVW  $p = 1.70E-36$ ) and left (OR = 2.46, IVW  $p = 8.94E-44$ ) eye Goldman-correlated measures calculated using a Reichert ocular response analyser (ORA) were demonstrated to be causal for POAG. Additionally, ORA corneal compensated IOPs (that were measured in the same study) were also significantly related to POAG risk (left eye OR = 3.79, IVW  $p = 1.86E-30$ ; right eye OR = 3.62, IVW  $p = 3.85E-37$ ). This is in keeping with our existing understanding of POAG and validates our methodology.

Waist circumference, a measure of abdominal obesity, was inversely related to POAG risk (OR = 0.79, IVW  $p = 1.66E-05$ , weighted media  $p = 0.01$ , MR Egger  $p = 0.02$ , weighted mode  $p = 0.004$ ). Whilst body mass index (BMI) was significantly related to POAG risk in IVW (OR = 0.84,  $p = 1.93E-05$ ), this did not persist throughout

Trait	nSNV	OR	FDR IVW $p$ value	MRE IVW $p$ value	Weighted Median $p$ value	MR egger $p$ value	Weighted mode $p$ value
Intra-ocular pressure, Goldmann-correlated (left)	53	2.47	2.88E-40	8.94E-44	1.76E-19	1.79E-04	7.97E-05
Intra-ocular pressure, corneal-compensated (right)	27	3.62	9.30E-34	3.85E-37	2.14E-18	3.15E-02	3.75E-05
Intra-ocular pressure, Goldmann-correlated (right)	54	2.46	3.29E-33	1.70E-36	2.20E-23	4.85E-04	2.29E-05
Intra-ocular pressure, corneal-compensated (left)	16	3.79	3.00E-27	1.86E-30	1.98E-10	6.53E-03	1.26E-03
Treatment speciality of consultant: Urology	7	2.39E+03*	3.75E-07	4.27E-10	4.68E-05	4.06E-02	1.95E-02
Cadherin-5	15	1.06	4.08E-04	1.31E-06	4.45E-04	1.18E-02	5.11E-03
Family history of Alzheimer's disease	8	0.88*	4.77E-04	1.63E-06	3.58E-04	2.52E-02	1.09E-02
Angiotensin-1 receptor, soluble	10	1.11	6.11E-04	2.34E-06	1.95E-04	7.38E-03	1.56E-02
Waist circumference	283	0.79	2.77E-03	1.66E-05	1.04E-02	1.55E-02	4.41E-03
Squamous cell lung cancer	13	0.92*	1.01E-02	1.35E-04	3.30E-03	2.04E-02	3.28E-02
Diabetes diagnosed by doctor	48	5.17*	3.87E-02	9.68E-04	8.02E-03	4.28E-03	2.54E-02

**Table 1.** Traits passing the criteria applied within this study for causal inference. OR: odds ratio, nSNV: number of single nucleotide variants, FDR: false discovery rate, IVW: inverse variance weighted, MRE: multiplicative random effects, \*: binary exposure trait, odds ratio can be used to indicate protective/causal effect but does not represent an accurate estimate of the magnitude of effect. It is likely that the traits 'Treatment speciality of consultant: Urology', 'Family history of Alzheimer's disease' and 'squamous cell lung cancer' violate the assumptions of MR, as discussed in the results and discussion.

all other MR analyses that we conducted (weighted median  $p=0.01$ , weighted mode  $p=0.13$ , MR Egger  $p=0.78$ ) and, as such, the evidence for a protective effect was less strong.

Our screen identified two serum proteins with a causal effect on POAG. These are angiotensin-1 receptor (OR = 1.11, MRE IVW  $p=2.34E-06$ ) and cadherin-5 (OR = 1.06, MRE IVW  $p=1.31E-06$ ). Orthogonally we discovered that rare deleterious missense variants within the gene encoding angiotensin-1 receptor (*TEK*) increases risk of POAG (OR = 2.48,  $p=0.009$ ). In contrast, rare variant burden testing did not reveal a relationship between changes in the gene encoding cadherin-5 (*CDH5*) and POAG<sup>17</sup>.

