

Life behaviors, educational attainment, and risk for endometriosis: a Mendelian randomization study

Shiyu Kong¹, Caining Zhuang¹, He Chen², Wanlei Sun², Wenfeng He^{1*}

¹Department of Medical Genetics, the Second Affiliated Hospital of Nanchang University, China

²Department of Clinical Laboratory, the Second Affiliated Hospital of Nanchang University, China

Submitted: 20 February 2025; **Accepted:** 19 July 2025

Online publication: 31 August 2025

Arch Med Sci

DOI: <https://doi.org/10.5114/aoms/208430>

Copyright © 2025 Termedia & Banach

***Corresponding author:**

Wenfeng He
Department of
Medical Genetics
the Second Affiliated
Hospital of
Nanchang University
China
E-mail:
fengfeng_982366@126.com

Endometriosis is characterized by endometrial-like mucosa outside the uterine cavity, affecting 5–10% of women of reproductive age worldwide [1, 2]. Lesion development involves genetic, hormonal, and immunological factors as well as endometrial progenitor cells [3]. Despite its prevalence, the disease remains poorly understood, and available treatments often fail to meet patient expectations [2, 3].

In literature, life behaviors (LBs) have been related to human health [4–6], and the association between LBs and endometriosis has been explored in observational epidemiological research. According to large-scale meta-analyses and observational studies, smoking was associated with a lower prevalence of endometriosis among women [7]. However, a related meta-analysis found no significant association between tobacco use and endometriosis risk [8], leaving the causal association between the two uncertain.

Recent studies suggest a potential association between sleep disturbances and endometriosis [9]. Ramin-Wright (2018) observed a negative association between endometriosis and sleep quality [10]. Similarly, retrospective and prospective studies have not demonstrated a significant association between different types of physical activity and endometriosis [11, 12]. Moreover, many of these behaviors are often considered as consequences of the disease, rather than contributors to its onset. To overcome these limitations, it is necessary to apply a method that minimizes the influence of reverse causality and confounding. Here, we employed a two-sample Mendelian randomization (MR) approach to examine the underlying associations between LBs and endometriosis.

Methods. Study design. Data were curated on 14 LB-related factors using genome-wide association studies (GWAS) summary data, available in Supplementary Table S1. A two-sample MR analysis was performed to explore the causality of endometriosis (Figure 1). Statistical analysis was done using TwoSampleMR R package, version 0.6.4, and MR-PRESSO R package, version 1.0.

GWAS data for LB-related factors. The study divided 14 phenotypes into educational attainment, physical activity, smoking behavior, and sleep behavior [1–3]. Supplementary Table S1 provides the detailed definition and description of each phenotype.

GWAS data for endometriosis. Statistics of endometriosis patients were obtained from FinnGen (FinnGen, <https://www.finnngen.fi/>). In Fin-

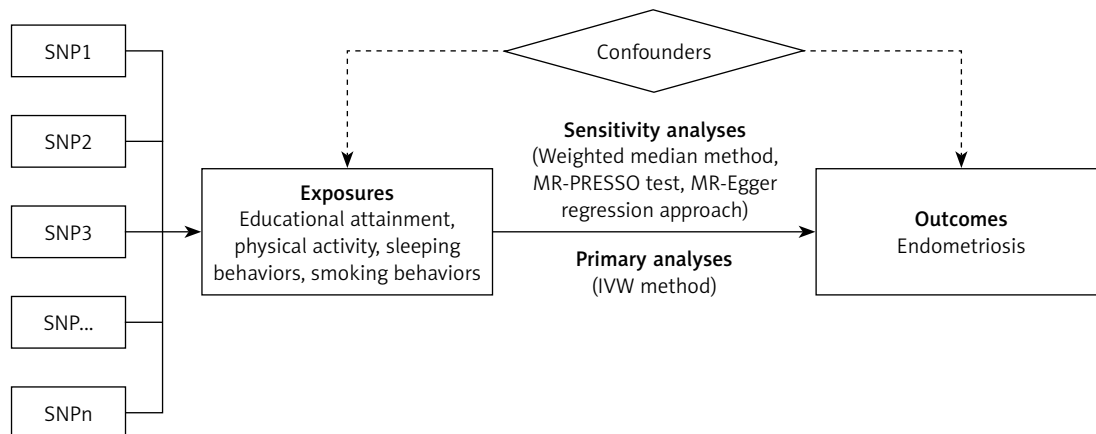


Figure 1. Flowchart depicting the current MR design

SNP – single nucleotide polymorphism, MR-PRESSO – Mendelian randomization-pleiotropy residual sum and outlier.

nGen, endometriosis is defined by N80 in ICD-10, 617 in ICD-9, and 6253 in ICD-8. FinnGen is a research project in genomics and personalized medicine. It is a large public-private partnership, which has collected and analyzed genome and health data from 412,181 Finnish Biobank donors to better comprehend the genetic bases of diseases. This data set includes 128,171 endometriosis individuals (16,588 cases and 111,583 controls; European blood). Supplementary Table SII displays the detailed description of the GWAS endometriosis data.

