Body Shape Phenotypes And Breast Cancer Risk: A Mendelian Randomization Analysis.

Laia Peruchet Noray

International Agency for Research on Cancer (IARC/WHO)

Niki Dimou

International Agency for Research on Cancer (IARC/WHO)

Anja M. Sedlmeier

University of Regensburg

Béatrice Fervers

Centre Léon Bérard

Isabelle Romieu

National Institute of Public Health

Vivian Viallon

International Agency for Research on Cancer (IARC/WHO)

Marc J. Gunter

International Agency for Research on Cancer (IARC/WHO)

Pietro Ferrari

International Agency for Research on Cancer (IARC/WHO)

Robert Carreras-Torres

Institut d'Investigació Biomèdica de Girona Dr. Josep Trueta (IDIBGI)

Heinz Freisling (✉ freislingh@iarc.fr)

International Agency for Research on Cancer (IARC/WHO)

Research Article

Keywords: Breast cancer, body shape, Mendelian randomization, causal inference.

Posted Date: February 21st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1366779/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Observational and genetic studies have linked different anthropometric traits to breast cancer risk, with inconsistent results. We aimed to investigate the association between body shape defined by principal component (PC) analysis of anthropometric traits (body mass index (BMI), height, weight, waist-to-hip ratio (WHR), waist and hip circumference) and overall breast cancer risk and by sub-type (luminal A, luminal B, HER2+, triple negative and luminal B/HER2 negative).

Methods: We performed two-sample Mendelian randomization analyses to assess the association between 188 genetic variants robustly linked to the first three PCs and breast cancer (133,384 combined pre- and post-menopausal cases/113,789 controls from the Breast Cancer Association Consortium (BCAC)).

Results: PC1 (general adiposity) was inversely associated with overall breast cancer risk (odds ratio (OR) 0.89 [95% confidence interval (CI) 0.81-0.98]; p-value = 0.016). PC2 (tall with low WHR) was weakly positively associated with overall breast cancer risk (OR = 1.05 [95% CI: 0.98-1.12]; p-value = 0.135), but with a confidence interval including the null. PC3 (tall with large WHR) was not associated with overall breast cancer risk. Some of these associations differed by breast cancer sub-types. For instance, PC2 was positively associated with risk of luminal A breast cancer sub-type (OR = 1.09 [95% CI: 1.01-1.18]; p-value = 0.02).

Conclusions: Our study provides evidence for potential causal associations between body shape and breast cancer risk and breast cancer sub-types highlighting the importance to also assess body morphology holistically.

Background

Breast cancer is the most frequently diagnosed cancer in women worldwide (2.26 million new cases in 2020) and is the leading cause of cancer mortality in women (684,996 deaths in 2020).¹ Established risk factors for female breast cancer include sex, age, family history, alcohol consumption, ionising radiation exposure, reproductive factors such as early menarche, late menopause, late age at first pregnancy, low parity, and elevated levels of both endogenous and exogenous estrogen.²

There is inconsistent evidence from prospective observational and genetic studies linking body fatness, mostly assessed by body mass index (BMI), to breast cancer incidence.³⁻⁵ BMI has been reported to be either positively or inversely associated with breast cancer depending on the menopausal status and the molecular breast cancer sub-type.³⁻⁵ Waist circumference (WC) has been positively associated with breast cancer risk among post-menopausal women, while the association with waist-to-hip ratio (WHR) remains inconclusive.^{6,7} Height has been positively associated with risk of overall and hormone receptor-positive breast cancer.⁸

While a correlation between different anthropometric variables exists, individuals similar in one trait (e.g. BMI) may considerably differ for others (e.g. WHR) resulting in different body shapes. A promising approach to capture the complexity of body shape phenotypes is to combine multiple traits. Recently, a principal component analysis (PCA) was performed on six anthropometric traits (height, weight, BMI, WC, hip circumference (HC), and WHR) to define body shape.⁹ Each principal component (PC) represented a specific pattern of the six anthropometric traits and captured distinct body shape dimensions.

Different biological mechanisms may underlie the heterogeneity of body morphology with breast cancer risk, and the different body shapes may confer differences in breast cancer risk. However, considerable uncertainty remains about the causal effect of these anthropometric phenotypes, individually or in combination, on overall breast cancer risk and according to sub-type.

