STROBE-MR checklist of recommended items to address in reports of Mendelian randomization studies[1 2]

Item No.	Section	Checklist item	Page No.	Relevant text from manuscript
1	**TITLE and ABSTRACT**	Indicate Mendelian randomization (MR) as the study's design in the title and/or the abstract if that is a main purpose of the study	2	Methods section of the abstract states that bidirectional Mendelian randomization (MR) analyses were performed.
	INTRODUCTION			
2	**Background**	Explain the scientific background and rationale for the reported study. What is the exposure? Is a potential causal relationship between exposure and outcome plausible? Justify why MR is a helpful method to address the study question	3-5	Introduction describes the epidemiological association between cataract and MADs, limitations of observational studies, and rationale for using genetic approaches including MR.
3	**Objectives**	State specific objectives clearly, including pre-specified causal hypotheses (if any). State that MR is a method that, under specific assumptions, intends to estimate causal effects	4-5	Study objective includes examining potential causal relationships using bidirectional two-sample MR.
	METHODS			
4	**Study design and data sources**	Present key elements of the study design early in the article. Consider including a table listing sources of data for all phases of the study. For each data source contributing to the analysis, describe the following:		
	a)	Setting: Describe the study design and the underlying population, if possible. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection, when available.	5-6; Supplementary Table 1	Publicly available GWAS summary statistics for cataract and six MAD phenotypes were obtained from FinnGen, PGC, UK Biobank, IEU OpenGWAS, and published GWAS datasets. Detailed dataset characteristics are provided in Supplementary Table 1.
	b)	Participants: Give the eligibility criteria, and the sources and methods of selection of participants. Report the sample size, and whether any power or sample size calculations were carried out prior to the main analysis	5-6; Supplementary Table 1	Sample sizes, case-control numbers, ancestry information, and phenotype definitions for all GWAS datasets are reported in Methods 2.1 and Supplementary Table 1.
	c)	Describe measurement, quality control and selection of genetic variants	11-12	Instrumental variables were selected at genome-wide significance ($P < 5 \times 10^{-8}$) and LD clumped at $r^2 < 0.001$ within a 10,000-kb window.
	d)	For each exposure, outcome, and other relevant variables, describe methods of assessment and diagnostic criteria for diseases	5-6	Cataract and MAD phenotypes were defined according to the original GWAS studies and are

					described in Methods 2.1 and Supplementary Table 1.
		e)	Provide details of ethics committee approval and participant informed consent, if relevant	N/A	All analyses were based on publicly available GWAS summary statistics from studies that had received relevant ethical approval and participant consent.
5	**Assumptions**		Explicitly state the three core IV assumptions for the main analysis (relevance, independence and exclusion restriction) as well assumptions for any additional or sensitivity analysis	11-12	Instrument strength was ensured through genome-wide significant IV selection. MR-Egger, MR-PRESSO, Cochran's Q, and Steiger tests were used to assess potential violations of MR assumptions.
6	**Statistical methods: main analysis**		Describe statistical methods and statistics used		
		a)	Describe how quantitative variables were handled in the analyses (i.e., scale, units, model)	11-12	Effect estimates were evaluated using IVW, MR-Egger, weighted median, weighted mode, and GSMR.
		b)	Describe how genetic variants were handled in the analyses and, if applicable, how their weights were selected	11	LD clumping was performed at $r^2 < 0.001$ within a 10,000-kb window using the European 1000 Genomes reference panel.
		c)	Describe the MR estimator (e.g. two-stage least squares, Wald ratio) and related statistics. Detail the included covariates and, in case of two-sample MR, whether the same covariate set was used for adjustment in the two samples	11-12	IVW was used as the primary MR estimator, with MR-Egger, weighted median, weighted mode, and GSMR used as complementary approaches.
		d)	Explain how missing data were addressed	N/A	Not applicable. Analyses were based on publicly available GWAS summary statistics.
		e)	If applicable, indicate how multiple testing was addressed	12	Statistical significance was defined as $P < 0.05$ for MR analyses.
7	**Assessment of assumptions**		Describe any methods or prior knowledge used to assess the assumptions or justify their validity	11-12	MR-Egger, MR-PRESSO, Cochran's Q, leave-one-out analysis, and Steiger test were used to assess the validity of MR assumptions.
8	**Sensitivity analyses and additional analyses**		Describe any sensitivity analyses or additional analyses performed (e.g. comparison of effect estimates from different approaches, independent replication, bias analytic techniques, validation of instruments, simulations)	11-12	Sensitivity analyses included MR-Egger, MR-PRESSO, leave-one-out analysis, Cochran's Q test, and Steiger test.
9	**Software and pre-registration**				