Diabetes (diagnosed by a doctor) was identified to be causally related to POAG (OR = 5.17, MRE IVW beta = 1.64, MRE IVW  $p=9.99E-04$ ). We performed multivariable MR (MVMR) analysis to assess the role of potential confounders including systolic blood pressure, diastolic blood pressure, BMI, waist circumference, physical activity levels, serum high density lipoprotein, and serum low density lipoprotein, but this analysis was inconclusive (Supplementary table 1.4). Individual diabetes subtypes were also explored. Type 2 diabetes was not found to be causally related to POAG risk (IVW  $p=0.21$ ). However when data from a genome-wide association study (GWAS) that adjusted for BMI were used as input, type 2 diabetes was found to be causally related to POAG in two measures (OR = 1.09, IVW beta = 0.09, IVW  $p=0.003$ , weighted median  $p=0.009$ , MR Egger  $p=0.10$ , weighted mode  $p=0.08$ )<sup>18</sup>. Whilst acknowledging the potential significance of this finding, it is important to note that using covariable-adjusted exposure instruments can lead to bias in MR (albeit to a lesser extent than the bias created by covariable-adjusted outcome instruments)<sup>19</sup>. Type 1 diabetes was nominally significant in IVW analysis in relation to POAG ( $p=0.03$ ) but no statistically significant signals were obtained for this trait by the other MR methods used in this study (weighted median  $p=0.07$ , weighted mode  $p=0.10$ , MR Egger  $p=0.19$ ).

Our screen identified family history of Alzheimer's disease as a protective factor for POAG (OR = 0.88, beta = -0.13, IVW  $p=1.63E-06$ ). Given this is a family history trait, it appears probable that the analysis would violate the third key assumption of MR (i.e. that the genetic instruments are likely to be associated with the outcome via horizontal pleiotropy rather than through vertical pleiotropy). Alzheimer's disease itself was not related to POAG risk (IVW  $p=0.96$ ).

Squamous cell lung cancer is suggested to be a protective trait for POAG from our MR results (OR = 0.92, MRE IVW beta = -0.08, MRE IVW  $p=1.35E-04$ ). It is however known that survival bias can distort MR results, and thus this result should be treated with caution<sup>20,21</sup>.

Intriguingly, individuals under the care of a urology service were identified to have a higher risk of POAG (OR = 2.39E + 03, beta = 7.78, IVW  $p=4.27E-10$ ). This of course represents a composite trait confounded by urological phenotypes, risk factors and medicine use. It is not possible to discern the nature of these confounding factors within this study.

## Discussion

In this study, we used MR analysis to advance our understanding of POAG pathogenesis. We outline a robust approach to causal inference that can be applied in other disorders with wide-reaching implications<sup>22</sup>. The validity of our methodology is supported by the identification of IOP as the leading causal trait for POAG; highlighting this well-established risk factor essentially serves as a positive control. Notably, our analysis supported causal roles for several disputed risk factors and identified previously unreported causal traits.

Our findings support a causal relationship between a diagnosis of diabetes and POAG. This is in accordance with an extensive biomedical literature base which also supports this relationship<sup>6</sup>. Furthermore, in patients with existing POAG, diabetes has been shown to advance the severity of glaucoma in the form of raising IOP, worsening visual field defects and resulting in a larger optic cup to disc ratio<sup>23</sup>. Several mechanisms have been proposed to account for the relationship between diabetes and POAG. First, it has been suggested that IOP could be raised due to increased formation of advanced glycation end products causing trabecular meshwork cell damage and resulting in blockage of aqueous humour outflow through the iridocorneal angle; serum/aqueous hyperosmolarity and hyperviscosity may also increase aqueous humour retention which could also impact IOP<sup>24</sup>. It has also been proposed that long standing hyperglycaemia and lipid abnormalities might increase the vulnerability of neurons to stress and, as such, increase vulnerability to POAG development<sup>6</sup>. Whilst this result is interesting, it is important to note that we were not able to rule out a role for other confounding traits, and therefore it remains possible that the relationship between diabetes and POAG is mediated by a third factor, including patterns of behaviour, comorbidities, or levels of healthcare engagement.

There is controversy surrounding the role of adiposity in POAG. There are numerous measures of adiposity including BMI, waist circumference, waist-hip ratio, and body fat percentage; each of these measures has its own strengths and weaknesses<sup>25</sup>. For instance, it is understood that BMI in isolation does not serve to deliver sufficient information regarding adiposity to fully evaluate cardiometabolic disease risk. As a result, it has been recommended that waist circumference is adopted as an additional routine measurement to identify and classify obesity<sup>25-27</sup>. Although lower BMI has been associated with worse glaucoma outcomes in some studies, this finding is not universal<sup>28</sup>. Further, some studies have identified a causal relationship between obesity and POAG<sup>29</sup>. A possible explanation between conflicting results between studies of adiposity and POAG, is that there is a non-linear relationship between these two variables. Whilst the pathophysiology underpinning any relationship between adiposity and glaucoma is as of yet unclear, some authors have suggested that elevated levels of hormones such as ghrelin, oestrogen and leptin in those with higher BMI might have neuroprotective effects, or that elevated cerebrospinal fluid pressure in those with greater body fat may reduce the translamellar pressure gradient and thus counteract the effects of raised IOP<sup>28</sup>. Here, we find that waist circumference is inversely associated with POAG risk and, as such, increasing waist circumference is protective for POAG. The evidence for a relationship with BMI is less compelling but it is noted that this measure does not necessarily serve as a good measure of fat distribution, abdominal obesity or fat/fat free mass<sup>25</sup>. Our findings are at odds with the conclusions of a

