



OPEN

A Mendelian randomization-based approach to explore the relationship between leukocyte counts and breast cancer risk in European ethnic groups

Zhitao Zhang^{1,3}, Lei Li²✉ & Jianbin Wu^{1,3}✉

Exploring the potential association between peripheral blood leukocyte counts and breast cancer risk by Mendelian randomization (MR) analysis methods. Genetic data related to peripheral blood sorting counts of leukocytes were collected from a genome-wide association study by Blood Cell Consortium (BCX). Single nucleotide polymorphic loci predicting peripheral blood sorting counts of these leukocytes were selected as instrumental variables according to the correlation assumption, independence assumption and exclusivity assumption of MR. The data on breast cancer and its subtypes were obtained from Breast Cancer Association Consortium (BCAC) and FinnGen Consortium. In this study, the Inverse-Variance Weighted (IVW), Weighted Median, MR-Egger, Maximum Likelihood (ML), MR-PRESSO and Constrained Maximum Likelihood and Model Averaging (cML-MA) methods of random effects models were used for MR analysis. Cochran's Q analysis, and MR-Egger intercept analysis were applied for sensitivity analysis. IVW and cML-MA were considered the primary analytical tools, and the results of the other 4 MRs were used as complementary and validation. The results suggest that there is no significant causal relationship between leukocyte count and breast cancer risk (IVW OR = 0.98 [95% CI: 0.93–1.03], p-value = 0.35; cML-MA OR = 1.01 [95% CI: 0.98–1.05], p-value = 0.51). In addition, we analyzed whether there was a potential correlation between the five main types of categorized leukocyte counts and different breast cancer subtypes. We did not find significant evidence to support a significant correlation between leukocyte counts and breast cancer subtypes.

Breast cancer (BC) is a common female cancer, and due to its high mortality and morbidity, BC has become a significant burden on women's health^{1–3}. Notably, in 2018, the United States reported the diagnosis of approximately 268,670 new BC cases^{4,5}. Moreover, BC exhibits pronounced heterogeneity, showcasing varying clinical presentations among individual patients, thus posing intricate challenges in its therapeutic approach. Drawing upon both molecular and histological evidence, BC may be systematically categorized into three primary subtypes: hormone receptor-positive BC (estrogen receptor (ER+) or progesterone receptor (PR+)), human epidermal receptor 2-positive BC (HER2+), and triple-negative breast cancer (TNBC)^{6–8}.

The significance of the immune microenvironment in the initiation and advancement of BC has garnered growing recognition^{9,10}. Tumor-associated immune cells have exhibited prospective prognostic utility across diverse malignancies, with elevated quantities of tumor-infiltrating lymphocytes frequently linked to more favorable clinical outcomes in BC¹¹. However, research in the realm of the association between circulating immune cell subtypes and the clinical attributes or risk profiles of BC patients remains notably underexplored.

A patient's leukocyte levels may provide clues to BC risk. A study by Kresovich et al.¹² suggests that the percentage of certain leukocyte types in a woman's blood may predict her risk of being diagnosed with breast cancer in the short and long term. The immune system plays an important role in the occurrence and development of BC, and it protects the health of the human body. Gene mutations are not uncommon in normal human

¹Department of Breast, Fujian Maternity and Child Health Hospital, Fuzhou, China. ²Department of Pathology, University of Otago, Dunedin 9016, New Zealand. ³These authors contributed equally: Zhitao Zhang and Jianbin Wu. ✉email: wshililei@buaa.edu.cn; wjbfmu@163.com

beings, but not every harmful mutation will develop into a malignant tumor. This is because the immune system eliminates these mutant cells in real time and resolves the threat of cancer at the initial stage. Only when the immune system malfunctions or the mutated cells establish an immune escape mechanism can the mutated cells, which were originally very small, have the opportunity to continue to divide and proliferate, and eventually form tumors^{13,14}. Therefore, detecting the immune function status of the body may be beneficial to assess the risk of malignant tumors such as BC^{15,16}. White blood cells are important immune cells in the body, including neutrophils, eosinophils, basophils, lymphocytes and monocytes^{17,18}. Different types of white blood cells participate in the body's immune defense response in different ways. A previous retrospective study has provided preliminary evidence of a conceivable correlation between peripheral leukocyte and neutrophil counts and the clinical attributes of BC patients¹⁹. Larsson et al.²⁰ reported significant associations between peripheral blood mononuclear cell populations and prognosis in patients afflicted with metastatic BC. Furthermore, there is emerging evidence suggesting potential links between peripheral blood lymphocyte levels and clinical characteristics, as well as chemotherapy responsiveness, in BC patients²¹. The level of peripheral blood leukocyte count may be one of the most important indicators for evaluating the risk characteristics of BC patients, while the leukocyte categorical count may better reflect the dynamic characteristics of the microenvironment of BC patients.

The majority of the research evidence regarding the association between peripheral blood leukocyte counts and BC risk is derived from observational studies. Nevertheless, these studies are susceptible to various limitations, including confounding factors, selection bias, and other biases that may compromise the validity of their findings. Mendelian randomization (MR) analysis employs genetic variants as instrumental variables (IVs) to enhance the causal inference in exposure-outcome relationships, particularly when pleiotropic effects are absent^{22,23}. This approach mitigates the confounding issues inherent in observational epidemiology, ultimately yielding more dependable and rigorous results. MR is now widely used in the study of BC^{24–26}. In this study, we performed MR analysis of two samples to investigate the potential association between peripheral blood leukocyte counts and breast cancer risk. In addition, we further investigated the potential relationship between categorized leukocyte cell counts (Neutrophils, eosinophils, basophils, lymphocytes and monocytes) and breast cancer subtypes (ER+ Breast cancer, ER- Breast cancer, HER- Breast cancer, HER+ Breast cancer, and HER2- Breast cancer). The flowchart of this study is shown in Fig. 1.