Mendelian randomization (MR) analysis uses genetic variants as instrumental variables (IVs) to strengthen causality in exposure-outcome associations. This methodology can overcome some of the limitations of conventional approaches such as reverse causation and unmeasured confounding, provided that the key assumptions of MR are not violated.¹⁰

In this study a two-sample MR analysis was conducted to investigate causal associations between three body shape phenotypes, as derived in Ried et al.,⁹ and the risk of breast cancer using summary level data from the Breast Cancer Association Consortium (BCAC).¹¹ We also investigated associations by hormonal receptor status breast cancer sub-types.

Methods

Body shape phenotypes and related genetic variants

Ried et al. performed a PCA of six anthropometric traits (BMI, height, weight, WC, HC and WHR) obtaining three body shape phenotypes.⁹ The first three PCs explained 96.7% of the variation of these six anthropometric traits.⁹ PC1 had equally high positive loadings on all traits, except for height, characterizing individuals with general adiposity.

PC2 had high but opposite loadings on height and WHR, capturing variation in a composite body shape that represents tall individuals with a low WHR or vice versa, short individuals with a large WHR.

PC3 had high positive loadings on height and WHR, and an opposite loading of nearly the same size on HC. PC3 characterizes tall individuals with a large WHR resulting from a smaller HC or vice versa, short individuals with a low WHR and a high HC.⁹ Further details regarding PC loadings can be found in Supplementary Table 1.

Ried et al. also reported a fourth body shape (PC4) characterizing individuals with high BMI and weight with low WC and HC (athletic individuals) and vice versa.⁹ Although we also attempted to analyze PC4 as

an additional exposure, results obtained are not reported in this manuscript due to their imprecision and the impossibility to build solid conclusions.

MR is based on three key assumptions: (1) the genetic variants are associated with the exposure, (2) the genetic variants are not associated with any potentially confounding variable and (3) the genetic variants are only associated with the outcome through the risk factor.^{12,13} To validate the first MR assumption, we used genetic variants identified from a two-stage meta-analysis of GWAS performed on the three first body shapes.⁹ The two GWAS steps included more than 170,000 individuals of European descent from 65 studies from the GIANT Consortium obtaining significant associations ($p < 5 \times 10^{-8}$) of 200 loci (31 for PC1, 124 for PC2 and 45 for PC3), some of them not previously associated with single anthropometric traits.⁹ On this basis, we excluded 11 genetic variants after linkage disequilibrium analysis ($LD\ r^2 \leq 0.01$) to avoid correlation between genetic variants, using the 1000 Genomes Project European samples. Genetic variant summary statistics for associations with PC1-3 were obtained for the remaining 189 genetic variants (30 for PC1, 115 for PC2 and 44 for PC3, considering 4 genetic variants related to both PC2 and PC3).

Breast cancer risk data

Genetic variant summary statistics for associations with risk of overall breast cancer and 5 sub-types (triple negative tumors, HER2 positive tumors regardless of estrogen receptor (ER) status, luminal B/HER2-negative-like and two types of ER positive breast cancer: luminal-A- and -B-like)¹⁴ were extracted from a GWAS of 247,173 individuals (133,384 breast cancer cases and 113,789 controls) of European ancestry carried out by the Breast Cancer Association Consortium (BCAC).¹¹ For 6 genetic variants, the use of a proxy genetic variant was needed ($r^2 \geq 0.8$; European samples); while for 1 genetic variant, no data were obtained. More detailed information on sample size and tumor heterogeneity is provided in Supplementary Table 2.

Statistical analysis

We performed a two-sample MR analysis to estimate the causal relation between independent body shape phenotypes (PC1, PC2 and PC3) and breast cancer risk (Fig. 1).

We used the fixed-effects inverse-variance weighted (IVW) method as the primary approach.¹⁵ This analysis can be understood as a meta-analysis of genetic variant effects which are calculated as the ratio of the genetic variant-breast cancer over the genetic variant-body shape association.^{12,16}

Sensitivity analyses

As the second (i.e. genetic variants are not associated with any potentially confounding variable) and the third (i.e. genetic variants are only associated with the outcome through the risk factor) MR assumptions are challenging to test,¹⁷ further analyses were performed to assess potential violation of both assumptions.

