Yu *et al. Infectious Agents and Cancer* (2024) 19:60
https://doi.org/10.1186/s13027-024-00626-y

Infectious Agents and Cancer

MATTERS ARISING Open Access

Enhancing the robustness of Mendelian randomization studies: lessons from a two-sample analysis of viral infections and colorectal cancer

Tianfei Yu[1,2*], Jinyong Xia[1,2], Haichang Yin[1,2], Nana Yi[1,2], Lanlan Zhang[1,2] and Ming Li[3*]

Abstract

This Matters Arising article critically examines the study "Genetic susceptibility association between viral infection and colorectal cancer risk: a two-sample Mendelian randomization analysis" by Li et al., highlighting both its contributions and methodological limitations. Their study employed two-sample Mendelian randomization (MR) to explore potential causal links between viral infections and colorectal cancer (CRC), identifying significant associations with infections such as herpes simplex virus and measles. However, several aspects of the methodology warrant scrutiny, including the relaxation of instrumental variable selection thresholds, the handling of potential pleiotropy, and the interpretation of biologically implausible findings. While leveraging advanced MR techniques such as MR-RAPS, cML, ConMix, and dIVW to address challenges like pleiotropy and weak instruments, the study encountered issues related to heterogeneity, insufficient exploration of biological plausibility, and a lack of detailed reporting on instrumental variable (IV) selection and preprocessing. This Matters Arising calls for more rigorous sensitivity analyses, improved transparency in IV selection criteria and harmonization of genome-wide association study (GWAS) datasets, particularly in addressing differences between self-reported and clinically diagnosed infections. Additionally, the Matters Arising article calls for a deeper exploration of biological mechanisms, such as the role of immune modulation and inflammation, to better interpret the observed associations. By addressing these limitations, future MR studies can enhance methodological rigor, improve reproducibility, and provide more robust insights into the causal pathways linking viral infections to CRC risk.

Keywords Mendelian randomization, Colorectal cancer, Viral infections, Instrumental variables, Genetic epidemiology

*Correspondence:
Tianfei Yu
yutianfei2001@163.com
Ming Li
fionalee629@163.com
[1]Department of Biotechnology, College of Life Science and Agriculture Forestry, Qiqihar University, Qiqihar 161006, China
[2]Department of Computer Science and Technology, College of Computer and Control Engineering, Qiqihar University, Qiqihar 161006, China
[3]Heilongjiang Provincial Key Laboratory of Resistance Gene Engineering and Protection of Biodiversity in Cold Areas, Qiqihar University, Qiqihar 161006, China

© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

Introduction

Mendelian randomization (MR) has emerged as a powerful tool in genetic epidemiology, offering a unique approach to infer causal relationships between risk factors and health outcomes by using genetic variants as instrumental variables (IVs) [1]. This method has gained widespread acceptance due to its ability to mitigate confounding and reverse causation, issues that often plague observational studies. Two-sample MR (TSMR) further enhances analytical power by using separate genome-wide association study (GWAS) datasets for exposures and outcomes, allowing researchers to investigate complex causal pathways [2]. The study "Genetic susceptibility association between viral infection and colorectal cancer risk: a two-sample Mendelian randomization analysis" by Li et al. contributed to this growing body of literature by investigating the genetic susceptibility associations between various viral infections and colorectal cancer (CRC) [3]. Li et al. identified significant associations between specific viral infections and CRC risk, including herpes simplex virus (HSV), herpes zoster, viral hepatitis, and infectious mononucleosis. However, their methodology raised concerns regarding the relaxation of IV selection thresholds ($P<5\times10^{-6}$), limited sensitivity analyses, and insufficient exploration of biological mechanisms underlying these associations. Such limitations could compromise the reliability of their findings, particularly in addressing the heterogeneity across CRC subtypes and potential biases from weak instruments.

This Matters Arising provides a detailed critique of the study, addressing issues such as IV selection thresholds, the handling of pleiotropy, subgroup heterogeneity, and the transparency of reporting. Additionally, it emphasizes the importance of exploring biological plausibility, particularly the role of immune responses and inflammation, to better contextualize the observed associations. By proposing methodological improvements and highlighting unresolved questions, this critique aims to guide future MR studies in addressing similar challenges.