Methods

Study populations

“Exposure” in this study was defined as a categorical count of multiple leukocytes in peripheral blood, and the data were obtained from a GWAS study conducted by the Blood Cell Consortium (BCX)^{27,28}. The study included 563,946 European ethnic subjects and provided GWAS data on five types of peripheral blood leukocyte counts: neutrophils, eosinophils, basophils, lymphocytes and monocytes (Supplementary Tables 1). In this study, “outcome” was selected as breast cancer and its subtypes, and data for Breast cancer, ER+ Breast cancer, and ER- Breast cancer were obtained from a published GWAS meta-analysis study conducted by the Breast Cancer Association Consortium (BCAC)²⁹. Data for HER2- breast cancer, HER- breast cancer, and HER+ breast cancer were obtained from the FinnGen Consortium. The FinnGen study represents a research initiative that combines genetic data obtained from Finnish biobanks with health records derived from Finnish health registries³⁰. All breast cancer cases had a diagnosis established according to standard procedures and were supported by a

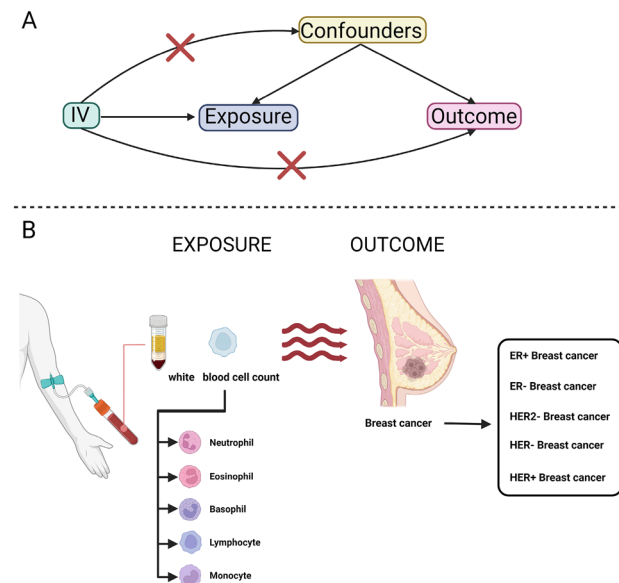


Figure 1. Research flow chart. (A) The selection of instrumental variables needed to fulfill 3 assumptions. (B) The exposure in the study was the white blood cells and their subtypes. The outcome in the study is breast cancer and its subtypes.

pathological basis. The exposure and outcome data were downloaded from the “GWAS summary statistics” database at <https://gwas.mrcieu.ac.uk/>^{31,32}. Given that the present study utilized publicly available summary data, there was no necessity for additional ethical approval or participant consent. The exposure and outcome specific information is shown in Table 1.

Selection of instrumental variables

Based on the relevance, independence and exclusivity assumptions of MR, single nucleotide polymorphism loci (SNPs, Single nucleotide polymorphisms) that met the requirements were screened from the above exposure GWAS data and used as instrumental variables (IV) in this study (Fig. 1A)^{33,34}. Specifically, all SNPs had to meet the following criteria: (1) Exposure correlation P-value less than 5×10^{-8} , along with an F-statistic greater than 10, thus satisfying the correlation assumption of MR. (2) SNPs with significant linkage disequilibrium with the measured SNPs ($r^2 = 0.001$) were removed in the range of 10,000 kb, thus satisfying the independence assumption of MR. (3) SNPs with significant associations with outcomes or confounders were manually removed via the PhenoScanner website, thus satisfying the exclusivity hypothesis. These eligible instrumental variables were then extracted from the outcome GWAS with the necessary coordination so that the effect of a single instrumental variable on exposure corresponded to the same allele as the effect on outcome. The data of the IVs are shown in Supplementary data S1–S6.

Statistical analysis

In this study, the Inverse-Variance Weighted (IVW), Weighted Median, MR-Egger, Maximum Likelihood (ML), MR-PRESSO and Constrained Maximum Likelihood and Model Averaging (cML-MA) methods of random effects models were used for MR analysis^{35–38}. IVW is one of the most commonly used MR analysis methods^{39,40}. The IVW method has the strongest causal inference and provides the most reliable findings when the instrumental variables are not confounded by horizontal pleiotropy. The weighted median and MR-Egger methods, on the other hand, serve as complements to the IVW method^{41,42}. Because these two methods can be used under a wider range of conditions and provide more conservative results. Of these, the weighted median method allows no more than 50% of the instrumental variables to be confounded by horizontal pleiotropy, whereas the MR-Egger method allows all instrumental variables to be pleiotropic, but this pleiotropy cannot interfere with the correlation of instrumental variables with exposure⁴³. In contrast to the IVW method, the ML method offers the advantage of a reduced standard error, and its findings remain unbiased in the absence of heterogeneity or horizontal polymorphism. We also used the MR-PRESSO method to identify and remove any outlier variants⁴⁴. This method regresses the genetic variance results on the genetic variance exposure and uses residual squares to identify outliers. The cML-MA method, a MR technique that incorporates ML and model averaging, is applied to address both correlated and uncorrelated pleiotropic effects. Importantly, it does not rely on the InSIDE (Instrument Strength Independent of Direct Effect) assumption, setting it apart from other MR approaches. cML-MA has better type I error control. Xue et al.⁴⁵ demonstrated that, in certain scenarios, the cML-MA method might offer advantages over the IVW method. As such, in this research, the primary analytical tools employed were the cML-MA and IVW methods, which were complemented by results obtained from four additional methods. MR-Egger and MR-PRESSO methods were employed to identify horizontal pleiotropy. In the MR-PRESSO analysis, the number of distributions was specified as 1000. Furthermore, the detection of heterogeneity was carried out using the IVW method and MR-Egger regression, with the quantification of heterogeneity achieved through the Cochran Q statistic. A $P < 0.05$ was considered statistically significant, while Bonferroni-corrected adjusted p-values ($BP < 0.05/N$, $N =$ number of comparisons) in analyses involving multiple comparisons were considered statistically significant in order to minimize the potential risk of type I error and to improve the overall validity and interpretability of the results. The selection of all instrumental variables and the MR analysis process were done by applying the “TwoSampleMR” and “MRcML” R package (version 4.3.1)^{45–48}.

