A two sample Mendelian randomization analysis indicates no causal relationship between COVID-19 and tinnitus.

Abstract

INTRODUCTION: COVID-19's recovery period can give rise to illness that affects various body systems, these symptoms and illnesses are collectively known as Long COVID. Tinnitus has been observed to be one of the many possible symptoms that arise from Long COVID, but a causal pathway has yet to be tested for. Two sample Mendelian randomization is an inexpensive, statistically powerful, and relatively robust method under the right assumptions that uses data from pre-existing genome-wide association studies (GWAS) to infer causality from observations based on genetics.

METHODS: Data used for COVID-19 defined three phenotypes with their own respective dataset: critical illness, hospitalization, and reported infection. Data used for tinnitus defined only one phenotype. Both datasets only contained individuals of European ancestry. Instruments were obtained by filtering COVID-19 SNPs that did not meet a significance threshold of $1*10^{-6}$ and subsequently clumping the remaining SNPs to remove SNPs in high linkage disequilibrium. Tinnitus data were filtered to keep only shared SNPs with the remaining COVID SNPs, and then the datasets were harmonized. The Egger, Inverse Variance Weighted with Fixed Effects, Weighted Median, Weighted Mode, and RadialMR methods were used to determine causality and identify outliers. The analyses were facilitated by the R software and its TwoSampleMR package. R Software was also used to visualize results into scatter, forest, leave-one-out, and Radial MR plots.

RESULTS: Egger estimates for critical illness, hospitalization, and reported infection returned p-values of *0.092*, *0.135*, and *0.23*, respectively. Heterogeneity tests of the Egger estimates returned p-values of *0.65*, *0.85*, and *0.129*, respectively. Weighted median estimates for critical illness, hospitalization, and reported infection returned p-values of *0.115*, *0.144*, and *0.43*, respectively. Weighted mode estimates for critical illness, hospitalization, and reported infection returned p-values of *0.132*, *0.192*, and *0.47*, respectively. Inverse variance weighted with fixed effects estimates for critical illness, hospitalization, and reported infection returned p-values of *0.55*, *0.92*, and *0.68*, respectively. Heterogeneity tests of the IVW estimates returned p-values of *0.50*, *0.71*, and *0.064*, respectively. Tests for horizontal pleiotropy on the analyses of critical illness, hospitalization, and reported infection returned p-values of *0.086*, *0.064*, and *0.099*, respectively. None of the RadialMR plots suggested any outliers.

DISCUSSION: These results signify that tinnitus is most likely not causally related to any of the three COVID-19 phenotypes defined for this study. Instead, it is possibly more likely that if tinnitus does arise in someone suffering from Long COVID, then it is as a result of the combination of age-related hearing loss and/or environmental factors along with COVID-related complications. Further analyses, especially using different traits/phenotypes and demographically varied samples could provide more generalizable results and give a more comprehensive insight into COVID-19 and/or tinnitus.

Introduction

Overview of the Exposures and Outcome of Interest

COVID-19 is an infectious disease known for its effects on the respiratory system as well as the multisystemic conditions that arise in its recovery period known as Long COVID. One of these multisystemic conditions has been observed to be tinnitus, phantom noise in the ears.[1] Tinnitus is almost always a symptom of an underlying condition, and its exact origin is hard to pinpoint due to its polygenic nature. Risk factors range from psychiatric distress to cardiovascular disease, among numerous others. Classic risk factors for tinnitus are age-related hearing loss as well as frequent environmental exposure to loud noise.[2]

Considering the polygenic nature of tinnitus along with observations of tinnitus manifesting from Long COVID, a multisystemic condition, it is possible that COVID-19 is causally related to tinnitus through some sort of biological pathway. One area of particular interest could be neuroinflammation as a result of Long COVID. Long COVID has a wide array of neurocognitive impacts that result in cognitive decline, cognitive dysfunction and even damaging neurons in a similar manner to Alzheimer's disease. The pathway of Long COVID's neuroinflammation is believed to lead to hearing loss, vertigo, and tinnitus as audiovestibular manifestations of Long COVID.[1]

Rationale for the Use of Mendelian Randomization

Mendelian randomization is a method that could be used to determine the validity of this hypothesis. Mendelian randomization is based on the idea that genetic variants are randomly and independently assigned to an organism at conception and that if there is a genetic variant that is associated with an exposure then it can be used as a sort of reference point for estimating the effect size of the exposure on an outcome.