recent MR study by Lin et al., which suggested that obesity may be associated with increased POAG risk<sup>30</sup>. The principal difference between our analysis and that of Lin et al. is our use of a larger POAG population (15,229 cases and 177,473 controls versus 1824 POAG cases and 93,036 controls). It is also noted that the conclusions by Lin et al. are not strongly supported by some of their MR analyses which did not reveal statistically significant signals; this may suggest that the relevant results are vulnerable to outlier SNVs and horizontal pleiotropy. In summary, our findings suggest that abdominal obesity may be associated with decreased risk of POAG. Given that these results are not in agreement with observations from a previous study by Lin et al., there is remaining controversy<sup>27</sup> and further work is warranted in this area, perhaps including the exploration of a non-linear relationship between adiposity and POAG.

Cadherins are a group of transmembrane proteins responsible for cell adhesion. These molecules have received little attention in relation to POAG, although a small number of studies describe a potential role in this condition. In a previous genome-wide copy number variant study, cadherin signalling was implicated in POAG susceptibility<sup>31</sup>. The authors of that study noted that cell adhesion proteins determine trabecular meshwork resistance to aqueous outflow and, as such, may impact IOP and therefore have a role in glaucoma. Our MR results have shown a causal effect of cadherin-5 on POAG (although further evidence on this relationship could not be obtained through rare variant burden testing). Cadherin 5 (also known as vascular endothelial (VE) cadherin) is not known to be expressed in trabecular meshwork cells, but it is expressed in the endothelial cells of Schlemm's canal<sup>32</sup>. Notably, a recent study by Kelly et al. reported that primary Schlemm's canal cells from individuals with glaucoma had higher levels of cadherin 5 protein expression compared to cells from healthy donor eyes<sup>33</sup>. It was therefore speculated by the authors that cadherin 5 may increase aqueous outflow resistance through decreasing paracellular permeability. Whilst it is important to note that our study relates to serum cadherin 5 rather than expression local to the anterior segment, our results do support a role for this protein in POAG pathophysiology.

Angiotensin-1 receptor is a cell surface receptor encoded by the gene *TEK*. Angiotensin proteins are vascular growth factors involved in the development and remodelling of blood vessels, and angiotensin-TEK signalling is pivotal to the development of Schlemm's canal<sup>34</sup>. Previous work in mice has demonstrated that angiotensin has a critical role in preserving the integrity of Schlemm's canal and it has been suggested that abnormalities of angiotensin physiology may have a role on POAG<sup>35</sup>. Similarly, human genetic studies have provided support for a role for angiotensin physiology in POAG, with common variants in the angiotensin-1 receptor ligand (*ANGPT1*) having been found to influence IOP<sup>36</sup> and mutations in *TEK* causing primary congenital glaucoma<sup>37</sup>. Our MR analysis supports a causal role for serum angiotensin-1 receptor levels in POAG and our rare variant burden test suggests that loss of function of the corresponding gene (*TEK*) is linked to POAG development. Whilst the MR and rare variant analyses are not concordant with respect to the direction of effect, this may represent differential expression of angiotensin-1 receptor intraocularly and in the serum. As for cadherin 5, it is outside the scope of this study to investigate the local effects of angiotensin-1 receptor in POAG but our findings suggest a role for serum protein levels and support a causal role for this molecule in POAG pathophysiology.

Family history of Alzheimer's disease and a diagnosis of squamous cell lung cancer were both found to have relationships with POAG within this study. We treat these results with a high degree of caution. The former is likely a product of violation of key assumptions of MR, whilst the latter is at high risk of survival bias. As such, whilst it is not possible to comment with certainty, we believe these results to be a consequence of bias. It is understood that MR in elderly populations may produce results distorted by bias (as the analysis surrounds a non-random subset of the population who have survived long enough to develop a given disease)<sup>9,21</sup>. With that in mind, it is evident that traits with a strong effect on survival are most likely to be biased and as such should be treated with a greater degree of caution. Additionally, recipients of care under urology were found to be more likely to develop POAG; this represents a broad composite trait inclusive of several confounding medical conditions, medications, other phenotypes and changes to patterns of healthcare engagement and as such this result is not interpretable.