Type	GWAS ID	Trait	Source Institution	Race	Sample size	Year
Exposure	ieu-b-29	Basophil cell count	BCX	European	563,946	2020
	ieu-b-30	Leukocyte count	BCX	European	563,946	ieu-b-30
	ieu-b-31	Monocyte cell count	BCX	European	563,946	2020
	ieu-b-32	Lymphocyte cell count	BCX	European	563,946	2020
	ieu-b-33	Eosinophil cell count	BCX	European	563,946	2020
	ieu-b-34	Neutrophil cell count	BCX	European	563,946	2020
Outcome	ieu-a-1126	Breast cancer	BCAC	European	228,951	2017
	ieu-a-1127	ER+ Breast cancer	BCAC	European	175,475	2017
	ieu-a-1128	ER– Breast cancer	BCAC	European	127,442	2017
	finn-b-C3_BREAST_HER2NEG	HER2– Breast cancer	FinnGen	European	–	2021
	finn-b-C3_BREAST_HERNEG	HER– Breast cancer	FinnGen	European	–	2021
	finn-b-C3_BREAST_HERPLUS	HER+ Breast cancer	FinnGen	European	–	2021

Table 1. Characteristics of peripheral blood leukocytes and breast cancer data. BCX, blood cell consortium; BCAC, Breast Cancer Association Consortium.

Ethics approval and consent to participate

Animal and human experiments were not conducted in this study.

Results

Analyzing the correlation between leukocyte counts and breast cancer risk

Previous studies have suggested that leukocyte counts in breast cancer patients may play a key role in breast cancer risk¹². We further explored this by the method of MR analysis. IVW and cML-MA were considered the primary analytical tools, and the results of the other 4 MRs were used as complementary and validation. When neither IVW nor cML-MA was statistically significant, we considered that there was no causal relationship between exposure and outcome.

The results suggest that there is no significant causal relationship between leukocyte count and breast cancer risk (Fig. 2; Supplementary data S7; IVW OR = 0.98 [95% CI: 0.93–1.03], p-value = 0.35; CML-MA OR = 1.01 [95% CI: 0.98–1.05], p-value = 0.51). Cochran's Q-test showed that there was no heterogeneity among the IVs. The results of the MR-Egger regression intercept and the MR-PRESSO test showed that there was no horizontal pleiotropy among the IVs ($P > 0.05$).

In addition, we analyzed whether there was a potential correlation between leukocyte counts and different breast cancer subtypes (ER+ Breast cancer, ER- Breast cancer, HER2- breast cancer, HER- breast cancer, and HER+ breast cancer). We did not find significant evidence to support a significant correlation between leukocyte counts and breast cancer subtypes (Fig. 3; Supplementary Table S8; All MR methods had p-value > 0.05). Therefore, we further collected different types of leukocytes in the hope of finding evidence of a correlation with breast cancer risk in categorized leukocyte counts.

We used six different MR analyses, which will be labeled with one star when the P value is < 0.05 and two stars when the P value satisfies $P = BP < 0.05/N$. CI, confidence interval; OR, odds ratio.

MR analysis of categorized white blood cell counts and breast cancer and its subtypes

We further investigated the causal relationship between the five major leukocyte types and breast cancer and its subtypes (Fig. 4; Supplementary Table S9–S13). The results showed that no significant causal relationship was found between basophil and eosinophil counts and the risk of breast cancer and its subtypes. In all analyses, horizontal pleiotropy was absent ($P > 0.05$).

In addition, it is worth noting that the results of the weighted median method suggest a potential correlation between lymphocyte counts, neutrophil counts and breast cancer risk. However, the results of both IVW and CML-MA were not significant. Therefore, we do not consider this result to be significant.

Another important finding was that the results of IVW indicated a statistically significant correlation between elevated peripheral blood monocyte counts and reduced risk of ER+ breast cancer at the genetic level (OR = 0.96, 95% CI = 0.91–1.00, $P = 0.043$). However, considering multiple comparisons, we extended the P value to BP. the IVW result did not fulfill the requirement of BP, and therefore, this result was also not significant.

In conclusion, the results of MR analysis indicated that there was no significant correlation between the counts of peripheral blood leukocytes and their subtypes and the risk of breast cancer and its subtypes.