Mendelian randomization is an effective method for this study because it is simple to conduct with data from genome wide association studies. While Mendelian randomization is not the definitive way to determine causality, it is very statistically powerful and helpful in triangulating causality when done correctly and under the right assumptions. Those assumptions are known as the three core assumptions of Mendelian randomization and are as follows: First, that a genetic

variant is associated with an exposure. Second, that a genetic variant is independent of confounding variables, and that a genetic variant is associated with an outcome only through the exposure.[3] A variant in this case is usually a SNP, a single nucleotide polymorphism.

This study aims to determine if there is a causal relationship between COVID-19 and tinnitus. Three COVID-19 phenotypes will be used as the exposures and tinnitus as the outcome in two sample Mendelian randomization analyses. Mendelian randomization carries its own assumptions, but if those are not violated and the methods used are robust then the study should have enough statistical power to estimate causal effects of COVID-19 on tinnitus.

Methods

Methods Used by the GWAS of Interest

The Human Genetics Initiative's (HGI) genome wide association study on COVID-19 defined and studied three different phenotypes: critically ill COVID-19, hospitalized COVID-19, and reported SARS-CoV-2 infection. The critically ill COVID-19 group contained $N = 6,179$ cases and $N = 1,483,780$ controls. The hospitalized COVID-19 group contained $N = 13,641$ cases and $N = 2,070,709$ controls. The reported SARS-CoV-2 infection group contained $N = 49,562$ cases and $N = 1,770,206$ controls. A majority of the individuals in this study were of European ancestry. The three phenotypes defined have overlap in their samples.[4]

Bhatt et al.'s genome wide association study on tinnitus started with the 500,000 individuals in the UK Biobank, but then filtered out the individuals that responded with "do not know" or "prefer not to answer" to the question, "Do you get or have you had noises (such as ringing or buzzing) in your head or in one or both ears that last for more than five minutes at a time?" on the UK Biobank questionnaire. They further filtered out those who did not report 'British White', 'Irish White', or similar ancestry/ethnicity. After this filtering process, they ended up having $N = 132,438$ individuals and, of those individuals, $N = 38,525$ cases based on questionnaire answers for use in the rest of their analyses.[5]

The HGI's data were used as data for the exposure in this analysis. Bhatt et al.'s data were used as data for the outcome in this analysis. This is to fit in with two sample Mendelian randomization's framework.

Comments Before Main Analysis

Before proceeding with the analysis, whether or not the core assumptions of Mendelian randomization are violated is worth considering. The first core assumption of Mendelian randomization is proven by the data only containing significant SNPs from genome wide association studies, since it indicates that the SNPs are associated with the respective phenotype. The second and third core assumptions of Mendelian randomization are impossible to prove and thus the robustness of the analysis should theoretically account for this.

This analysis was conducted using R 4.3.0, R Studio software, and its TwoSampleMR, tidyverse, and RadialMR packages were used to facilitate the analyses along with data handling.

Overview of R Script's Functions (Script Available on GitHub)
The respective COVID-19 and tinnitus datasets are imported into R. Then the TwoSampleMR package's formatting function is used on both of the datasets to make them comply with the rest of the TwoSampleMR package's functions. The formatted exposure data is filtered to remove SNPs that do not meet a significance threshold of $1*10^{-6}$ and subsequently clumped to remove SNPs in high linkage disequilibrium. Afterwards, the formatted outcome data is filtered to only keep shared SNPs with the exposure data. The clumped exposure data and the filtered outcome data are then harmonized to make the alleles and allele frequencies of each dataset match up with each other. Mendelian randomization is then performed using the Egger, Weighted Median, Weighted Mode, and Inverse Variance Weighted with fixed effects methods. Results of this process were recorded on tables and visualized via scatter, forest, and leave-one-out plots using the respective functions for those plots. Afterwards, tests for Heterogeneity and Horizontal Pleiotropy are conducted. Finally, Radial MR is performed and its respective plot is created to see if there are outliers in the analysis.

This process is done separately for each of the three COVID-19 phenotypes defined for this study.