IV selection and thresholds

IV selection is a critical aspect of MR studies, as the validity of the causal inference hinges on the strength and relevance of the selected IVs. The criteria for IV selection in Li et al.'s study followed these steps: (1) Extracting summary statistics for genetic variants associated with viral infections (exposures) from large GWAS datasets, such as 23andMe and FinnGen; (2) Filtering variants based on the relaxed $P<5\times10^{-6}$ threshold; (3) Clumping variants to ensure independence (linkage disequilibrium threshold: $r^2<0.001$, within a 10,000 kb window); and (4) Excluding SNPs with F-statistics less than 10 to minimize weak instrument bias. In their study, Li et al. opted to relax the conventional genome-wide significance threshold for selecting IVs from $P<5\times10^{-8}$ to a more lenient threshold of $P<5\times10^{-6}$ [3]. This decision was likely made to ensure an adequate number of IVs for the MR analysis, given the limited availability of strongly associated genetic variants.

While these measures aim to enhance analytical power, the use of a relaxed threshold without sufficient justification raises concerns. Relaxing the threshold increases the likelihood of selecting weak IVs, which can lead to biased causal estimates, particularly if pleiotropic effects are present. Li et al. mentioned excluding SNPs with low F-statistics, but they did not provide detailed information on the distribution of F-statistics across the selected IVs, nor did they conduct sensitivity analyses to evaluate the impact of this relaxation.

To address this limitation, the decision to relax IV selection thresholds should be accompanied by robust sensitivity analyses, including Steiger directionality tests, Leave-one-out analyses, and Cochran's Q tests. The Steiger test can confirm the correct causal directionality [4], while Leave-one-out analysis assesses the influence of individual IVs on overall estimates [5]. Cochran's Q test evaluates heterogeneity among IVs, identifying potential violations of MR assumptions [6]. By implementing these analyses, researchers can validate the robustness of causal estimates despite relaxed thresholds.

Furthermore, harmonization procedures for GWAS datasets should be explicitly reported [7]. These include aligning alleles to the same reference strand, excluding palindromic SNPs to prevent strand ambiguity, and ensuring consistent IV presence in both exposure and outcome datasets. Clear documentation of these steps enhances transparency and facilitates reproducibility in MR studies.

To improve the robustness of future MR studies, researchers should adhere to the conventional threshold of of $P<5\times10^{-8}$ unless strong justification exists for relaxation. If a relaxed threshold is used, comprehensive sensitivity analyses and detailed reporting of IV characteristics, including F-statistics, are essential to maintain credibility. Additionally, leveraging larger GWAS datasets or employing polygenic risk scores may help reduce the need for threshold relaxation by increasing statistical power.

Potential pleiotropy and confounding

MR relies on the assumption that the selected genetic variants influence the outcome solely through the exposure of interest and not through other pathways—a condition known as no pleiotropy [8]. Pleiotropy occurs when a genetic variant affects multiple traits, potentially introducing bias if the pleiotropic effects are related to the outcome through pathways other than the exposure. Addressing pleiotropy is critical for ensuring the validity

of causal inferences in MR studies. In their analysis, Li et al. utilized several MR methods, including MR-Egger and MR-PRESSO, to account for horizontal pleiotropy [3]. The study reported that no significant pleiotropy was detected, which the authors interpret as evidence of the robustness of their findings.

However, MR-Egger has significant limitations. It assumes that pleiotropic effects are uncorrelated with the strength of the IV-exposure association (InSIDE assumption), which may not always hold in real-world data. Furthermore, MR-Egger is highly sensitive to weak instruments, potentially producing biased estimates when the F-statistics of IVs are low [9]. MR-PRESSO, on the other hand, identifies and removes outlier variants contributing to horizontal pleiotropy, providing corrected causal estimates [10]. While effective in reducing the influence of pleiotropic outliers, MR-PRESSO assumes that the remaining IVs are valid and does not address scenarios where pleiotropy is widespread across most IVs. Additionally, the outlier removal process can reduce the number of IVs, potentially affecting statistical power and precision.

To address these limitations, researchers should consider incorporating additional methods to enhance the robustness of MR analyses. Weighted median and mode-based estimators are promising alternatives. The weighted median method provides valid causal estimates if at least 50% of the weight comes from valid instruments, offering protection against invalid IVs [11]. The mode-based estimator, which relies on identifying the most common causal effect among IVs, is effective even when a majority of the IVs are pleiotropic, provided that the mode represents the true causal effect [12]. These methods are less sensitive to violations of the no-pleiotropy assumption and complement MR-Egger and MR-PRESSO by addressing their weaknesses.