First, we used PhenoScanner (available at <http://www.phenoscanter.medschl.cam.ac.uk>) and the GWAS catalog (available at <https://www.ebi.ac.uk/gwas/>) to search whether our selected genetic variants were associated with other traits in previous GWAS.^{17–19}

Second, we used the Cochran's Q test to evaluate heterogeneity of causal effects from each variant, with a p-value < 0.05 indicating statistical significance.²⁰ In case of heterogeneity, likely indicating pleiotropy²¹, a random-effects IVW MR analysis was used.²² This method relaxes the third assumption as the total pleiotropic effect of a single genetic variant no longer needs to be null but assumes zero mean between all the genetic variants (i.e. balanced pleiotropy).²² In addition, we implemented MR-Egger regression²³ and the weighted median approach.²⁴ The MR-Egger methodology provides a statistical test for overall directional pleiotropy based on whether the intercept term is different from zero, as well as the I^2_{GX} statistic, which indicates unreliability of MR-Egger inferences at values below 90%.^{23,25} The weighted median estimator allows for the violations of the second and the third assumptions when up to 50% of the genetic variants are invalid (i.e. violation of one or more MR assumptions).^{21,24} Both tests provide valid MR estimates in the presence of overall directional pleiotropy but suffer from reduced power.^{23,24} Finally, we also used the MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method to identify and remove any outlying variants²⁶ and leave-one genetic variant-out analysis were used to assess whether any association was driven by specific genetic variants.

We used scatter plots to present the genetic associations between body shape phenotypes and breast cancer risk, in combination with funnel plots, to visually examine the consistency of MR estimates and the potential associated bias (Supplementary Fig. 1).

A two-sided p-value < 0.05 was considered statistically significant in all analyses.

All analyses were conducted in the statistical software R 4.0.4 and RStudio 1.4.1106 using the MendelianRandomization and MRPRESSO R packages.^{26,27}

Results

Body shape phenotypes and risk of overall breast cancer

PC1 was inversely associated with breast cancer risk (IVW_{random-effects} OR per 1 SD equal to 0.89 [95% CI: 0.81–0.98]; p-value = 0.016) (Fig. 2). There was evidence of heterogeneity across genetic variants (p-value < 0.001) and of overall directional pleiotropy (Egger intercept = 0.029 [95% CI: 0.014–0.043]; p-value < 0.001) (Supplementary Table 3). Nevertheless, results of the random IVW after performing MR PRESSO, MR Egger, and the weighted median supported the inverse association (Egger OR = 0.59 [95% CI: 0.48–0.74]; p-value < 0.001; weighted median OR = 0.81 [95% CI: 0.76–0.88]; p-value < 0.001) (Supplementary Fig. 2.A). The leave-one genetic variant-out analysis showed an estimate closer to the null when excluding the genetic variant rs1121980 (a variant in the FTO gene) (Supplementary Fig. 3). This genetic

variant, along with the rs1582874 variant (associated with height and other traits) (Supplementary File 1), were identified as outliers by the MR-PRESSO method. No distortion of the overall risk estimate by the presence of these two outlier genetic variants was detected (p distortion = 0.13) (Supplementary Table 4).

PC2 was weakly positively associated with the risk of breast cancer (IVW_{random-effects} OR = 1.05 [95% CI: 0.98–1.12]; p -value = 0.135) (Fig. 2), but the confidence interval included the null. We found presence of heterogeneity across genetic variants (p -value < 0.001), but directional pleiotropy was not detected (Supplementary Table 3). Accordingly, the dispersion of the estimates provided by the genetic variants in the funnel plot showed balanced pleiotropy (Supplementary Fig. 4). However, MR PRESSO indicated that the presence of five outlier genetic variants was potentially biasing the initial risk estimate (p distortion = 0.013) (Supplementary Table 4). The weighted median and IVW random MR approaches without outlier genetic variants estimations were attenuated toward the null (Supplementary Fig. 2.B).

There was no evidence of an association between PC3 and breast cancer risk (IVW_{random-effects} OR = 1.04 [95% CI: 0.89–1.21]; p -value = 0.619) (Fig. 2). We obtained evidence of heterogeneity (p -value < 0.001), but without presence of overall directional pleiotropy (Supplementary Table 3). The sensitivity analysis supported a null association (Supplementary Fig. 2.C).

Body shape phenotypes and risk of breast cancer sub-types

PC1 was inversely associated with risk of all five breast cancer sub-types, of which associations were most consistent for HER2+ (IVW_{random-effects} OR = 0.81 [95% CI: 0.67–0.99]; p -value = 0.043) and triple negative sub-types (IVW_{random-effects} OR = 0.82 [95% CI: 0.73–0.91]; p -value < 0.001) (Fig. 2). Nevertheless, all of the associations showed heterogeneity and overall directional pleiotropy (Supplementary Table 3). Sensitivity analysis accounting for unbalanced pleiotropy and outliers supported the inverse associations of the main analysis (Supplementary Fig. 2.A). MR-PRESSO classified genetic variants as outliers in all the sub-types but only in luminal A sub-type identified biased association estimates (p distortion = 0.02) (Supplementary Table 4).