		a)	Name statistical software and package(s), including version and settings used	7-12; Figure 1	Statistical analyses were performed using established software packages as described in the Methods.
		b)	State whether the study protocol and details were pre-registered (as well as when and where)	N/A	Not preregistered.
	RESULTS				
10	**Descriptive data**				
		a)	Report the numbers of individuals at each stage of included studies and reasons for exclusion. Consider use of a flow diagram	N/A	Not applicable. Individual-level inclusion and exclusion information was not available because analyses were based on publicly available GWAS summary statistics.
		b)	Report summary statistics for phenotypic exposure(s), outcome(s), and other relevant variables (e.g. means, SDs, proportions)	N/A	Not applicable. Individual-level phenotypic data were not available because all analyses were based on publicly available GWAS summary statistics.
		c)	If the data sources include meta-analyses of previous studies, provide the assessments of heterogeneity across these studies	N/A	Not applicable. Formal between-study heterogeneity assessments were not available because all analyses were based on publicly available GWAS summary statistics.
		d)	For two-sample MR: i. Provide justification of the similarity of the genetic variant-exposure associations between the exposure and outcome samples ii. Provide information on the number of individuals who overlap between the exposure and outcome studies	i. 9-11 ii. 23	i. Exposure and outcome GWAS datasets were derived predominantly from large population-based cohorts of European ancestry. ii. Exact overlap was unavailable. Potential sample overlap was evaluated using LDSC genetic covariance intercepts, which showed no evidence of substantial bias.
11	**Main results**				
		a)	Report the associations between genetic variant and exposure, and between genetic variant and outcome, preferably on an interpretable scale	18-19	MR effect estimates derived from genetic variant–exposure and genetic variant–outcome associations are reported.
		b)	Report MR estimates of the relationship between exposure and outcome, and the measures of uncertainty from the MR analysis, on an interpretable scale, such as odds ratio or relative risk per SD difference	18-19; Figure 4; Supplementary Table 25	No evidence of a significant causal effect was observed in either direction across the primary MR analyses.
		c)	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	N/A	Not applicable.

		d)	Consider plots to visualize results (e.g. forest plot, scatterplot of associations between genetic variants and outcome versus between genetic variants and exposure)	19; Figure 4	Results of MR analyses are presented in Figure 4.
12	**Assessment of assumptions**				
		a)	Report the assessment of the validity of the assumptions	18-19; Figure 4; Supplementary Table 25	MR assumptions were evaluated using F-statistics, MR-Egger regression, MR-PRESSO, Cochran's Q test, and the Steiger test.
		b)	Report any additional statistics (e.g., assessments of heterogeneity across genetic variants, such as I^2, Q statistic or E-value)	18	Sensitivity analyses showed no evidence supporting a causal relationship.
13	**Sensitivity analyses and additional analyses**				
		a)	Report any sensitivity analyses to assess the robustness of the main results to violations of the assumptions	18-19; Supplementary Table 25	Multiple sensitivity analyses were conducted to evaluate robustness.
		b)	Report results from other sensitivity analyses or additional analyses	18; Figure 4; Supplementary Table 25	Additional analyses included generalized summary-data-based Mendelian randomization (GSMR), which yielded results consistent with the primary MR analyses.
		c)	Report any assessment of direction of causal relationship (e.g., bidirectional MR)	18-19; Figure 4; Supplementary Table 25	Bidirectional MR analyses were conducted to evaluate causal direction.
		d)	When relevant, report and compare with estimates from non-MR analyses	12-18	MR findings were interpreted together with genetic correlation, local correlation, and cross-trait GWAS analyses.
		e)	Consider additional plots to visualize results (e.g., leave-one-out analyses)	Figure 4	Forest plot of bidirectional MR results.
	DISCUSSION				
14	**Key results**		Summarize key results with reference to study objectives	19	MDD exhibited the strongest and most consistent shared genetic architecture with cataract, while MR analyses provided no evidence for a direct causal relationship.

15	**Limitations**		Discuss limitations of the study, taking into account the validity of the IV assumptions, other sources of potential bias, and imprecision. Discuss both direction and magnitude of any potential bias and any efforts to address them	23-24	Limitations related to GWAS summary statistics, ancestry composition, phenotype heterogeneity, and statistical power are discussed.
16	**Interpretation**				
		a)	Meaning: Give a cautious overall interpretation of results in the context of their limitations and in comparison with other studies	19-23	Findings support pleiotropic overlap rather than direct causality.
		b)	Mechanism: Discuss underlying biological mechanisms that could drive a potential causal relationship between the investigated exposure and the outcome, and whether the gene-environment equivalence assumption is reasonable. Use causal language carefully, clarifying that IV estimates may provide causal effects only under certain assumptions	20-23	Shared biological pathways and pleiotropic mechanisms are discussed; causal interpretations are made cautiously.
		c)	Clinical relevance: Discuss whether the results have clinical or public policy relevance, and to what extent they inform effect sizes of possible interventions	24	Findings improve understanding of the biological basis of cataract–MDD comorbidity and may inform future prevention and management strategies.
17	**Generalizability**		Discuss the generalizability of the study results (a) to other populations, (b) across other exposure periods/timings, and (c) across other levels of exposure	24	Generalizability across ancestries and populations is discussed.
	OTHER INFORMATION				
18	**Funding**		Describe sources of funding and the role of funders in the present study and, if applicable, sources of funding for the databases and original study or studies on which the present study is based	25	Funding sources for the present study are reported in the Funding section. Funding information for the original GWAS datasets is available in the corresponding source publications.
19	**Data and data sharing**		Provide the data used to perform all analyses or report where and how the data can be accessed, and reference these sources in the article. Provide the statistical code needed to reproduce the results in the article, or report whether the code is publicly accessible and if so, where	25	All datasets analyzed during the current study are publicly available summary-level data. Detailed dataset information is provided in Supplementary Table 1.
20	**Conflicts of Interest**		All authors should declare all potential conflicts of interest	25	The authors declare no competing interests.

1. Skrivankova VW, Richmond RC, Woolf BAR, Yarmolinsky J, Davies NM, Swanson SA, et al. Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomization (STROBE-MR) Statement. JAMA. 2021;under review.

2. Skrivankova VW, Richmond RC, Woolf BAR, Davies NM, Swanson SA, VanderWeele TJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomisation (STROBE-MR): Explanation and Elaboration. BMJ. 2021;375:n2233.