Beyond addressing pleiotropy, confounding factors such as population stratification and linkage disequilibrium (LD) must also be carefully managed. Population stratification, where allele frequencies vary across ancestral groups, can introduce spurious associations if not properly accounted for [13]. To mitigate this, replication studies in diverse populations and the inclusion of principal components in single-sample MR analyses are recommended. In two-sample MR, where covariates cannot be directly included, researchers should focus on harmonizing datasets and testing consistency across subpopulations. LD, where selected IVs are correlated with other variants influencing the outcome through unrelated pathways, presents another challenge [14]. To minimize LD-related bias, IVs should be clumped to ensure independence and evaluated for their individual contributions to the causal estimate.

Future MR studies should consider employing more advanced methods to address pleiotropy and confounding. Multivariable Mendelian randomization (MVMR) is one such approach that allows for the simultaneous consideration of multiple exposures, helping to disentangle the effects of correlated risk factors [15]. While including principal components as covariates in MR regression can account for population stratification, this approach is feasible in single-sample MR studies only. In two-sample MR studies, covariates cannot be directly included in the two-stage least squares (2SLS) regression due to data constraints, so researchers should instead aim to conduct replication studies in different populations to help validate findings and address potential confounding [16]. The sensitivity analyses that account for population stratification, such as including principal components as covariates in the analysis, can help mitigate the risk of confounding [17]. Researchers should also consider conducting replication studies in different populations or using different datasets to validate the findings and ensure they are not driven by unrecognized confounders.

Interpretation of results and biological plausibility

The results of the study by Li et al. indicated associations between specific viral infections and CRC risk, such as the protective effect of HSV on CRC and the increased risk associated with measles virus infection [3]. While these associations were statistically significant, their biological plausibility requires further exploration to strengthen the scientific validity of the findings.

For example, the finding that HSV is associated with a reduced risk of CRC is counterintuitive, as chronic viral infections are generally considered pro-inflammatory and potentially carcinogenic [18]. However, HSV has demonstrated oncolytic properties in preclinical studies, selectively infecting and destroying tumor cells, particularly in the central nervous system [19]. Oncolytic herpesviruses engineered to target cancer cells have shown promise in clinical trials, likely due to their ability to activate anti-tumor immune responses and induce direct cytotoxic effects [20]. In the context of CRC, the neurotropic nature of HSV might limit its direct interaction with colorectal tissues, potentially explaining the observed protective effect. Experimental studies using colorectal cell lines and animal models are needed to confirm these mechanisms.

Similarly, the study reported that measles virus infection was associated with an increased risk of colon cancer. Measles virus is known to induce immune suppression, which could impair tumor surveillance and allow cancerous cells to proliferate unchecked [21]. A study has highlighted the overexpression of the measles virus receptor (PVRL4) in colorectal cancer tissues, suggesting a plausible link between viral entry and tumor progression [22].

This pathway warrants further investigation to clarify the role of measles virus in CRC development.

For herpes zoster, the study identified an increased risk of rectal cancer and a protective association with colon cancer. This dichotomy may reflect site-specific immune responses and microenvironmental factors. Varicella-zoster virus (VZV) replication in $CD8^+$ T cells has been linked to localized inflammation, particularly in the rectum, which could increase cancer risk [23]. Conversely, high $CD8^+$ T cell infiltration in the colon is associated with better cancer prognosis, potentially mitigating the carcinogenic effects of VZV in this region [24]. These contrasting effects underscore the importance of conducting site-specific analyses to disentangle the underlying mechanisms.

Infectious mononucleosis, caused by Epstein-Barr virus (EBV), was associated with a reduced risk of both colon and rectal cancer [25]. EBV has been implicated in hematological malignancies but appears to play a less prominent role in solid tumors [26]. Experimental evidence suggests that EBV infection can modulate the tumor immune microenvironment, enhancing anti-tumor immunity in certain contexts [27]. This immunomodulatory effect could explain the observed protective association, but further research is required to delineate its role in CRC risk.