Our study has several limitations. First, we applied very strict criteria with respect to which traits we considered notable as we wanted to focus on the most robust results (which could then be explored in detail). As a result, our phenome-wide screen may have failed to identify true causal relationships for several traits (due to type 2 error). MR Egger is particularly underpowered relative to other MR tests and, whilst the inclusion of this method is likely to reduce the number of false positive findings, it may also lead to overlooking important traits. For instance, two potentially interesting proteins with roles in the trabecular meshwork (serum inter-alpha-trypsin inhibitor heavy chain H and 72 kDa inositol polyphosphate 5-phosphatase) are significant in all measures except for MR Egger and could represent novel glaucoma-related molecules. This limitation applies to the other MR methods that we used, and a table presenting all results which were significant in FDR corrected IVW can be found in Supplementary table 1.2 for readers to explore and research further. Second, as with all MR analyses, our results are vulnerable to violations of the assumptions of MR (including horizontal pleiotropy and unmeasured confounding) and, as such, no MR result in isolation can be categorically regarded to prove causation. Thirdly, in order to avoid bias due to population structure, the present analysis utilised data derived from populations of European ancestries only. This increases the risk of the results not being generalisable across populations. Finally, in this study we elected to exclude eQTL and imaging traits.

In conclusion, this study provides support for a role for intraocular pressure and diabetes in POAG, suggests a protective effect for adiposity and highlights potential causal roles for serum cadherin 5 and angiotensin-1 receptor. Future research on the effects of waist circumference, serum cadherin 5 and serum angiotensin-1 receptor on POAG progression will provide key insights with a view to developing lifestyle modification advice and drug targets. Further, we suggest study of the local ocular effects of cadherin 5 and angiotensin-1 receptor in order to glean mechanistic insights into the effects that topical interventions might have.

## Methods

**Data sources.** *Exposure data.* In this study we performed a phenome-wide screen to make causal inferences on the role of an extensive range of traits on POAG. To achieve this, we used both published and unpublished GWAS data from the MRC IEU open GWAS database (available online via <https://gwas.mrcieu.ac.uk/>); these were accessible via the TwoSampleMR programme in R<sup>38</sup>. We included all ‘European GWAS’ within the database but removed imaging phenotypes and expression quantitative trait loci (eQTL) data due to difficulty in biological interpretation. ‘Finn’ trait GWAS data (derived from the FinnGen project) were excluded due to sample overlap. As it was not possible to manually inspect the degree of population overlap for all traits prior to conducting the analysis, the degree of overlap for all significant traits was inspected after the analysis.

*Outcome data.* The Gharahkhani et al. GWAS for POAG was utilised to provide the outcome data. Data were obtained from individuals of European ancestries only (16,677 POAG cases and 199,580 controls). Full details regarding the diagnosis of POAG and data sources contributing to the meta-analysis are detailed in the original paper<sup>7</sup>. Where an exposure study was indexed in the MRC Integrative Epidemiology Unit (IEU) database with the prefix “ukb” or “met-d” (those which are explicitly 100% UK Biobank subjects) we utilised a version of the Gharahkhani et al. GWAS in which UK Biobank subjects were omitted (n = 15,229 POAG cases and 177,473 controls)<sup>39</sup>. For all other studies, we manually inspected the composite populations if traits were identified as significant in order to ensure there was not significant sample overlap which might increase the risk of bias.

**Instrument selection.** In this study, we used a statistically-driven approach to instrumental variable selection. Briefly, an arbitrary  $p$  value threshold was set for the identification of appropriate SNVs that could be used as instrumental variables (referred to thereafter as instruments). A conventional value for selection of instruments is  $p < 5E-08$ , but this can, in some cases, be problematic. For example, when the number of instruments exceeding this threshold is small, the analysis can be underpowered or, in certain cases of unbiased screens, the results can be inflated<sup>40</sup>. With this in mind, we have set the  $p$  value for instrument selection for each trait to the level where  $> 5$  instruments are available for each analysis. More specifically, for each trait, the analysis would first be conducted with a  $p$  value for inclusion of  $5E-8$ , and this would be sequentially increased by a factor of 10 each time until  $> 5$  eligible instruments are identified to a predefined maximum of  $p < 5E-04$ . Ultimately, all significant traits were generated with instruments of  $p < 5E-06$ .

**Proxies.** Where an exposure instrument was not present in the outcome dataset, we sought to identify a suitable proxy<sup>41</sup>. Proxy variants were SNVs which were in high linkage disequilibrium with exposure-associated variants which were not present in the outcome dataset. Proxies were identified using the ensembl server with a linkage disequilibrium  $R^2$  value of 0.9 or above for identification of SNVs<sup>38,42</sup>.

**Clumping.** SNVs were clumped using a linkage disequilibrium  $R^2$  value of 0.001 and a genetic distance cut-off of 10,000 kilo-bases. This was achieved using the 1000 genomes European superpopulation reference panel and the TwoSampleMR PLINK clumping function.

**Harmonisation.** The effects of instruments on outcomes and exposures were harmonised in order to ensure that the beta values (i.e. the regression analysis estimates of effect size) were signed with respect to the same alleles<sup>41</sup>. For palindromic alleles (i.e. alleles that are the same on the forward as on the reverse strand), those with minor allele frequency (MAF)  $> 0.42$  were omitted from the analysis in order to reduce the risk of errors due to strand issues.