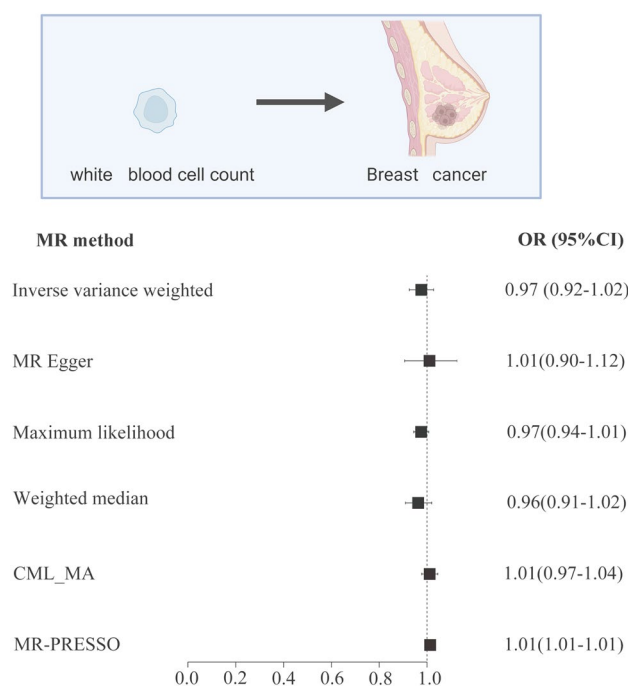


Figure 2. MR analysis between white blood cell count and breast cancer risk.

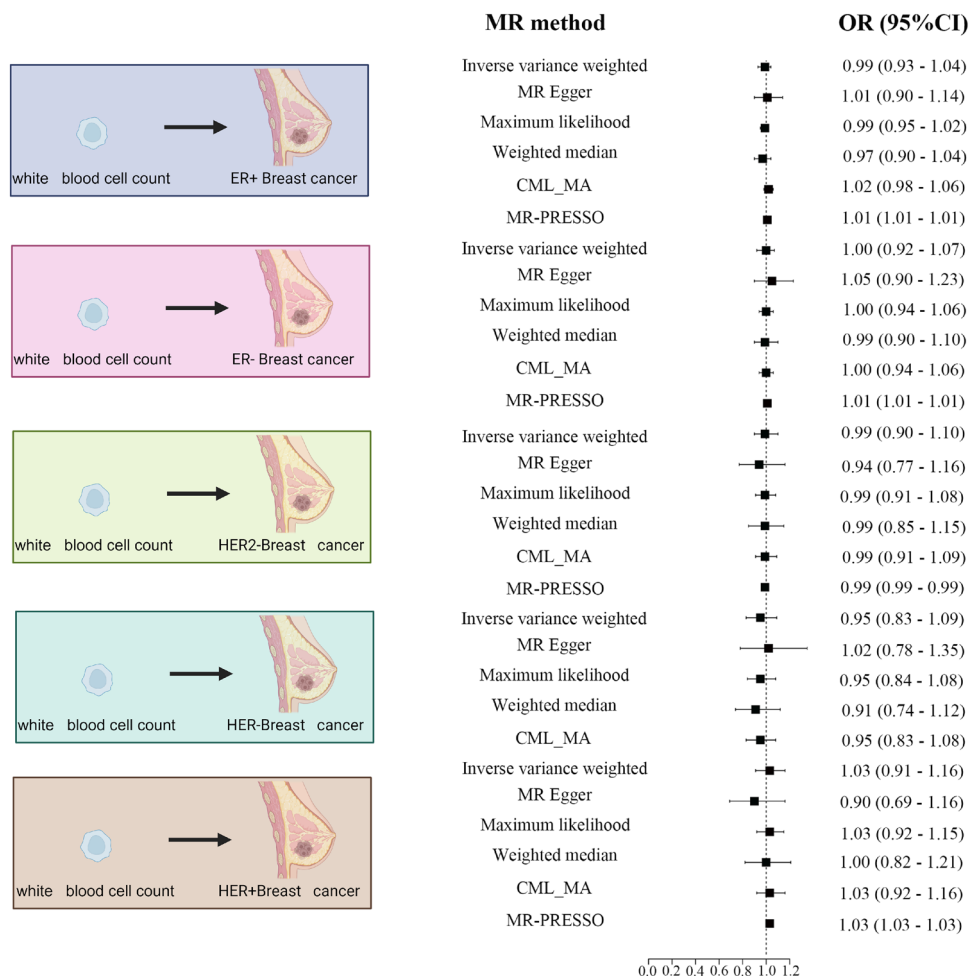


Figure 3. MR analysis of the association between white blood cell count and risk of breast cancer subtypes. The five common subtypes of breast cancer are shown from top to bottom. We used six different MR analyses, which will be labeled with one star when the P value is <0.05 and two stars when the P value satisfies $P = BP < 0.05/N$. CI, confidence interval; OR, odds ratio.