Comparable to the overall breast cancer risk analysis, PC2 was positively associated with all breast cancer sub-types, except for HER2+, where we found a suggestive inverse association (IVW_{random-effects} OR = 0.96 [95% CI: 0.82–1.11]; p -value = 0.549) (Fig. 2). The only consistent positive association was observed for luminal A (IVW_{random-effects} OR = 1.09 [95% CI: 1.01–1.18]; p -value = 0.02) with presence of balanced pleiotropy (heterogeneity p -value_{luminal A} < 0.001; Egger intercept p -value_{luminal A} = 0.256) (Supplementary Table 3).

PC3 was suggestively inversely associated with the risk of the HER2 + sub-type breast cancer (IVW_{random-effects} OR = 0.75 [95% CI: 0.53–1.08]; p -value = 0.122) (Fig. 2). In contrast, positive associations were observed with risks of triple negative and luminal B/HER2 negative sub-types of breast cancer, although risk estimates were imprecise (OR_{triple negative} = 1.14 [95% CI 0.88–1.48]; OR_{luminal B/HER2 negative} = 1.13 [95% CI 0.88–1.44]). After accounting for unbalanced pleiotropy and outliers, results

supported these suggestive associations except the Egger test which presented unreliable results ($I^2_{GX} \sim 0\%$) like in the overall breast cancer analysis (Supplementary Fig. 2.C).

Discussion

We performed a large-scale MR analysis to estimate causal associations between three distinct body shape phenotypes and the risk of overall breast cancer and five breast cancer sub-types. A body shape phenotype capturing general adiposity (PC1), with positive loadings on all anthropometric traits except height, was inversely associated with risk of overall breast cancer, with consistent associations across breast cancer sub-types. A body shape phenotype contrasting tall women with a low WHR to short women with a large WHR (PC2) was positively associated with risk of the luminal A sub-type of breast cancer, while associations with risk of overall breast cancer or other sub-types were less clear. There was no evidence of an association between PC3, a body shape phenotype contrasting tall women with a large WHR resulting from a small HC to short women with a low WHR and a high HC, and risk of overall breast cancer. There was suggestive evidence that PC3 was differentially associated with breast cancer sub-types (inversely associated with HER2+, but positively associated with triple negative and luminal B/HER2 negative).

To the best of our knowledge, this is the first MR analysis investigating the association between body shape phenotypes and breast cancer risk. A few MR studies reported associations between individual anthropometric exposures used to construct the body shape PCs and breast cancer risk.^{4,7,8,28–30}

PC1 was characterized by general adiposity, often defined by BMI. A 2019 MR analysis found that BMI was inversely associated with the risk of breast cancer, as well as both ER+ (luminal A, luminal B and luminal B/HER2 negative) and ER- (HER2 + and triple negative sub-types). The study used the same genetic data (BCAC) as we did and reported inverse associations between BMI and overall breast cancer risk.²⁸ Conversely, the associations observed for the ER + and ER- sub-groups slightly differed from our results. While our results pointed to a 10% and 20% reduced risk, in ER + and ER- related sub-types breast cancer, respectively, Ooi et al. did not find any difference between ER+/- and reported an OR close to 0.8 for both sub-types.²⁸ This difference could be due to the specificity of PC1, which captures a body shape that goes beyond BMI. PC1 could thus be complementary to and potentially more complete than BMI in characterizing general adiposity. This assertion is supported by the finding that two novel loci for PC1 were identified by Ried et al. that were not identified before in larger GWAS analyses for BMI, WHRadjBMI, and height.⁹

Observational studies reported a positive association between BMI and breast cancer in post-menopausal women and an inverse association in pre-menopausal women.^{31–34} In post-menopausal women, adipose tissue largely catalyzes the conversion of adrenal androgens into estrogen by aromatase enzyme, processed known as estrogen biosynthesis. The presence of estrogen, alone and with progestin, is associated with post-menopausal breast cancer risk.³⁵ In pre-menopausal women, large BMI is generally

associated with longer anovulatory cycles resulting in low levels of estrogen and progestin.³⁶ These were potential arguments proposed to explain why ER + breast cancer is more common in post-menopausal women than ER-, which is more likely in pre-menopausal stages. Although we were not able to differentiate between pre- and post-menopausal breast cancer with our summary-level genetic data, our results remain consistent with other MR analyses that used individual-level data and found that BMI was inversely associated with breast cancer risk, independently of menopausal stage.^{4,29}