**Removal of pleiotropic genetic variants and outliers.** In addition to using a range of MR methods and quality control measures (as detailed below), we endeavoured to remove pleiotropic instruments and outliers from the analysis. We operated this using ‘Steiger filtering’, a statistical approach in which instruments which were more significant (in terms of  $p$  value) for the outcome than the exposure are removed<sup>43</sup>. Radial MR, an approach that detects and removes or downweights outlying instruments was also utilised for outlier identification<sup>44</sup>.

**Causal inference.** MR estimation was primarily performed using a multiplicative random effects (MRE) inverse variance weighted (IVW) analysis. MRE IVW was selected over a fixed effects (FE) approach as it allows inclusion of heterogeneous instruments, certain to occur within the breadth of this screen<sup>45</sup>. MR relies on the selected instruments satisfying three core assumptions, these being that: (1) the instrumental SNVs are robustly associated with the exposure; (2) there must be no confounders of the SNV’s and the outcome; (3) the SNVs are associated with the disease risk outcome only through the exposure of interest<sup>9</sup>. A range of methods were utilised to ensure the accuracy of the results and to cover for a range of violations of the aforementioned MR assumptions<sup>45</sup>. The following approaches were selected: weighted median<sup>46</sup>, MR Egger<sup>47</sup>, weighted mode<sup>48</sup> and radial MR with modified second order weights (for the purpose of outlier identification only)<sup>44</sup>. These robust methods to reduce vulnerability to violation of specific assumptions, whilst each having its own weakness which should be considered when evaluating their estimates. The MR Egger’s weaknesses are that it is a low power method and that it relies on the InSIDE (Instrument Strength Independent of Direct Effect) assumption, which assumes that the correlation of the genetic variants with the exposure and the direct effect of the variants on the outcome is zero<sup>49</sup>. The power of this method, however, is that the causal relationship can still be evaluated using the MR Egger even in the context of directional pleiotropy and with 100% invalid instruments. The weighted

median's main strength is that the estimates provided are robust to outliers, with the primary weakness being the 'majority valid' assumption that > 50% of the instruments utilised are valid<sup>45</sup>. The weighted mode analysis is also a relatively low power measure which assumes that more variants assume the true causal estimate than any other quantity (plurality valid). Causal estimates from the weighted mode are valid even with mostly invalid instruments and the measure benefits from a low type 1 error rate<sup>48</sup>.

**Further quality control.** Instrument strength was determined using the F statistic (which tests the association between the instruments and the exposure)<sup>50</sup>. F statistics were calculated against the final included set of instruments. Instrumental F statistics were calculated as  $((N-1-k)/k)/(R^2/(1-R^2))$  where N = sample size, k = the number of instruments and  $R^2$  was the proportion of variance explained by the instrument. A mean F statistic > 10 was considered sufficiently strong.

The Cochran's Q test was performed for each analysis. Cochran's Q is a measure of heterogeneity among causal estimates and serves as an indicator of the presence of horizontal pleiotropy (which occurs when an instrument exhibits effects on the outcome through pathways other than the exposure)<sup>51</sup>. It is noted that a heterogeneous instrument is not necessarily invalid, but rather calls for a primary assessment with an MRE IVW rather than an FE approach; this has been conducted as standard throughout our analysis.

The MR Egger intercept test was used to detect horizontal pleiotropy. When this occurs, the MR Egger regression is robust to horizontal pleiotropy, under the assumption that that pleiotropy is uncorrelated with the association between the SNV and the exposure<sup>52</sup>.

The  $I^2$  statistic was calculated as a measure of heterogeneity between variant specific causal estimates. An  $I^2 < 0.9$  indicates that MR Egger is more likely to be biased towards the null through violation of the 'NO Measurement Error' (NOME) assumption<sup>53</sup>.

Leave one out cross-validation was performed for every analysis to determine if any particular SNV was driving the significance of the causal estimates.

**Management of duplicate traits/GWAS.** As the GWAS database used contains multiple different GWAS studies for certain traits, some exposures were analysed on multiple occasions. Where this occurred, we considered the largest sample size study to be the primary analysis. Where there were duplicate studies in the same population, we used the study with the largest F statistic. Where a binary trait was identified to have a causal relationship with POAG, we additionally ensured that the data utilised were the one that had the largest number of cases for the condition of interest.