Discussion

Breast cancer is a malignant tumor that poses a serious health risk to women⁴⁹. Immune mechanisms play an obvious role in the occurrence and development of breast cancer. Therefore, in-depth studies targeting this area are beneficial to further reveal the pathogenesis of breast cancer and to improve its prevention and treatment measures. Peripheral blood leukocyte count is a routine clinical test, which is mainly used for the assessment of patients' general condition and the diagnosis of acute bacterial infections, with the characteristics of being easy to perform and low cost^{50,51}. As important immune cells in the organism, the levels of various leukocytes in peripheral blood can also reflect changes in immune function. Wei et al.⁵² recruited 140 Chinese women, 75 with breast cancer and 65 healthy controls. Their results showed that breast cancer patients had higher white blood cell counts, neutrophil counts and monocyte counts. Okutural et al.⁵³ showed that neutrophil levels are associated with risk of breast cancer. A meta-analysis evaluated the association between neutrophil-to-lymphocyte ratio as a biomarker and breast cancer prognosis using leukocyte subtypes⁵⁴. Therefore, the present study investigates the correlation between peripheral blood leukocyte sorting count levels and the risk of breast cancer development, giving researchers the opportunity to deepen their understanding of the immune mechanism of breast cancer development and clinicians the opportunity to find a simple method to assess the risk of breast cancer.

In this study we performed a large-scale MR analysis using six different MR analysis methods. We first investigated the potential relationship between leukocyte counts and the risk of breast cancer and its subtypes. The results of the MR analyses did not support a significant causal relationship between them. Then, we further investigated the relationship between leukocyte subtypes and breast cancer risk. The IVW results showed that there was a statistically significant correlation between an elevated peripheral blood monocyte count and a reduced risk of ER+ breast cancer (OR = 0.96, 95% CI = 0.91–1.00, $P = 0.043$). However, considering multiple comparisons, we extended the P-value to BP. The IVW results did not fulfill the requirement of BP, and therefore this result was also not significant. The results of our MR study showed no significant causal relationship between leukocyte count and breast cancer risk.

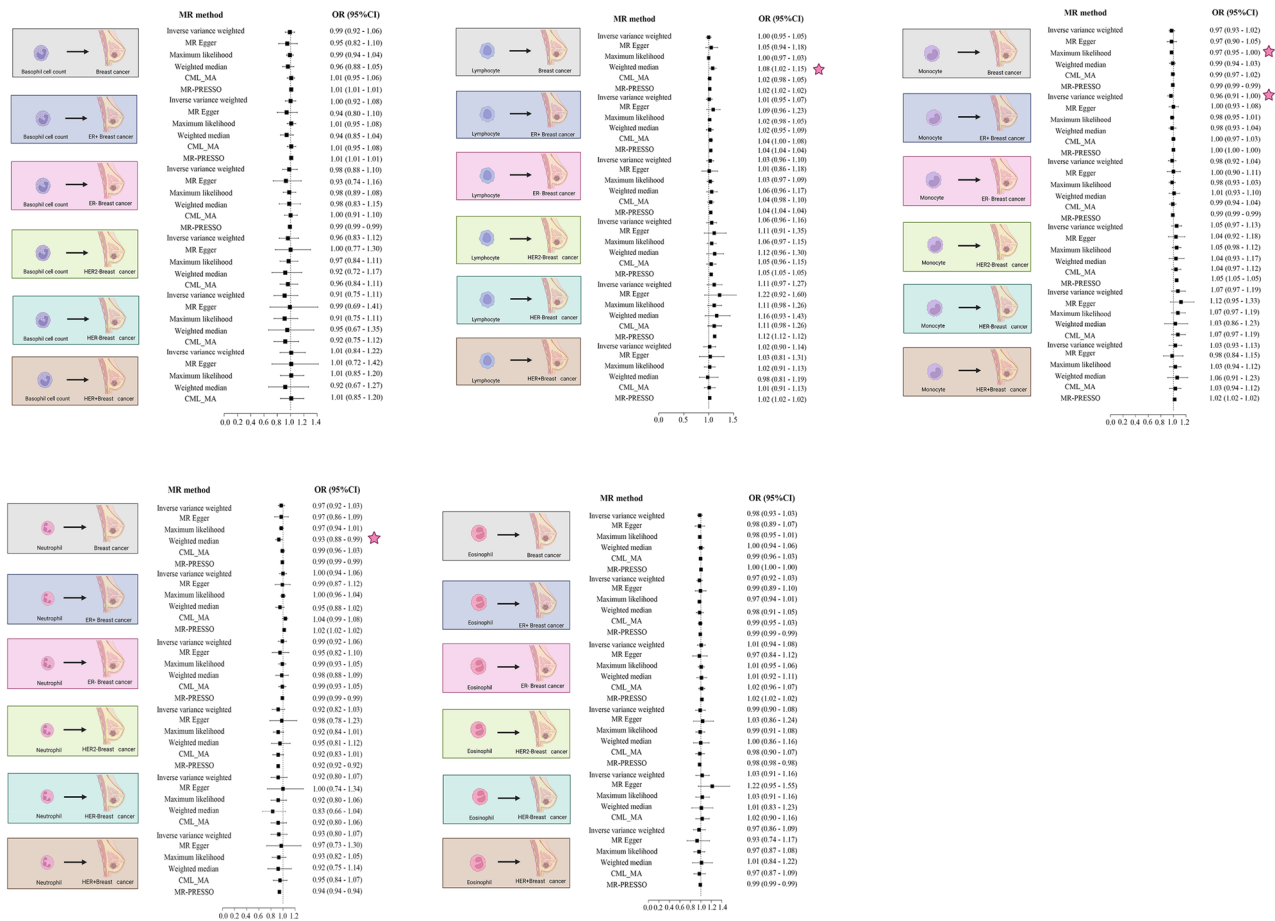


Figure 4. MR analysis of peripheral blood categorical leukocyte count and risk of breast cancer and its subtypes. Five different leukocyte types were analyzed by MR with breast cancer and its subtypes, respectively. We used six different MR analyses, which will be labeled with one star when the P value is < 0.05 and two stars when the P value satisfies $P = BP < 0.05/N$. CI, confidence interval; OR, odds ratio.