Overview of Sensitivity Analyses Used
The Egger, Inverse Variance Weighted, Weighted Mode, and Weighted Median estimate sensitivity analyses were used to return various plots and determine if there was any causal relationship between COVID and tinnitus along with helping to identify factors like horizontal pleiotropy, heterogeneity, and other biases.

The Egger estimate assumes that a SNP's effect on the exposure is independent of any possible Horizontal Pleiotropy and then estimates the causal effect of SNPs on an exposure and outcome. The regression line produced from this method is not fixed to pass through the origin.[6]

The inverse variance weighted estimate with fixed effects assumes that there is no Horizontal Pleiotropy and is essentially a linear regression of SNP-exposure effect on SNP-outcome effects with weights based on the inverse variance of their effects on the outcome. The regression line produced from this method is fixed to pass through the origin.[6]

The weighted median estimate takes the median effect size while allowing SNPs with larger effect sizes to have more weight in the estimate based on the inverse variance of its association with the outcome.[6]

The weighted mode estimate groups up SNPs with similar effect sizes. The group with the largest number of SNPs is then used as the basis of this estimate. Like the other weighted estimates, SNPs are weighted by the inverse variance of their effect on the outcome.[6]

Additionally, Radial MR was used. Radial MR uses a lot of the same methods as the inverse variance weighted estimator, but is especially effective at identifying outliers. If there were outliers present, Radial MR would make it much simpler to identify and remove them.[6]

Results

Results for COVID-19 Critical Illness

The Egger estimate for critical illness returned a p-value of *0.092*. The heterogeneity test of this Egger estimate returned a p-value of *0.65*. The weighted median estimate for critical illness returned a p-value of *0.115*. The weighted mode estimate for critical illness returned a p-value of *0.132*. The inverse variance weighted with fixed effects estimate for critical illness returned a p-value of *0.55*. The heterogeneity test of this IVW estimate returned a p-values of *0.50*. The test for horizontal pleiotropy on the analyses of critical illness returned a p-value of *0.086*. The RadialMR plot for critical illness did not suggest any outliers.

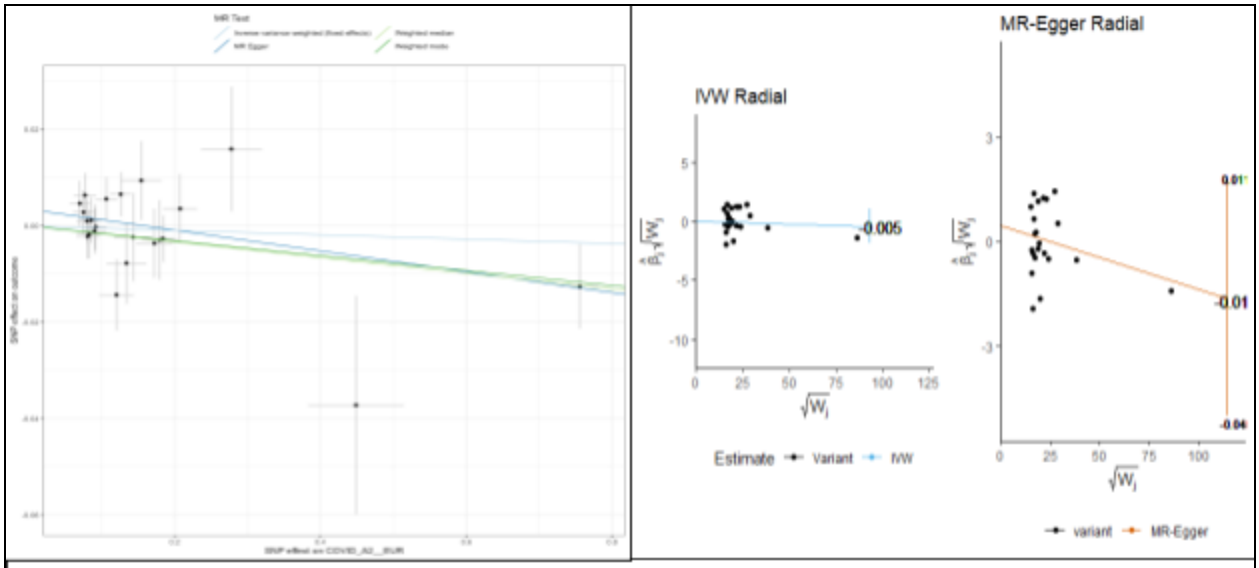

Fig. 1. Scatter Plot (Left) and Radial MR Plot (Right) of Critically ill COVID-19 on Tinnitus
(Left) Each dot represents a SNPs effect size point estimate and confidence interval. The regression line represents the overall relationship between SNP effect on exposure (x-axis) and SNP effect on outcome (y-axis). Based on the regression line, this plot suggest no significant causal relationship between COVID-19 related critical illness and tinnitus.