**Identification of significant results.** Before considering an MR result to be significant, the results of a range of causal inference and quality control tests should be taken into account. Notably, it is not necessary for a study to find significance in all measures to support a causal relationship. MR is a low power study type and, as such, an overly conservative approach to multiple testing can be excessive<sup>45</sup>. However, in the context of the present study we have chosen to explore and discuss the results of our POAG phenome-wide screen only if they remained: significant after FDR correction in the MRE IVW; nominally significant in weighted mode, MR Egger and weighted median; and nominally significant throughout the leave one out analysis (MRE IVW)<sup>54</sup>. This approach was selected as a large number of phenotypes were considered and because we wanted to focus on high-confidence signals to ensure reproducibility. Importantly, we expect our approach to overlook some important signals given the low power of our methodology. All our results are presented in the Supplementary Information section and traits significant according to their FDR corrected *p* value in the IVW analysis represent phenotypes that may be of interest for further research and follow-up outside of this study. FDR corrections were applied using the Benjamini–Hochberg procedure, with correction applied for the total number of traits (9661)<sup>54</sup>.

**Presentation of effect sizes.** Effect sizes are presented to enable readers to appraise the impact of putative risk factors. In univariable MR, these are shown as odds ratio (OR) per 1 standard deviation (SD) change in the exposure for continuous traits, and as beta values plus the OR for binary exposure traits. This approach was chosen as there are differences of opinion with regards to how best to present effect sizes for binary exposure traits in MR, particularly when a binary trait is a dichotomised continuous variable<sup>55</sup>.

**Rare variant burden testing.** For significant protein phenotypes, we performed an orthogonal analysis using a gene-based rare variant burden test to aid functional interpretation. We utilised the AstraZeneca PheWas portal which includes whole genome sequencing data from the UK Biobank<sup>17</sup>. Rare variants were defined as those having minor allele frequency less than 0.001 within both the UK Biobank and the Genome Aggregation Database (gnomAD)<sup>56</sup>. A variant was assigned as deleterious if it was associated with a change in the amino-acid sequence and reached a REVEL score > 0.25<sup>57</sup>. Variants that met these criteria were collapsed into a single two-sided Fisher's exact test to assess the relative frequency in individuals with POAG and in unaffected controls. Our analysis included 3,088 POAG cases and 339,310 controls. A rare variant burden reaching nominal significance was considered significantly associated with POAG.

**Multivariable MR (MVMR).** Multivariable MR was performed in circumstances where it was important to estimate the effect of > 1 closely related (and/or potentially confounding) exposure trait. *P*-values for the inclusion of instruments for the exposures of interest were optimised to obtain sufficiently high (> 10) conditional *F*-statistics for reliable analysis<sup>38</sup>. Selection for exposures began at a *p*-value threshold of > 5E–08. Where trait's

instruments had a conditional  $F$ -statistic  $< 10$ , the  $p$ -value for selection was reduced in an automated manner by factor of 10 until an  $F$ -statistic  $> 10$  was obtained. The same clumping procedure as in the univariable MR analysis was used. Adjusted Cochran's  $Q$ -statistics were calculated. Where exposure traits were derived from overlapping cohorts, we analysed the data assuming both low (0.1) and high (0.9) levels of phenotypic correlation to account for the full range of potential conditional  $F$  and adjusted  $Q$  statistics<sup>58</sup>.

**Ethical approval.** This study utilised publicly available data and no additional ethical approval was required.

### Data availability

All data associated with this study are published in the Supplementary Information. **Software:** R version 4.1.0. TwoSampleMR version 0.5.6. RadialMR version 1.0. MVMR version 0.3.