The present study is a two-sample MR study, which has the greatest advantage of avoiding causal reversal and minimizing the effect of confounding factors by using instrumental variables in place of phenotype for causal inference. The subjects of the breast cancer GWAS data in this study were all female, whereas the leukocyte GWAS data in this study were from a population of both sexes. This may affect the reliability of the study results to some extent. Current data do not support making gender distinctions. The results of previous studies have shown that a patient’s leukocyte level can be affected by a variety of factors, such as time of diagnosis, body weight, hormone levels, and menopausal status^{12,55}. In a recent study, Farrell et al.⁵⁶ collected clinical characteristics of 37,052 men and 15,004 women, and their results indicated that the white blood cell count levels were approximately 6.0 ± 1.4 ($10^9/L$) in men and 5.9 ± 1.5 ($10^9/L$) in women. Gender may not be the main influencing factor of white blood cell count. The patient’s body mass index and age may be more important influencing factors.

Overall, the results of the MR study showed no significant correlation between white blood cell counts and breast cancer risk. The results of this paper need to be further validated in clinical trials and in larger patient cohorts.

Data availability

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Received: 28 April 2023; Accepted: 7 October 2023
 Published online: 09 October 2023

References

1. Veronesi, U., Boyle, P., Goldhirsch, A., Orecchia, R. & Viale, G. Breast cancer. *Lancet* **365**(9472), 1727–1741. [https://doi.org/10.1016/S0140-6736\(05\)66546-4](https://doi.org/10.1016/S0140-6736(05)66546-4) (2005).
2. Partridge, A. H. & Carey, L. A. Unmet needs in clinical research in breast cancer: Where do we need to go?. *Clin. Cancer Res.* **23**(11), 2611–2616. <https://doi.org/10.1158/1078-0432.CCR-16-2633> (2017).
3. Maughan, K. L., Lutterbie, M. A. & Ham, P. S. Treatment of breast cancer. *Am. Fam. Phys.* **81**(11), 1339–1346 (2010).