Finally, the study identified blood metabolites linked to viral infections and CRC risk. For instance, the cysteinylglycine-to-taurine ratio was found to be inversely associated with rectal cancer risk. Cysteine plays a critical role in maintaining redox balance and immune function, while taurine has demonstrated anti-inflammatory properties in preclinical models [28]. These metabolites may act synergistically to counteract inflammation-driven tumorigenesis. Future research integrating metabolomic and genetic data may help uncover novel biomarkers and therapeutic targets for CRC.

By providing a deeper exploration of these biological mechanisms, researchers can better contextualize their findings and enhance the credibility of MR studies.

Heterogeneity and subgroup analysis

Heterogeneity in effect estimates is a common issue in epidemiological studies, including MR analyses [29]. In the study by Li et al., subgroup analyses were conducted to explore the associations between viral infections and different CRC subtypes, such as colon cancer and rectal cancer [3]. These analyses are valuable, as they help to identify potential differences in the impact of exposures across various forms of CRC. However, the study did not adequately explore potential sources of heterogeneity, nor did it address how these sources might influence the reliability of the findings.

One potential source of heterogeneity is the genetic diversity across study populations. Differences in allele frequencies and LD patterns across populations may lead to population-specific effects, limiting the generalizability of the findings. To mitigate this, rreplication studies in diverse populations with varying genetic backgrounds are necessary to validate the findings and reduce population-specific biases.

Another source of heterogeneity may arise from differences in study designs. For instance, GWAS datasets such as 23andMe and FinnGen use distinct approaches to measure exposures: 23andMe relies on self-reported infection history [30], which is prone to recall bias, while FinnGen utilizes clinically diagnosed cases [31]. Such discrepancies could result in variations in effect estimates and should be addressed by harmonizing definitions of exposures and outcomes across datasets. Stratified analyses by dataset source can also help disentangle dataset-specific biases.

Heterogeneity might also reflect anatomical and biological differences between CRC subtypes. Colon and rectal cancers have distinct etiologies, immune infiltration patterns, and microbiome compositions, all of which could modify the effects of viral infections. For example, the protective association observed for herpes zoster with colon cancer, contrasted with its increased risk for rectal cancer, suggests site-specific factors influencing the tumor microenvironment. Stratifying analyses by tumor location and incorporating covariates such as inflammatory biomarkers or microbiome profiles could provide more refined insights into these differences.

The authors mentioned the use of Cochran's Q statistic to assess heterogeneity, but this test alone is not sufficient to address the complexities of the data. Cochran's Q evaluates overall heterogeneity but does not identify specific sources or moderators of variability [32, 33]. Complementing this approach with meta-regression techniques [34] to investigate moderators, such as age, sex, or genetic ancestry, could help pinpoint factors contributing to heterogeneity. Random-effects models [35] could also account for variability across subgroups, providing more reliable estimates in the presence of significant heterogeneity.

Additionally, the lack of harmonization in defining CRC subtypes or outcomes across datasets may have introduced additional variability. Future studies should ensure consistent definitions of outcomes and exposures, particularly when combining data from multiple sources. This could involve applying standardized criteria for CRC subtyping, such as TNM staging, and ensuring consistent diagnostic thresholds for viral infections.

To better account for heterogeneity, future MR studies should consider using more sophisticated methods, such as random-effects models, which allow for variability

across subgroups and provide a more nuanced understanding of the data [35]. Stratified analyses based on other relevant factors (e.g., age, sex, genetic ancestry) should also be conducted to assess whether the effects of viral infections on CRC risk differ across different population subgroups. Additionally, interaction terms could be included in the MR analysis to evaluate potential effect modification by these factors [36]. Exploring the biological differences between CRC subtypes in greater detail, particularly in relation to viral infections, would also enhance the study's contributions to our understanding of cancer etiology.

Reporting and transparency

Transparency in reporting is essential for the reproducibility and credibility of scientific research. While the study by Li et al. presented valuable data and results, there are areas where reporting could be improved. For example, the specific criteria for selecting IVs, the steps taken during data preprocessing, and the results of the sensitivity analyses were not thoroughly described. This lack of detail makes it challenging for readers to fully assess the validity of the study's findings or to replicate the analysis in future research.