(Right) Each dot represents a SNP plotted by its weight (x-axis) and by its weight multiplied by the estimated effect size (y-axis). The regression line also displays a confidence interval for the overall effect size. Since the confidence interval includes 0, this plot suggests no significant causal relationship between COVID-19 related critical illness and tinnitus.

Results for COVID-19 Hospitalization

The Egger estimate for hospitalization returned a p-value of *0.135*. The heterogeneity test of this Egger estimate returned a p-value of *0.85*. The weighted median estimate for hospitalization returned a p-value of *0.144*. The weighted mode estimate for hospitalization returned a p-value of *0.192*. The inverse variance weighted with fixed effects estimate for hospitalization returned a p-value of *0.92*. The heterogeneity test of this IVW estimate returned a p-value of *0.71*. The test for horizontal pleiotropy on the analyses of hospitalization returned a p-value of *0.064*. The RadialMR plot for hospitalization did not suggest any outliers.

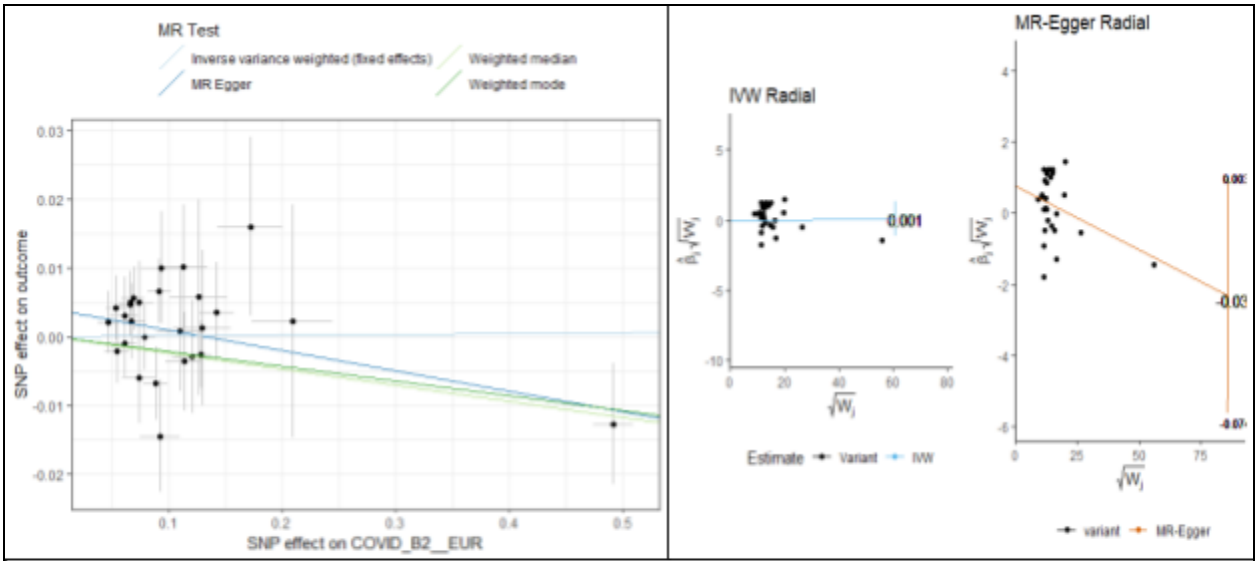

Fig. 2. Scatter Plot (Left) and Radial MR Plot (Right) of Hospitalized COVID-19 on Tinnitus
Similarly to Fig. 1 there is no significant evidence for a causal relationship between COVID-19 hospitalization and tinnitus in either of these plots.

See Fig. 1 for more details on the types of plots used.