4. Barzaman, K. *et al.* Breast cancer: Biology, biomarkers, and treatments. *Int. Immunopharmacol.* **84**, 106535. <https://doi.org/10.1016/j.intimp.2020.106535> (2020) (Epub 2020 Apr 29).
5. Makhoul, I., Atiq, M., Alwbari, A. & Kieber-Emmons, T. Breast cancer immunotherapy: An update. *Breast Cancer (Auckl)*. **30**(12), 1178223418774802. <https://doi.org/10.1177/1178223418774802> (2018).
6. Nag, S. *et al.* Risk factors for the development of triple-negative breast cancer versus non-triple-negative breast cancer: A case-control study. *Sci. Rep.* **13**(1), 13551. <https://doi.org/10.1038/s41598-023-40443-8> (2023).
7. Nishimura, T. *et al.* Evolutionary histories of breast cancer and related clones. *Nature* **620**(7974), 607–614. <https://doi.org/10.1038/s41586-023-06333-9> (2023) (Epub 2023 Jul 26).
8. Prat, A. *et al.* Clinical implications of the intrinsic molecular subtypes of breast cancer. *Breast*. **24**(Suppl 2), S26–35. <https://doi.org/10.1016/j.breast.2015.07.008> (2015) (Epub 2015 Aug 5).
9. Mehraj, U., Dar, A. H., Wani, N. A. & Mir, M. A. Tumor microenvironment promotes breast cancer chemoresistance. *Cancer Chemother. Pharmacol.* **87**(2), 147–158. <https://doi.org/10.1007/s00280-020-04222-w> (2021) (Epub 2021 Jan 9).
10. Xu, Q., Chen, S., Hu, Y. & Huang, W. Landscape of immune microenvironment under immune cell infiltration pattern in breast cancer. *Front. Immunol.* **27**(12), 711433. <https://doi.org/10.3389/fimmu.2021.711433> (2021).
11. Savas, P. *et al.* Clinical relevance of host immunity in breast cancer: From TILs to the clinic. *Nat. Rev. Clin. Oncol.* **13**(4), 228–241. <https://doi.org/10.1038/nrclinonc.2015.215> (2016) (Epub 2015 Dec 15).
12. Kresovich, J. K. *et al.* Prediagnostic immune cell profiles and breast cancer. *JAMA Netw. Open* **3**(1), e1919536. <https://doi.org/10.1001/jamanetworkopen.2019.19536> (2020).
13. Abbott, M. & Ustoyev, Y. Cancer and the immune system: The history and background of immunotherapy. *Semin. Oncol. Nurs.* **35**(5), 150923. <https://doi.org/10.1016/j.soncn.2019.08.002> (2019) (Epub 2019 Sep 13).
14. Xia, L. *et al.* The cancer metabolic reprogramming and immune response. *Mol. Cancer* **20**(1), 28. <https://doi.org/10.1186/s12943-021-01316-8> (2021).
15. Dieci, M. V., Griguolo, G., Miglietta, F. & Guarneri, V. The immune system and hormone-receptor positive breast cancer: Is it really a dead end?. *Cancer Treat. Rev.* **46**, 9–19. <https://doi.org/10.1016/j.ctrv.2016.03.011> (2016) (Epub 2016 Mar 28).
16. Basu, A. *et al.* Immunotherapy in breast cancer: Current status and future directions. *Adv. Cancer Res.* **143**, 295–349. <https://doi.org/10.1016/bs.acr.2019.03.006> (2019).
17. Rubin, R. White blood cells might provide clues to breast cancer risk. *JAMA* **323**(12), 1123. <https://doi.org/10.1001/jama.2020.2457> (2020).
18. Goff, S. L. & Danforth, D. N. The role of immune cells in breast tissue and immunotherapy for the treatment of breast cancer. *Clin. Breast Cancer*. **21**(1), e63–e73. <https://doi.org/10.1016/j.clbc.2020.06.011> (2021) (Epub 2020 Jul 2).
19. Tian, W., Wang, Y., Zhou, Y., Yao, Y. & Deng, Y. effects of prophylactic administration of granulocyte colony-stimulating factor on peripheral leukocyte and neutrophil counts levels after chemotherapy in patients with early-stage breast cancer: A retrospective cohort study. *Front. Oncol.* **25**(12), 777602. <https://doi.org/10.3389/fonc.2022.777602> (2022).
20. Larsson, A. M. *et al.* Peripheral blood mononuclear cell populations correlate with outcome in patients with metastatic breast cancer. *Cells* **11**(10), 1639. <https://doi.org/10.3390/cells11101639> (2022).
21. Li, M., Xu, J., Jiang, C., Zhang, J. & Sun, T. Predictive and prognostic role of peripheral blood T-cell subsets in triple-negative breast cancer. *Front. Oncol.* **15**(12), 842705. <https://doi.org/10.3389/fonc.2022.842705> (2022).
22. Sekula, P., Del Greco, M. F., Pattaro, C. & Köttgen, A. Mendelian randomization as an approach to assess causality using observational data. *J. Am. Soc. Nephrol.* **27**(11), 3253–3265. <https://doi.org/10.1681/ASN.2016010098> (2016) (Epub 2016 Aug 2).
23. Ellingjord-Dale, M. *et al.* Coffee consumption and risk of breast cancer: A Mendelian randomization study. *PLoS One*. **16**(1), e0236904. <https://doi.org/10.1371/journal.pone.0236904> (2021).
24. Zhu, M. *et al.* C-reactive protein and cancer risk: A pan-cancer study of prospective cohort and Mendelian randomization analysis. *BMC Med.* **20**(1), 301. <https://doi.org/10.1186/s12916-022-02506-x> (2022).
25. Nounu, A., Kar, S. P., Relton, C. L. & Richmond, R. C. Sex steroid hormones and risk of breast cancer: A two-sample Mendelian randomization study. *Breast Cancer Res.* **24**(1), 66. <https://doi.org/10.1186/s13058-022-01553-9> (2022).
26. Yu, X. *et al.* The association between plasma chemokines and breast cancer risk and prognosis: A mendelian randomization study. *Front. Genet.* **4**(13), 1004931. <https://doi.org/10.3389/fgene.2022.1004931> (2023).
27. Vuckovic, D. *et al.* The polygenic and monogenic basis of blood traits and diseases. *Cell* **182**(5), 1214–1231. <https://doi.org/10.1016/j.cell.2020.08.008> (2020).
28. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **30**(7), e34408. <https://doi.org/10.7554/eLife.34408> (2018).
29. Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature*. **551**(7678), 92–94 (2017).
30. Kurki, M. I. *et al.* FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**(7944), 508–518. <https://doi.org/10.1038/s41586-022-05473-8> (2023) (Epub 2023 Jan 18).
31. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**(D1), D1005–D1012. <https://doi.org/10.1093/nar/gky1120> (2019).
32. Sollis, E. *et al.* The NHGRI-EBI GWAS Catalog: Knowledgebase and deposition resource. *Nucleic Acids Res.* **51**(D1), D977–D985. <https://doi.org/10.1093/nar/gkac1010> (2023).
33. Burgess, S., Small, D. S. & Thompson, S. G. A review of instrumental variable estimators for Mendelian randomization. *Stat. Methods Med. Res.* **26**(5), 2333–2355. <https://doi.org/10.1177/0962280215597579> (2017) (Epub 2015 Aug 17).
34. Lin, Y. *et al.* Effects of glutamate and aspartate on prostate cancer and breast cancer: A Mendelian randomization study. *BMC Genom.* **23**(1), 213. <https://doi.org/10.1186/s12864-022-08442-7> (2022).
35. Tang, H., Yang, D., Han, C. & Mu, P. Smoking, DNA methylation, and breast cancer: A Mendelian randomization study. *Front. Oncol.* **28**(11), 745918. <https://doi.org/10.3389/fonc.2021.745918> (2021).
36. Beeghly-Fadiel, A. *et al.* A Mendelian randomization analysis of circulating lipid traits and breast cancer risk. *Int. J. Epidemiol.* **49**(4), 1117–1131. <https://doi.org/10.1093/ije/dyz242> (2020).
37. Wen, Y. *et al.* Breast cancer risk in patients with polycystic ovary syndrome: A Mendelian randomization analysis. *Breast Cancer Res. Treat.* **185**(3), 799–806. <https://doi.org/10.1007/s10549-020-05973-z> (2021) (Epub 2020 Oct 31).
38. Ma, Y., Jian, Z., Xiang, L. & Jin, X. Higher genetically predicted low-density lipoprotein levels increase the renal cancer risk independent of triglycerides and high-density lipoprotein levels: A Mendelian randomization study. *Int. J. Cancer.* **151**(4), 518–525. <https://doi.org/10.1002/ijc.34032> (2022) (Epub 2022 May 3).
39. Burgess, S., Butterworth, A. & Thompson, S. G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **37**(7), 658–665. <https://doi.org/10.1002/gepi.21758> (2013) (Epub 2013 Sep 20).
40. Yavorska, O. O. & Burgess, S. MendelianRandomization: An R package for performing Mendelian randomization analyses using summarized data. *Int. J. Epidemiol.* **46**(6), 1734–1739. <https://doi.org/10.1093/ije/dyx034> (2017).
41. Bowden, J., Davey Smith, G., Haycock, P. C. & Burgess, S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet. Epidemiol.* **40**(4), 304–314. <https://doi.org/10.1002/gepi.21965> (2016) (Epub 2016 Apr 7).
42. Zhao, X., Yang, Y., Yue, R. & Su, C. Potential causal association between leisure sedentary behaviors, physical activity and musculoskeletal health: A Mendelian randomization study. *PLoS One*. **18**(3), e0283014. <https://doi.org/10.1371/journal.pone.0283014> (2023).

