TID
Tobacco Induced Diseases

Supplementary file

DOI:
10.18332/tid/226573

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	1-2	Line 1-31,line 1-11
	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	4	Line 11-21
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	Line 21-28
	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:	5	Line 3-11
	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	Table1
	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	Table1
	c)	Describe measurement, quality control and selection of genetic variants	5	Line 21-30
	d)	For each exposure, outcome, and other relevant variables, describe methods of assessment and diagnostic criteria for diseases	6	Line 10-17
	e)	Provide details of ethics committee approval and participant informed consent, if relevant		N/A
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	6	Line 19-25
6	**Statistical methods: main**	Describe statistical methods and statistics used		

	analysis				
		a)	Describe how quantitative variables were handled in the analyses (i.e., scale, units, model)	5	Line 3-17
		b)	Describe how genetic variants were handled in the analyses and, if applicable, how their weights were selected	5	Line 3-17
		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	6	Line 10-17
		d)	Explain how missing data were addressed	5	Line 29-30
		e)	If applicable, indicate how multiple testing was addressed		N/A
7	**Assessment of assumptions**		Describe any methods or prior knowledge used to assess the assumptions or justify their validity	6	Line 21-25
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)	6	Line 18-25
9	**Software and pre-registration**				
		a)	Name statistical software and package(s), including version and settings used	6	Line 10-11
		b)	State whether the study protocol and details were pre-registered (as well as when and where)		N/A
	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	8	Line 2-6;Table3
		b)	Report summary statistics for phenotypic exposure(s), outcome(s), and other relevant variables (e.g. means, SDs, proportions)	8	Line 2-6;Table3
		c)	If the data sources include meta-analyses of previous studies, provide the assessments of heterogeneity across these studies		N/A
		d)	For two-sample MR: i. Provide justification of the similarity of the genetic variant-exposure associations	8	Line 2-6;Table3

			between the exposure and outcome samples ii. Provide information on the number of individuals who overlap between the exposure and outcome studies		
11	**Main results**				
		a)	Report the associations between genetic variant and exposure, and between genetic variant and outcome, preferably on an interpretable scale	7	Line 1-26, Table2
		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	7	Table2
		c)	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		N/A
		d)	Consider plots to visualize results (e.g. forest plot, scatterplot of associations between genetic variants and outcome versus between genetic variants and exposure)		Figure 2/3
12	**Assessment of assumptions**				
		a)	Report the assessment of the validity of the assumptions	8	Line 10-27
		b)	Report any additional statistics (e.g., assessments of heterogeneity across genetic variants, such as I^2, Q statistic or E-value)	8	Line 10-17 Table 4
13	**Sensitivity analyses and additional analyses**				
		a)	Report any sensitivity analyses to assess the robustness of the main results to violations of the assumptions	8	Line 10-17
		b)	Report results from other sensitivity analyses or additional analyses	8	Line 29
		c)	Report any assessment of direction of causal relationship (e.g., bidirectional MR)	8	Line 19
		d)	When relevant, report and compare with estimates from non-MR analyses		N/A
		e)	Consider additional plots to visualize results (e.g., leave-one-out analyses)		Figure 4/Figure S1
	DISCUSSION				
14	**Key results**		Summarize key results with reference to study objectives	10	Line 21-28
15	**Limitations**		Discuss limitations of the study, taking into account the validity of the IV assumptions,	11	Line 15-30

			other sources of potential bias, and imprecision. Discuss both direction and magnitude of any potential bias and any efforts to address them		
16	**Interpretation**				
		a)	Meaning: Give a cautious overall interpretation of results in the context of their limitations and in comparison with other studies	9	Line 19-24
		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	9	Line 11-19
		c)	Clinical relevance: Discuss whether the results have clinical or public policy relevance, and to what extent they inform effect sizes of possible interventions	12	Line 8-9
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	11	Line 16-17
	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	12	Line 27
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	12	Line 24
20	**Conflicts of Interest**		All authors should declare all potential conflicts of interest	12	Line 14

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.

Suppl-Table 2 Mendelian randomization results for the association between the NBL1 and smoking initiation, IEU OpenGWAS 2018 and 2019 (N=311,629 for Smoking Initiation and N=3,301 for NBL1)

Datasets			**β**	**SE**	**OR (95% CI)**	**P**
Exposure	Outcome	Reverse MR				
NBL1	Smoking initiation	MR Egger	0.02	0.02	1.02 (0.97-1.06)	0.52
		Weighted median	0.02	0.01	1.02 (1.00-1.04)	0.13
		IVW	0.02	0.01	1.02 (1.00-1.04)	0.06
		Simple mode	0.03	0.02	1.03 (0.99-1.07)	0.18
		Weighted mode	0.02	0.02	1.02 (0.98-1.06)	0.27

Abbreviations: NBL1: Neuroblastoma suppressor of tumorigenicity 1; IVW: Inverse variance weighted; CI, confidence interval; OR, odds ratio; SE, standard error; P: P-value.

Suppl-Table 3 Mendelian randomization results for the association between the smoking initiation and Neuroblastoma, IEU OpenGWAS 2013 and 2019 (N=311,629 for Smoking Initiation and N=4881 for Neuroblastoma).

Datasets			**β**	**SE**	**OR (95% CI)**	**P**
Exposure	Outcome	Forward MR				
Smoking initiation	Neuroblastoma	MR Egger	-2.39	1.90	0.09 (0.00-3.79)	0.24
		Weighted median	-0.07	0.75	0.93 (0.21-4.04)	0.92
		IVW	0.05	0.55	1.05 (0.36-3.05)	0.93
		Simple mode	1.80	1.42	6.03 (0.38-96.82)	0.24
		Weighted mode	-1.24	1.01	0.29 (0.04-2.11)	0.25

Abbreviations: IVW: Inverse variance weighted; CI, confidence interval; OR, odds ratio; SE, standard error; P: P-value.

Supplementary Figure 1 Leave-one-out analysis assessing the influence of individual SNPs on the estimated effect of NBL1 expression on smoking initiation, IEU OpenGWAS 2018 and 2019 (N=314,930).