Moreover, the study would benefit from a more comprehensive discussion of the limitations of MR analysis, particularly in the context of the specific methodological choices made by the authors. For instance, the decision to relax the IV selection threshold and the potential for weak instrument bias should be discussed more openly. Additionally, the study did not address the limitations of using self-reported infection history data from 23andMe, which could introduce recall bias or other forms of misclassification [37].

The study also lacked sufficient detail on the sensitivity analyses conducted. While the authors mentioned using MR-Egger and MR-PRESSO to account for pleiotropy, they did not provide the specific results of these analyses or discuss how they were interpreted. Furthermore, the study did not explore alternative MR methods that could provide additional robustness to the findings, such as the use of weighted median or mode-based estimators.

To enhance transparency, future MR studies should provide a detailed description of their data preprocessing steps, criteria for IV selection, and the results of all sensitivity analyses conducted. This could be achieved by including supplementary materials or appendices that contain this information. Additionally, the study should follow established reporting guidelines for MR studies, such as the STROBE-MR checklist, to ensure that all relevant details are reported comprehensively [38]. By being more transparent about the methodological choices and potential limitations, researchers can improve the credibility of their findings and facilitate replication by others in the field.

Application of novel MR techniques

The study by Li et al. employed several advanced MR techniques, such as cML, ConMix, MR-RAPS, and dIVW, to assess the causal effects of viral infections on CRC risk [3]. These methods are valuable for addressing some of the challenges inherent in MR analysis, such as pleiotropy and weak instrument bias.

The cML and ConMix methods were chosen due to their ability to handle both correlated and uncorrelated pleiotropy, which occurs when genetic variants affect the outcome through pathways independent of the exposure. cML and ConMix are particularly useful in scenarios where the IVs are suspected to include pleiotropic variants, as it models the distribution of pleiotropic effects across all IVs rather than assuming all IVs are valid. This feature allows for a more robust causal estimate even in the presence of pleiotropic bias. However, cML and ConMix assume that the distribution of pleiotropic effects is symmetric, which may not hold in all datasets.

MR-RAPS, on the other hand, was selected for its robustness against weak instrument bias, which can distort causal estimates if the IVs are not strongly associated with the exposure. MR-RAPS is beneficial when the F-statistic is close to the threshold for weak instruments, as it uses a robust adjusted profile score that reduces the influence of weak IVs. Nonetheless, MR-RAPS requires careful tuning of parameters, and its effectiveness may vary depending on the strength and number of IVs used.

The dIVW approach provides an alternative to the standard IVW method by debiasing the variance of IV estimates, making it less sensitive to outliers and more reliable in the presence of heterogeneity. However, dIVW assumes that the majority of IVs are valid, which may limit its applicability if pleiotropy is widespread.

However, the study did not fully explore the potential benefits and limitations of these novel techniques, nor did it discuss the rationale for selecting these specific methods over others. cML and ConMix, for example, is designed to handle correlated pleiotropy by modeling the distribution of pleiotropic effects across genetic variants [39, 40]. MR-RAPS, on the other hand, provides a robust adjusted profile score that is less sensitive to outliers and weak instruments [41]. While these methods offer significant advantages, they also have limitations that should be acknowledged. For instance, cML and ConMix assumes that pleiotropic effects are normally distributed, which may not always be the case in real-world data [42, 43]. Similarly, MR-RAPS requires careful selection of tuning parameters, and its performance can be sensitive to the choice of these parameters [44].

The study would benefit from a more detailed discussion of the strengths and weaknesses of these novel MR techniques, as well as a justification for their use in the context of this particular analysis. Additionally, the study could explore the potential for combining multiple MR methods to provide a more comprehensive assessment of causal relationships, as different methods may be better suited to addressing different sources of bias.

Future MR studies should provide a more detailed rationale for the selection of specific MR methods, particularly when using novel techniques. A comparative analysis of the performance of different methods in the context of the study's data could offer valuable insights into their relative strengths and weaknesses. Additionally, researchers should consider the use of complementary MR methods, such as the weighted median [45] or mode-based estimators [46], to provide a more robust assessment of causal relationships. By thoroughly discussing the rationale for method selection and the potential limitations of each approach, researchers can enhance the credibility and impact of their findings.

Conclusion

The study by Li et al. represents an important contribution to the field of genetic epidemiology by exploring the genetic susceptibility associations between viral infections and colorectal cancer risk using two-sample MR [3]. While the study provides valuable insights, several areas for improvement have been identified. These include the need for more rigorous IV selection, better accounting for pleiotropy and confounding, a more in-depth exploration of biological plausibility, and improved reporting and transparency.