Results for Reported SARS-CoV-2 Infection

The Egger estimate for reported infection returned a p-value of *0.23*. The heterogeneity test of this Egger estimate returned a p-value of *0.129*. The weighted median estimate for reported infection returned a p-value of *0.43*. The weighted mode estimate for reported infection returned a p-value of *0.47*. The inverse variance weighted with fixed effects estimate for reported infection returned a p-value of *0.68*. The heterogeneity test of this IVW estimate returned a p-value of *0.064*. The test for horizontal pleiotropy on the analyses of reported infection returned a p-value of *0.099*. The RadialMR plot for reported infection did not suggest any outliers.

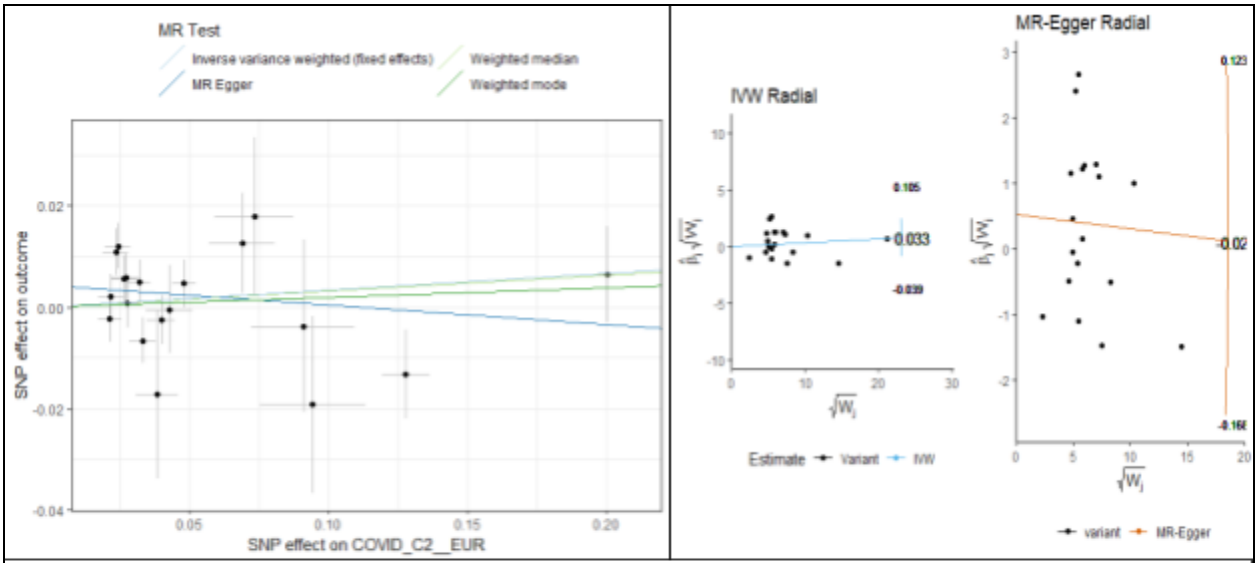

Fig. 3. Scatter Plot (Left) and Radial MR Plot (Right) of Reported SARS-CoV-2 Infection on Tinnitus
Similarly to Fig. 1 and Fig. 2 there is no significant evidence for a causal relationship between reported COVID-19 infection and tinnitus in either of these plots.

See Fig. 1 for more details on the types of plots used.

Exposure	Outcome	Method	β	SE	P
COVID-19 critical illness	tinnitus	MR Egger	-0.02158830	0.012212878	0.0924448
COVID-19 critical illness	tinnitus	Weighted Median	-0.016348848	0.010386487	0.1154871
COVID-19 critical illness	tinnitus	Weighted Mode	-0.015444682	0.009855375	0.1320294
COVID-19 critical illness	tinnitus	Inverse Variance Weighted (Fixed Effects)	-0.004561528	0.007728582	0.5550462
COVID-19 hospitalization	tinnitus	MR Egger	-0.0295318728	0.01911960	0.1350119
COVID-19 hospitalization	tinnitus	Weighted Median	-0.02324331381	0.01606726	0.1447191
COVID-19 hospitalization	tinnitus	Weighted Mode	-0.0214221746	0.01600454	0.1923193
COVID-19 hospitalization	tinnitus	Inverse Variance Weighted (Fixed Effects)	0.0009976225	0.01082898	0.9265985
Reported SARS-CoV-2 infection	tinnitus	MR Egger	0.03302645	0.02761559	0.2317220
Reported SARS-CoV-2 infection	tinnitus	Weighted Median	0.03095213	0.03937767	0.4318484
Reported SARS-CoV-2 infection	tinnitus	Weighted Mode	-0.03852774	0.05246599	0.4727513
Reported SARS-CoV-2 infection	tinnitus	Inverse Variance Weighted (Fixed Effects)	0.01906795	0.04589756	0.6827260

Table 1. Betas, Standard Errors, and P-Values for the Sensitivity Methods Used
Using the standard significance threshold of 0.05, there are no significant results from any of the methods used with any of the exposures.