Selection of genetic instruments for LB-related phenotypes. Independent genetic variants strongly associated with LB ($p < 5 \times 10^{-8}$) were identified by clumping single nucleotide polymorphisms (SNPs) with a linkage disequilibrium threshold of $r^2 < 0.001$ within a 10,000-kb window, using the 1,000 Genomes European reference panel (Supplementary Table SIII). To determine whether there was a weak instrumental variable bias, F-statistics were calculated to quantify the strength of instrumental variables, wherein F-statistics > 10 indicated a low possibility of weak instrumental variable bias [5, 6].

MR estimates. Primary MR analysis employed the random-effects IVW method that combines causal estimates from individual SNPs through meta-analysis, to determine the pooled causal effect of the exposure on the outcome. This approach offers a more conservative causal inference by considering uncertainty resulting from pleiotropy [7]. Throughout the study, 'IVW' specifically referred to the random-effects IVW method. Statistical significance was determined using the Bonferroni-corrected threshold of $p < 4 \times 10^{-3}$ (0.05 divided by 14 exposures multiplied by one outcome), where $p < 0.05$ indicated association and $p > \text{Bonferroni designated threshold}$.

Quality control of MR estimation. Quality control was conducted on significant findings ($p < 0.05$). Heterogeneity in the IVW model was first

assessed using the Cochran's Q test and quantified with I^2 statistics [8]. Cochran's Q test results with $p < 0.05$ and $I^2 > 25\%$ suggested potential heterogeneity. MR-Egger regression was employed to investigate the potential pleiotropic effects of instrumental variables, with the intercept term serving as an indicator of directional horizontal pleiotropy in causal estimates. Furthermore, a leave-one-out analysis was conducted by systematically excluding each SNP and performing MR analysis on the remaining SNPs, so as to identify any potential outliers among instrumental variables [8]. The Steiger test of directionality was applied to evaluate the causal relationship between the exposure and outcome.

Sensitivity analysis. Three databases were searched, i.e., Ensembl, GWAS Catalog, and PhenoScanner to explore any potential associations of the selected SNPs with other known risk factors (confounders) of endometriosis identified in previous MR studies.

A method described by Brion *et al.* was used to calculate the statistical power [9] (<https://shiny.cnsgenomics.com/mRnd/>). A sufficient power of $> 80\%$ was recommended. Summary-level statistics only were employed, which waived off the need for ethical approval.

Results. Study overview. The study evaluated the causal effects of 14 genetically predicted LB-related phenotypes on endometriosis. The detailed assumptions for each MR method are provided in Supplementary Table SV. After rigorous single nucleotide polymorphism (SNP) filtering, the number of SNPs used per phenotype ranged from 2 to 95 (Supplementary Table SVI). The F-statistics ranged from 15.48 to 965.05, suggesting that the likelihood of bias due to weak instruments was minimal (Supplementary Table SVII). In the primary analysis, one causal association reached the significance threshold of $p < 4 \times 10^{-3}$ (Figure 2). The full effect estimates from various

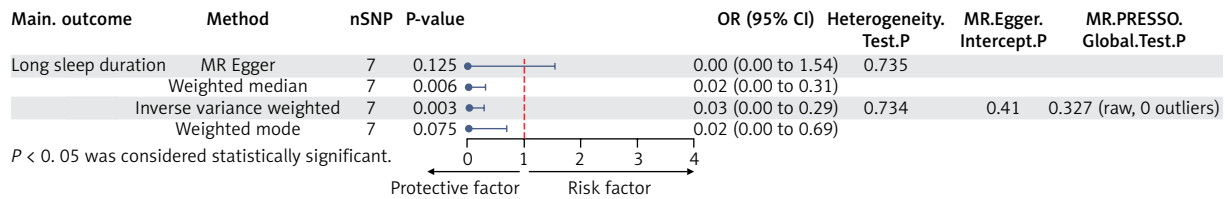


Figure 2. Associations between long sleep duration and the risk of endometriosis

CI – confidence interval, OR – odds ratio, SNP – single nucleotide polymorphism.