By addressing these issues, future research can build on the findings of Li et al. and provide more robust evidence on the role of viral infections in colorectal cancer etiology. Additionally, the application of novel MR techniques and the use of complementary methods can enhance the reliability and validity of MR analyses, contributing to a deeper understanding of the complex relationships between genetics, infections, and cancer risk.

Acknowledgements
Not applicable.

Author contributions
T.Y. wrote the original draft of the manuscript. J.X., H.Y., N.Y. and L.Z. undertook the analysis. M.L. contributed to the conception and design of the study. All authors read and approved the final manuscript.

Funding
This work was supported by the Fundamental Research Funds in Heilongjiang Provincial Universities (145109136) and the Heilongjiang Province Leading Talent Echelon Reserve Leader Funding Project. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Data availability
No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 5 September 2024 / Accepted: 2 December 2024
Published online: 05 December 2024

References

1. Lu Z, Chai X, Pan Y, Li S. The causality between $CD8^{+}$NKT cells and $CD16^{-}$ CD56 on NK cells with hepatocellular carcinoma: a mendelian randomization study. Infect Agent Cancer. 2024;19(1):3. https://doi.org/10.1186/s13027-024-00565-8.
2. Gkatzionis A, Burgess S, Newcombe PJ. Statistical methods for cis-mendelian randomization with two-sample summary-level data. Genet Epidemiol. 2023;47(1):3–25. https://doi.org/10.1002/gepi.22506.
3. Li G, Wang S, Ma J, Liu S. Genetic susceptibility association between viral infection and colorectal cancer risk: a two-sample mendelian randomization analysis. Infect Agent Cancer. 2024;19(1):37. https://doi.org/10.1186/s13027-024-00602-6.
4. Dong Q, Chen D, Zhang Y, Xu Y, Yan L, Jiang J. Constipation and cardiovascular disease: a two-sample mendelian randomization analysis. Front Cardiovasc Med. 2023;10:1080982. https://doi.org/10.3389/fcvm.2023.1080982.
5. Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. J Anim Sci Biotechnol. 2017;8:38. https://doi.org/10.1186/s40104-017-0164-6.
6. Malik T, Lee MJ, Harikrishnan AB. The incidence of stoma related morbidity - a systematic review of randomised controlled trials. Ann R Coll Surg Engl. 2018;100(7):501–8. https://doi.org/10.1308/rcsann.2018.0126.
7. Wang J, Huang P, Yu Q, Lu J, Liu P, Yang Y, et al. Epilepsy and long-term risk of arrhythmias. Eur Heart J. 2023;44(35):3374–82. https://doi.org/10.1093/eurheartj/ehad523.
8. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. Nat Genet. 2018;50(5):693–8. https://doi.org/10.1038/s41588-018-0099-7.
9. Burgess S, Thompson SG. Interpreting findings from mendelian randomization using the MR-Egger method. Eur J Epidemiol. 2017;32(5):377–89. https://doi.org/10.1007/s10654-017-0255-x.
10. Ong JS, MacGregor S. Implementing MR-PRESSO and GCTA-GSMR for pleiotropy assessment in mendelian randomization studies from a practitioner's perspective. Genet Epidemiol. 2019;43(6):609–16. https://doi.org/10.1002/gepi.22207.
11. Yang M, Wan X, Zheng H, Xu K, Xie J, Yu H, et al. No evidence of a genetic causal relationship between Ankylosing spondylitis and Gut Microbiota: a two-sample mendelian randomization study. Nutrients. 2023;15(4):1057. https://doi.org/10.3390/nu15041057.
12. Zeng Y, Cao S, Yang H. The causal role of gastroesophageal reflux disease in anxiety disorders and depression: a bidirectional mendelian randomization study. Front Psychiatry. 2023;14:1135923. https://doi.org/10.3389/fpsyt.2023.1135923.
13. Cinelli C, LaPierre N, Hill BL, Sankararaman S, Eskin E. Robust mendelian randomization in the presence of residual population stratification, batch effects and horizontal pleiotropy. Nat Commun. 2022;13(1):1093. https://doi.org/10.1038/s41467-022-28553-9.
14. Cui G, Li S, Ye H, Yang Y, Jia X, Lin M, et al. Gut microbiome and frailty: insight from genetic correlation and mendelian randomization. Gut Microbes. 2023;15(2):2282795. https://doi.org/10.1080/19490976.2023.2282795.