Received: 15 October 2022; Accepted: 16 June 2023

Published online: 20 June 2023

### References

- Jonas, J. B. *et al.* Glaucoma. *Lancet* **390**, 2183–2193 (2017).
- Weinreb, R. N. *et al.* Primary open-angle glaucoma. *Nat Rev Dis Primers* **2**, 16067 (2016).
- Weinreb, R. N., Aung, T. & Medeiros, F. A. The pathophysiology and treatment of glaucoma: A review. *JAMA* **311**, 1901–1911 (2014).
- Grzybowski, A., Och, M., Kanclerz, P., Leffler, C. & Moraes, C. G. D. Primary open angle glaucoma and vascular risk factors: A review of population based studies from 1990 to 2019. *J. Clin. Med. Res.* **9**, 761 (2020).
- 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).
- Zhou, M., Wang, W., Huang, W. & Zhang, X. Diabetes mellitus as a risk factor for open-angle glaucoma: A systematic review and meta-analysis. *PLoS ONE* **9**, e102972 (2014).
- Gharahkhani, P. *et al.* Genome-wide meta-analysis identifies 127 open-angle glaucoma loci with consistent effect across ancestries. *Nat. Commun.* **12**, 1258 (2021).
- Craig, J. E. *et al.* Multitrait analysis of glaucoma identifies new risk loci and enables polygenic prediction of disease susceptibility and progression. *Nat. Genet.* **52**, 160–166 (2020).
- Julian, T. H. *et al.* A review of Mendelian randomization in amyotrophic lateral sclerosis. *Brain* <https://doi.org/10.1093/brain/awab420> (2021).
- Sanderson, E. *et al.* Mendelian randomization. *Nat Rev. Methods Primers* **2**, 1–21 (2022).
- Liu, K. *et al.* Implicating causal brain magnetic resonance imaging in glaucoma using mendelian randomization. *Front. Med.* **9**, 956339 (2022).
- Xu, M. *et al.* Plasma lipid levels and risk of primary open angle glaucoma: A genetic study using Mendelian randomization. *BMC Ophthalmol.* **20**, 390 (2020).
- Bao, J., Yang, Z., Zheng, S., Li, J. & Shentu, X. Circulating fatty acids and risk of primary open-angle glaucoma: A mendelian randomization study. *Gene* **811**, 146078 (2022).
- Hu, Z., Zhou, F., Kamina, A. C. & Xu, H. Type 2 diabetes, fasting glucose, hemoglobin A1c levels and risk of primary open-angle glaucoma: A mendelian randomization study. *Invest. Ophthalmol. Vis. Sci.* **63**, 37 (2022).
- Li, X. *et al.* Habitual coffee consumption increases risk of primary open-angle glaucoma: A mendelian randomization study. *Ophthalmology* <https://doi.org/10.1016/j.ophtha.2022.04.027> (2022).
- Plotnikov, D. *et al.* High blood pressure and intraocular pressure: A Mendelian randomization study. *Invest. Ophthalmol. Vis. Sci.* **63**, 29 (2022).
- Wang, Q. *et al.* Rare variant contribution to human disease in 281,104 UK Biobank exomes. *Nature* **597**, 527–532 (2021).
- Mahajan, A. *et al.* Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. *Nat. Genet.* **50**, 559–571 (2018).
- Hartwig, F. P., Tilling, K., Davey Smith, G., Lawlor, D. A. & Borges, M. C. Bias in two-sample Mendelian randomization when using heritable covariable-adjusted summary associations. *Int. J. Epidemiol.* **50**, 1639–1650 (2021).
- Boef, A. G. C., le Cessie, S. & Dekkers, O. M. Mendelian randomization studies in the elderly. *Epidemiology* **26**, e15–e16 (2015).
- Smit, R. A. J., Trompet, S., Dekkers, O. M., Jukema, J. W. & le Cessie, S. Survival bias in Mendelian randomization studies: A threat to causal inference. *Epidemiology* **30**, 813–816 (2019).
- Julian, T. H. *et al.* Phenome-wide Mendelian randomisation analysis identifies causal factors for age-related macular degeneration. *Elife* **12**, e82546 (2023).
- Khatri, A., Shrestha, J. K., Thapa, M., Khatri, B. K. & Kharel, M. Severity of primary open-angle glaucoma in patients with hypertension and diabetes. *Diabetes Metab. Syndr. Obes.* **11**, 209–215 (2018).
- Li, Y., Mitchell, W., Elze, T. & Zebardast, N. Association between diabetes, diabetic retinopathy, and glaucoma. *Curr. Diab. Rep.* **21**, 38 (2021).
- van Dijk, S. B., Takken, T., Prinsen, E. C. & Wittink, H. Different anthropometric adiposity measures and their association with cardiovascular disease risk factors: a meta-analysis. *Neth. Heart J.* **20**, 208–218 (2012).
- Ross, R. *et al.* Waist circumference as a vital sign in clinical practice: A consensus statement from the IAS and ICCR working group on visceral obesity. *Nat. Rev. Endocrinol.* **16**, 177–189 (2020).
- Lin, S.-C., Pasquale, L. R., Singh, K. & Lin, S. C. The association between body mass index and open-angle glaucoma in a south Korean population-based sample. *J. Glaucoma* **27**, 239–245 (2018).
- Marshall, H. *et al.* Association between body mass index and primary open-angle glaucoma in three cohorts. *Am. J. Ophthalmol.* <https://doi.org/10.1016/j.ajo.2022.08.006> (2022).
- Liu, W., Ling, J., Chen, Y., Wu, Y. & Lu, P. The association between adiposity and the risk of glaucoma: A meta-analysis. *J. Ophthalmol.* **2017**, 9787450 (2017).
- Lin, Y. *et al.* The causal association between obesity and primary open-angle glaucoma: A two-sample mendelian randomization study. *Front. Genet.* **13**, 835524 (2022).
- Lo Faro, V., Ten Brink, J. B., Snieder, H., Jansonius, N. M. & Bergen, A. A. Genome-wide CNV investigation suggests a role for cadherin, Wnt, and p53 pathways in primary open-angle glaucoma. *BMC Genomics* **22**, 590 (2021).
- Perkumas, K. M. & Stamer, W. D. Protein markers and differentiation in culture for Schlemm's canal endothelial cells. *Exp. Eye Res.* **96**, 82–87 (2012).
- Kelly, R. A. *et al.* Fibrotic changes to Schlemm's canal endothelial cells in glaucoma. *Int. J. Mol. Sci.* **22**, 9446 (2021).