MR models are available in Supplementary Table SVI, while the statistical power calculations are presented in Supplementary Table SVII.

Causal effects of LB-related phenotypes on endometriosis. Following the Bonferroni correction, long sleep duration was the only phenotype showing a causal association with endometriosis in the IVW model (Supplementary Figure S1, Supplementary Table SVI). Specifically, long sleep duration (odds ratio: 0.03; 95% confidence interval: 2.77e-03–0.29; $p = 0.003$) was associated with a reduced risk of endometriosis. Smoking cessation showed a protective trend (odds ratio: 0.58; 95% confidence interval: 0.37–0.91; $p = 0.038$), but it did not meet the corrected significance threshold ($p < 4 \times 10^{-3}$). Therefore, this association was considered suggestive, indicating preliminary evidence, which warrants further validation.

No evidence of heterogeneity was found for either long sleep duration (heterogeneity test $Q_{pval} = 0.734$) or smoking cessation (heterogeneity test $Q_{pval} = 0.311$). The causal estimates were consistent across MR models, and the MR–Egger intercept indicated no directional pleiotropy (Supplementary Tables SVI, SVIII). The leave-one-out analysis confirmed that the observed effects were not driven by any single SNP (Supplementary Figure S2). Directionality testing using the Steiger method further supported the negative causal effect of long sleep duration on endometriosis risk (Supplementary Table SVI).

Genetic variants associated with potential confounders. Potential confounding SNPs were identified through systematic screening of three databases: Ensembl, PhenoScanner, and the GWAS Catalog (Supplementary Table SIX). The identified variants were then consolidated and excluded to evaluate the overall impact of confounders on the observed associations. Sensitivity analyses indicated that the observed robust associations were not significantly biased by these potential confounders (Supplementary Table SX).

Discussion. The findings of the study underscored the significant role of long sleep duration as a protective factor and suggested a potentially beneficial effect of smoking cessation on the risk of endometriosis.

While current smoking showed no causal effect ($p > 0.05$), we observed a suggestive protective as-

sociation for smoking cessation due to the following reasons: 1. It did not remain significant after the Bonferroni correction; 2. It was supported by biological plausibility, such as re-establishment of estrogen balance [13, 14]; and 3. Previous observational studies have yielded inconsistent results [15, 16]. These findings may indicate that cessation-specific biological mechanisms, rather than smoking itself, influence the risk of disease. Further mechanistic studies are warranted to clarify how smoking and its cessation influence endometriosis risk at molecular level.

Although our study did not establish a direct causal association between insomnia symptoms and endometriosis risk, the findings suggest that longer sleep duration may exert a protective effect. This observation appears somewhat contradictory to prior findings. For example, a study by Cakan *et al.* reported that half- and whole-night work shifts disrupted both sleep duration and quality, altered melatonin secretion, and suppressed estradiol levels [17]. These findings support a hypothesis that sleep disruption may negatively affect estrogen regulation and impair ovarian function. In contrast, prolonged and sufficient sleep may stabilize hormone secretion and promote ovarian health. The lack of a causal association for insomnia ($p > 0.05$) supports the interpretation that sleep disturbances in endometriosis are more likely secondary to pain [18, 19], rather than serving as a causative factor. Conversely, the potential protective role of long sleep duration may stem from its capacity to regulate endocrine function [17]. Despite these insights, the biological pathways associating sleep patterns and endometriosis remain inadequately understood, reinforcing the need for further mechanistic research [20].

Nevertheless, some limitations must be acknowledged. Our study relied on self-reported phenotypes, which may be subject to reporting bias. Also, our reliance on summary-level data precluded age- and sex-stratified analyses. Additionally, the lack of complete GWAS data for certain exposures prevented us from conducting reverse-direction MR analyses. Future studies should aim to perform bidirectional MR once full summary statistics are available.

In conclusion, this MR analysis provides strong evidence that long sleep duration may reduce the

risk of endometriosis. The suggestive protective effect of smoking cessation deserves further exploration. These insights underscore the potential of prioritizing sleep hygiene as a modifiable factor in preventive strategies for endometriosis.