43. Sang, N. *et al.* Causal relationship between sleep traits and risk of systemic lupus erythematosus: A two-sample mendelian randomization study. *Front. Immunol.* **17**(13), 918749. <https://doi.org/10.3389/fimmu.2022.918749> (2022).
44. Li, P. *et al.* Association between gut microbiota and preeclampsia-eclampsia: A two-sample Mendelian randomization study. *BMC Med.* **20**(1), 443. <https://doi.org/10.1186/s12916-022-02657-x> (2022).
45. Xue, H., Shen, X. & Pan, W. Constrained maximum likelihood-based Mendelian randomization robust to both correlated and uncorrelated pleiotropic effects. *Am. J. Hum. Genet.* **108**(7), 1251–1269. <https://doi.org/10.1016/j.ajhg.2021.05.014> (2021).
46. Amin, H. A. *et al.* Mendelian randomisation analyses of UK Biobank and published data suggest that increased adiposity lowers risk of breast and prostate cancer. *Sci. Rep.* **12**(1), 909. <https://doi.org/10.1038/s41598-021-04401-6> (2022).
47. Rasooly, D. & Patel, C. J. Conducting a reproducible mendelian randomization analysis using the R analytic statistical environment. *Curr. Protoc. Hum. Genet.* **101**(1), 82. <https://doi.org/10.1002/cphg.82> (2019) (Epub 2019 Jan 15).
48. Bhardwaj, P. *et al.* Estrogens and breast cancer: Mechanisms involved in obesity-related development, growth and progression. *J. Steroid Biochem. Mol. Biol.* **189**, 161–170. <https://doi.org/10.1016/j.jsmb.2019.03.002> (2019) (Epub 2019 Mar 6).
49. Woolston, C. Breast cancer: 4 big questions. *Nature* **527**(7578), S120. <https://doi.org/10.1038/527S120a> (2015).
50. Ghebeh, H., Elshenawy, M. A., AlSayed, A. D. & Al-Tweigeri, T. Peripheral blood eosinophil count is associated with response to chemoimmunotherapy in metastatic triple-negative breast cancer. *Immunotherapy* **14**(4), 189–199. <https://doi.org/10.2217/imt-2021-0149> (2022) (Epub 2022 Jan 5).
51. Jalali, A. *et al.* Peripheral blood cell ratios as prognostic indicators in a neoadjuvant chemotherapy-treated breast cancer cohort. *Curr. Oncol.* **29**(10), 7512–7523. <https://doi.org/10.3390/curroncol29100591> (2022).
52. Wei, C. T. *et al.* Elevated plasma level of neutrophil gelatinase-associated lipocalin (NGAL) in patients with breast cancer. *Int. J. Med. Sci.* **18**(12), 2689–2696. <https://doi.org/10.7150/ijms.58789> (2021).
53. Okuturlar, Y. *et al.* Utility of peripheral blood parameters in predicting breast cancer risk. *Asian Pac. J. Cancer Prev.* **16**(6), 2409–2412. <https://doi.org/10.7314/apjcp.2015.16.6.2409> (2015).
54. Wei, B. *et al.* The neutrophil lymphocyte ratio is associated with breast cancer prognosis: An updated systematic review and meta-analysis. *Onco Targets Ther.* **8**(9), 5567–5575. <https://doi.org/10.2147/OTT.S108419> (2016).
55. Park, B., Lee, H. S., Lee, J. W. & Park, S. Association of white blood cell count with breast cancer burden varies according to menopausal status, body mass index, and hormone receptor status: A case-control study. *Sci. Rep.* **9**(1), 5762. <https://doi.org/10.1038/s41598-019-42234-6> (2019).
56. Farrell, S. W. *et al.* Cardiorespiratory fitness, white blood cell count, and mortality in men and women. *J. Sport Health Sci.* **11**(5), 605–612. <https://doi.org/10.1016/j.jshs.2021.10.005> (2022) (Epub 2021 Nov 3).

Acknowledgements

We are grateful to all the studies that have made the public GWAS summary data available, and to all the investigators and participants who contributed to the breast cancer study.

Author contributions

Z.Z. and J.W. designed the overall study and performed the analysis and article writing. L.L. reviewed and revised the article. All the authors have approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-44397-9>.

Correspondence and requests for materials should be addressed to L.L. or J.W.

Reprints and permissions information is available at www.nature.com/reprints.

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

© The Author(s) 2023