Exposure	Heterogeneity Egger P	Heterogeneity IVW P	Pleiotropy P
COVID-19 Critical Illness	0.6548815	0.5079579	0.08696427
COVID-19 Hospitalization	0.8500624	0.7119360	0.06406207
Reported SARS-CoV-2 Infection	0.12924408	0.06420051	0.09998522

Table 2. P-Values of Heterogeneity and Pleiotropy Tests
Using the standard significance threshold of 0.05, there is no significant evidence for Heterogeneity or Horizontal Pleiotropy.

Discussion

Summation

This study aimed to establish causal relationships between three COVID-19 phenotypes (Critical Illness, COVID-related Hospitalization, Reported SARS-CoV2 Infection) and tinnitus using the two sample Mendelian randomization method to use genetics as a way of inferring causality based on GWAS data. R software facilitated handling the GWAS data, performing the MR Egger, Inverse Variance Weighted with Fixed Effects, Weighted Median, and Weighted Mode estimator methods, and the visualization of the aforementioned methods. We found that the results of all of these tests returned insignificant p-values, no indication of outliers, no indication of significant heterogeneity, and no indication of significant horizontal pleiotropy.

Limitations

Lack of generalizability is a concern due to the GWAS used for this analysis only containing individuals of European ancestry and thus the results may not be applicable to other groups.

Also Mendelian randomization is not the be-all end-all for determining causality. It is possible that through further analysis, a causal relationship between COVID-19 and tinnitus can be triangulated.

Conclusions

Based on the results, there seems to be no causal relationship between COVID-19 and tinnitus. No significant horizontal pleiotropy, heterogeneity, or outliers are present either so those are not significant sources of bias. It can be assumed there is only observational evidence of such a pathway. These results can help further the understanding of what COVID-19 is actually causally related to and what it is not while also being another contribution to demystifying the polygenic nature of tinnitus.

References

[1] Davis, H. E., McCorkell, L., Vogel, J. M. & Topol, E. J. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol 1–14 (2023) doi:10.1038/s41579-022-00846-2.

[2] Jarach, C. M. et al. Global Prevalence and Incidence of Tinnitus. Jama Neurol 79, 888–900 (2022).

[3] Sanderson, E., Glymour, M.M., Holmes, M.V. *et al.* Mendelian randomization. *Nat Rev Methods Primers* 2, 6 (2022). https://doi.org/10.1038/s43586-021-00092-5.

[4] Initiative, C.-19 H. G. et al. Mapping the human genetic architecture of COVID-19. Nature 600, 472–477 (2021).

[5] Bhatt, I. S., Wilson, N., Dias, R. & Torkamani, A. A genome-wide association study of tinnitus reveals shared genetic links to neuropsychiatric disorders. Sci Rep-uk 12, 22511 (2022).

[6] Hemani, G., Zheng, J., Elsworth, B., Wade, K. H., Haberland, V., Baird, D., Laurin, C., Burgess, S., Bowden, J., Langdon, R., Tan, V. Y., Yarmolinsky, J., Shihab, H. A., Timpson, N. J., Evans, D. M., Relton, C., Martin, R. M., Davey Smith, G., Gaunt, T. R., & Haycock, P. C. (2018). The MR-Base platform supports systematic causal inference across the human phenome. *eLife*, *7*, e34408. https://doi.org/10.7554/eLife.34408.

[7] Davies N M, Holmes M V, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians *BMJ* 2018; 362 :k601 doi:10.1136/bmj.k601.

[8] Uffelmann, E., Huang, Q.Q., Munung, N.S. *et al.* Genome-wide association studies. *Nat Rev Methods Primers* 1, 59 (2021). https://doi.org/10.1038/s43586-021-00056-9.