Acknowledgments

We acknowledge the investigators and participants of the FinnGen and original GWAS. We are thankful to all GWAS sharing summary data used in this study.

Data availability statement

All data used in our study are publicly available. For LB-related phenotypes in the study, corresponding summary statistics were available from each of the original GWAS research. Summary statistics for endometriosis were available on the FinnGen website: <https://www.finnngen.fi/>. Further inquiries can be directed to the corresponding author.

Funding

Natural Science Foundation of Jiangxi Province [20232BAB206055].

Ethical statement

Our study was a secondary analysis of publicly available data. The summary statistics for the MR study were acquired from GWAS: <https://www.ebi.ac.uk>, <https://www.finnngen.fi/>, and some other original GWAS studies. These data are freely accessible, and all ethical approvals for the GWAS were obtained by the original GWAS authors or organization.

Conflict of interest

The authors declare no conflict of interest.

References

- Vercellini P, Viganò P, Somigliana E, Fedele L. Endometriosis: pathogenesis and treatment. *Nat Rev Endocrinol* 2014; 10: 261-75.
- Taylor HS, Kotlyar AM, Flores VA. Endometriosis is a chronic systemic disease: clinical challenges and novel innovations. *Lancet* 2021; 397: 839-52.
- Horne AW, Saunders PTK. SnapShot: endometriosis. *Cell* 2019; 179: 1677.
- Huang G, Cai J, Li W, Zhong Y, Liao W, Wu P. Causal relationship between educational attainment and the risk of rheumatoid arthritis: a Mendelian randomization study. *BMC Rheumatol* 2021; 5: 47.
- Pickett KE, Wilkinson RG. Income inequality and health: a causal review. *Soc Sci Med* 2015; 128: 316-26.
- Carter AR, Gill D, Davies NM, et al. Understanding the consequences of education inequality on cardiovascular disease: mendelian randomisation study. *BMJ* 2019; 365: l1855.
- Baron JA. Beneficial effects of nicotine and cigarette smoking: the real, the possible and the spurious. *Br Med Bull* 1996; 52: 58-73.
- Bravi F, Parazzini F, Cipriani S, et al. Tobacco smoking and risk of endometriosis: a systematic review and meta-analysis. *BMJ Open* 2014; 4: e006325.
- Leone Roberti Maggiore U, Bizzarri N, Scala C, et al. Symptomatic endometriosis of the posterior cul-de-sac is associated with impaired sleep quality, excessive daytime sleepiness and insomnia: a case-control study. *Eur J Obstet Gynecol Reprod Biol* 2017; 209: 39-43.
- Ramin-Wright A, Schwartz ASK, Geraedts K, et al. Fatigue – a symptom in endometriosis. *Hum Reprod* 2018; 33: 1459-65.
- Bonocher CM, Montenegro ML, Rosa E, Silva JC, Ferriani RA, Meola J. Endometriosis and physical exercises: a systematic review. *Reprod Biol Endocrinol* 2014; 12: 4.
- Tennfjord MK, Gabrielsen R, Tellum T. Effect of physical activity and exercise on endometriosis-associated symptoms: a systematic review. *BMC Womens Health* 2021; 21: 355.
- Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinforma Oxf Engl* 2015; 31: 3555-7.
- Li B, Martin EB. An approximation to the F distribution using the chi-square distribution. *Comput Stat Data Anal* 2002; 40: 21-6.
- Ricci E, Viganò P, Cipriani S, et al. Physical activity and endometriosis risk in women with infertility or pain: systematic review and meta-analysis. *Medicine (Baltimore)* 2016; 95: e4957.
- Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet* 2019; 51: 237-44.
- Kim HJ, Lee HS, Kazmi SZ, et al. Familial risk for endometriosis and its interaction with smoking, age at menarche and body mass index: a population-based cohort study among siblings. *BJOG Int J Obstet Gynaecol* 2021; 128: 1938-48.
- Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. *Eur J Epidemiol* 2017; 32: 377-89.
- Koninckx PR, Ussia A, Adamyan L, Wattiez A, Gomel V, Martin DC. Pathogenesis of endometriosis: the genetic/epigenetic theory. *Fertil Steril* 2019; 111: 327-40.
- Chmaj-Wierzchowska K, Rzymiski P, Wojciechowska M, Parda I, Wilczak M. Health-related quality of life (Nottingham Health Profile) in patients with endometriomas: correlation with clinical variables and self-reported limitations. *Arch Med Sci* 2019; 16: 584-91.