15. Rasooly D, Peloso GM. R Curr Protoc. 2021;1(12):e335. https://doi.org/10.1002/cpz1.335. Two-Sample Multivariable Mendelian Randomization Analysis Using.
16. Si S, Li J, Tewara MA, Xue F. Genetically determined chronic low-Grade inflammation and hundreds of Health outcomes in the UK Biobank and the FinnGen Population: a phenome-wide mendelian randomization study. Front Immunol. 2021;12:720876. https://doi.org/10.3389/fimmu.2021.720876.
17. LaPierre N, Fu B, Turnbull S, Eskin E, Sankararaman S. Leveraging family data to design mendelian randomization that is provably robust to population stratification. Genome Res. 2023;33(7):1032–41. https://doi.org/10.1101/gr.277664.123.
18. Liu W, Deng Y, Li Z, Chen Y, Zhu X, Tan X, Cao G. Cancer Evo-Dev: a theory of inflammation-Induced Oncogenesis. Front Immunol. 2021;12:768098. https://doi.org/10.3389/fimmu.2021.768098.
19. Jahan N, Ghouse SM, Martuza RL, Rabkin SD. In situ Cancer vaccination and immunovirotherapy using Oncolytic HSV. Viruses. 2021;13(9):1740. https://doi.org/10.3390/v13091740.
20. Epstein AL, Rabkin SD. Safety of non-replicative and oncolytic replication-selective HSV vectors. Trends Mol Med. 2024;30(8):781–94. https://doi.org/10.1016/j.molmed.2024.05.014.
21. Dor E, Fluss R, Israel A, Huppert A. Quantifying the long-term effects of measles infection-a retrospective cohort study. Clin Microbiol Infect. 2024;30(11):1460–5. https://doi.org/10.1016/j.cmi.2024.08.007.
22. Delpeut S, Sisson G, Black KM, Richardson CD. Measles Virus enters breast and Colon cancer cell lines through a PVRL4-Mediated macropinocytosis pathway. J Virol. 2017;91(10):e02191–16. https://doi.org/10.1128/JVI.02191-16.
23. Mahale P, Yanik EL, Engels EA. Herpes Zoster and Risk of Cancer in the Elderly U.S. Population. Cancer Epidemiol Biomarkers Prev. 2016;25(1):28–35. https://doi.org/10.1158/1055-9965.EPI-15-1033. Epub.
24. Tada S, Kaito Y, Watanabe A, Sugiyama Y, Nishigaichi A, Miwa T, et al. Varicella-zoster meningitis and myelitis after herpes zoster Dermatitis Treatment with Amenamevir: a Case Series and Literature Review. Cureus. 2024;16(2):e54775. https://doi.org/10.7759/cureus.54775.
25. Chetty R, Capo-Chichi JM, Serra S. Colorectal large-cell neuroendocrine carcinoma with lymphoid stroma: further evidence confirming a unique subtype associated with MLH1/PMS2 loss, BRAF mutation, Epstein-Barr virus negativity, and the possibility of a better prognosis. Histopathology. 2019;75(2):247–53. https://doi.org/10.1111/his.13875.
26. Leopizzi M, Mundo L, Messina E, Campolo F, Lazzi S, Angeloni A, et al. Epstein-Barr virus-encoded EBNA2 downregulates ICOSL by inducing miR-24 in B-cell lymphoma. Blood. 2024;143(5):429–43. https://doi.org/10.1182/blood.2023021346.
27. Seo D, Byun H, Cho M, Lee SH, Youn S, Lee J, et al. Dihydrotestosterone enhances MICA-Mediated Immune responses to Epstein-Barr Virus-Associated gastric carcinoma. Cancers (Basel). 2024;16(18):3219. https://doi.org/10.3390/cancers16183219.
28. Ma N, He F, Kawanokuchi J, Wang G, Yamashita T. Taurine and its anticancer functions: in vivo and in Vitro Study. Adv Exp Med Biol. 2022;1370:121–8. https://doi.org/10.1007/978-3-030-93337-1_11.
29. Wang M, Li L, Li J, Li Y, Wang Z, Guo Y, et al. Causal effect of breast cancer on endometrial cancer risk: a two-sample mendelian randomization study. Eur J Obstet Gynecol Reprod Biol. 2024;304:16–22. https://doi.org/10.1016/j.ejogrb.2024.11.013.