34. Thomson, B. R. *et al.* Angiopoietin-1 is required for Schlemm's canal development in mice and humans. *J. Clin. Invest.* **127**, 4421–4436 (2017).
35. Kim, J. *et al.* Impaired angiopoietin/Tie2 signaling compromises Schlemm's canal integrity and induces glaucoma. *J. Clin. Invest.* **127**, 3877–3896 (2017).
36. MacGregor, S. *et al.* Genome-wide association study of intraocular pressure uncovers new pathways to glaucoma. *Nat. Genet.* **50**, 1067–1071 (2018).
37. Souma, T. *et al.* Angiopoietin receptor TEK mutations underlie primary congenital glaucoma with variable expressivity. *J. Clin. Invest.* **126**, 2575–2587 (2016).
38. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, e34408 (2018).
39. Elsworth, L., Alexander, L. & Matthews. The MRC IEU OpenGWAS data infrastructure. bioRxiv. 2020. *Google Scholar*.
40. Boddy, S. *et al.* Unbiased metabolome screen leads to personalized medicine strategy for amyotrophic lateral sclerosis. *Brain Commun.* **4**, fcac069 (2022).
41. Hartwig, F. P., Davies, N. M., Hemani, G. & Davey Smith, G. Two-sample Mendelian randomization: Avoiding the downsides of a powerful, widely applicable but potentially fallible technique. *Int. J. Epidemiol.* **45**, 1717–1726 (2016).
42. Cunningham, F. *et al.* Ensembl 2022. *Nucleic Acids Res.* **50**, D988–D995 (2022).
43. Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet.* **13**, e1007081 (2017).
44. Bowden, J. *et al.* Improving the visualization, interpretation and analysis of two-sample summary data Mendelian randomization via the radial plot and Radial regression. *Int. J. Epidemiol.* **47**, 1264–1278 (2018).
45. Burgess, S. *et al.* Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res.* **4**, (2019).
46. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**, 304–314 (2016).
47. Burgess, S. & Thompson, S. G. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* **32**, 377–389 (2017).
48. Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int. J. Epidemiol.* **46**, 1985–1998 (2017).
49. Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **44**, 512–525 (2015).
50. Burgess, S., Thompson, S. G., CRPCHD Genetics Collaboration. Avoiding bias from weak instruments in Mendelian randomization studies. *Int. J. Epidemiol.* **40**, 755–764 (2011).
51. Bowden, J. & Holmes, M. V. Meta-analysis and Mendelian randomization: A review. *Res. Synth. Methods* **10**, 486–496 (2019).
52. Burgess, S. & Thompson, S. G. Erratum to: Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur. J. Epidemiol.* **32**, 391–392 (2017).
53. Bowden, J. *et al.* Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: The role of the I<sup>2</sup> statistic. *Int. J. Epidemiol.* <https://doi.org/10.1093/ije/dyw220> (2016).
54. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57**, 289–300 (1995).
55. Burgess, S. & Labrecque, J. A. Mendelian randomization with a binary exposure variable: Interpretation and presentation of causal estimates. *Eur. J. Epidemiol.* **33**, 947–952. <https://doi.org/10.1007/s10654-018-0424-6> (2018).
56. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
57. Ioannidis, N. M. *et al.* REVEL: An ensemble method for predicting the pathogenicity of rare missense variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).
58. Sanderson, E., Spiller, W. & Bowden, J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat. Med.* **40**, 5434–5452 (2021).

## Acknowledgements

We acknowledge the following sources of funding: the Wellcome Trust (224643/Z/21/Z, Clinical Research Career Development Fellowship to P.I.S.; 200990/Z/16/Z, Transforming Genetic Medicine Initiative to G.C.B.); the UK National Institute for Health Research (NIHR) Clinical Lecturer Programme (CL-2017-06-001 to P.I.S.); the UK NIHR Academic Clinical Fellowship Programme (T.H.J.); Retina UK and Fight for Sight (GR586, RP Genome Project—UK Inherited Retinal Disease Consortium to G.C.B.); the University of Manchester's Wellcome Institutional Strategic Support Fund (Wellcome ISSF) grant (204796/Z/16/Z).

## Author contributions

T.H.J., J.C.K., E.S. and P.I.S. were involved in study design, data analysis, data interpretation, and manuscript write up. G.C.B., H.G., J.Y. and Z.G. were involved in data interpretation and manuscript write up. All authors approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-37144-7>.

**Correspondence** and requests for materials should be addressed to P.I.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.





**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023