30. Yuji K, Tanimoto T, Oshima Y. 23andMe and the FDA. N Engl J Med. 2014;370(23):2248. https://doi.org/10.1056/NEJMc1404692.
31. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613(7944):508–18. https://doi.org/10.1038/s41586-022-05473-8.
32. Hoaglin DC. Misunderstandings about Q and 'Cochran's Q test' in meta-analysis. Stat Med. 2016;35(4):485–95. https://doi.org/10.1002/sim.6632.
33. Aslam M. Cochran's Q test for analyzing categorical data under uncertainty. J Big Data. 2023;10(1):147. https://doi.org/10.1186/s40537-023-00823-3.
34. Bowden J, Holmes MV. Meta-analysis and mendelian randomization: a review. Res Synth Methods. 2019;10(4):486–96. https://doi.org/10.1002/jrsm.1346.
35. Sun JX, Xu JZ, Liu CQ, An Y, Xu MY, Zhong XY, et al. The association between human papillomavirus and bladder cancer: evidence from meta-analysis and two-sample mendelian randomization. J Med Virol. 2023;95(1):e28208. https://doi.org/10.1002/jmv.28208.
36. Lu K, Guo ZC, Zhang JJ, Yu X, Zhang ZY. Functional investigation and two-sample mendelian randomization study of Inguinal Hernia hub genes obtained by Bioinformatics Analysis. Curr Comput Aided Drug Des. 2024. https://doi.org/10.2174/0115734099282407240325054745. Online ahead of print.
37. Abbasi J. JAMA. 2020;323(10):916. https://doi.org/10.1001/jama.2020.2238. 23andMe Develops First Drug Compound Using Consumer Data.
38. 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: the STROBE-MR Statement. JAMA. 2021;326(16):1614–21. https://doi.org/10.1001/jama.2021.18236.
39. Lin Z, Pan W. A robust cis-mendelian randomization method with application to drug target discovery. Nat Commun. 2024;15(1):6072. https://doi.org/10.1038/s41467-024-50385-y.
40. Xiang N, Su S, Wang Z, Yang Y, Chen B, Shi R, et al. Exploring the causal relationship: bidirectional mendelian randomization study on benign prostatic hyperplasia and cardiovascular diseases. Front Genet. 2024;15:1432055. https://doi.org/10.3389/fgene.2024.1432055.
41. Karageorgiou V, Tyrrell J, Mckinley TJ, Bowden J. Weak and pleiotropy robust sex-stratified mendelian randomization in the one sample and two sample settings. Genet Epidemiol. 2023;47(2):135–51. https://doi.org/10.1002/gepi.22512.
42. Wang H, Chen X, Wang S, Zhang H. Exploration of the causal effects of leukocyte elomere length and four gastrointestinal diseases: a two-sample bidirectional mendelian randomization study. BMC Gastroenterol. 2023;23(1):446. https://doi.org/10.1186/s12876-023-03081-y.
43. Lin Z, Xue H, Pan W. Robust multivariable mendelian randomization based on constrained maximum likelihood. Am J Hum Genet. 2023;110(4):592–605. https://doi.org/10.1016/j.ajhg.2023.02.014.
44. Luo W, Yv H, Yu X, Wu X. Investigating the Causal Link between Rheumatoid Arthritis and Atrial Fibrillation in east Asian populations: a mendelian randomization Approach. Cardiol Res Pract. 2024;2024:3274074. https://doi.org/10.1155/2024/3274074.
45. Xu J, Zhang S, Tian Y, Si H, Zeng Y, Wu Y, et al. Genetic Causal Association between Iron Status and Osteoarthritis: a two-sample mendelian randomization. Nutrients. 2022;14(18):3683. https://doi.org/10.3390/nu14183683.
46. Wu F, Huang Y, Hu J, Shao Z. Mendelian randomization study of inflammatory bowel disease and bone mineral density. BMC Med. 2020;18(1):312. https://doi.org/10.1186/s12916-020-01778-5.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.