Positive associations between body height and breast cancer risk using both traditional epidemiological approaches (OR per 9.2cm = 1.08 [95% CI: 1.06–1.10]) and genetic approaches (OR per 9.2cm = 1.07 [95% CI: 1.03–1.11]) have been reported.³⁰ MR studies found similar associations, reporting a risk increased by a magnitude of 20% per 10cm increase.⁸ Our association (OR = 1.05 [95% CI: 0.98 – 0.13]) was slightly weaker than reported in the literature for height, which can potentially be explained by the fact that PC2 in our study was also characterized by differences in WHR. A meta-analysis of prospective observational studies reported a positive association between a large WHR and breast cancer risk (OR = 1.06 [95% CI: 0.99–1.13]) in postmenopausal women.⁶ In contrast, Gao et al. reported an inverse association between WHR and breast cancer risk (OR = 0.73 [95% CI: 0.53-1.00]) using MR.⁷ In our study, WHR had a weak positive contribution in PC1 but a strong negative one in PC2.⁹

Regarding breast cancer sub-types, only luminal A (ER+/PR+) breast cancer showed a positive association with PC2. Results on other sub-types were less clear. Zhang et al. reported a positive association between height and all hormone receptor-positive (ER+/PR+) sub-type breast cancers and no relationship for hormone receptor-negative (ER-/PR-) sub-types.⁸ The lack of a consistent positive association between PC2 and other hormone receptor-positive sub-types (i.e. luminal B and luminal B/HER2 negative) in our analysis could be due to the lower number of genetic variants linked to height (114 variants; see Supplementary File 1) as compared to Zhang et al. (168 variants).⁸

PC3, indicating body height combined with small HC and large WHR⁹, was not associated with risk of overall and most sub-types of breast cancer, except for the HER2 + sub-type that showed an inverse association. A high positive estimate between PC3 with triple negative and luminal B/HER2 negative breast cancer sub-types was observed, although they were imprecise. Similar as for PC1, Ried et al. identified two novel loci for PC3 not previously related to anthropometric traits, which indicates that additional phenotype information is captured by integrating multiple anthropometric traits. They also argue that these novel loci may highlight biological pathways that would go undetected with single-trait analyses.⁹ This could at least partly explain why body shape phenotypes – as compared to single anthropometric traits – appear be more specific in characterizing risk of breast cancer sub-types.

Two body shape phenotypes (PC1 and PC3) were inversely associated with the risk of HER2 + breast cancer sub-types, albeit this association was only statistically significant for PC1. In previous literature, HER2 + sub-type tumors have been identified as a distinct subset (comprising 20–25%) of breast carcinomas and being differently associated with hormone-related risk factors.^{37,38}

In this study we employed different MR approaches: IVW-based methods, methods based on Egger regression and median-based methods. IVW methods are generally more powerful compared to the other methods implemented but they do not account for directional (unbalanced) pleiotropy. MR-Egger accounts for heterogeneity due to pleiotropy, although it results in a less powerful method. Median-based methods are between both IVW and MR-Egger approaches as they are more robust to directional pleiotropy than IVW and they allow up to 50% of invalid (i.e. violating one or more of the three assumptions of MR) genetic variants to be discarded.³⁹

Although we checked for associations between our genetic variants and confounders, a violation of the second assumption of the MR approach cannot be entirely excluded due to the impossibility to check unknown confounding factors. Other limitations include that our study may be biased due to uncontrolled confounding from family effects such as assortative mating, dynastic effects, and population structure.⁴⁰ We minimized this bias from population stratification performing a two-sample MR analysis in samples composed of Caucasians individuals only. However, this may cause our findings to be not generalizable to other ethnic groups. Referring to the body shapes, it is difficult to quantify the effect of PCs on breast cancer as it is not evident to what 1 SD corresponds in terms of anthropometry. Finally, it is important to highlight that with these genetic data for breast cancer, we were not able to differentiate between pre- and post-menopausal breast cancer. However, ~ 85% of breast cancer cases were post-menopausal, which means that our findings are more representative of post-menopausal breast cancer.

Conclusions

Our MR analysis provides evidence of an inverse association between a body shape phenotype that reflects general adiposity and risk of overall and in particular HER2 + and triple negative sub-types of breast cancer. Furthermore, we found a positive association between a body shape phenotype related to body height with a low WHR and luminal A sub-type breast cancer risk. A body shape related to body height with a large WHR resulting from smaller HC was not associated with breast cancer risk. The present study adds a complementary and promising approach to interrogate multiple anthropometric measurements in relation to breast cancer risk.

Abbreviations

BC	Breast Cancer
BMI	Body Mass Index
BCAC	Breast Cancer Association Consortium
CI	Confidence Interval
DAG	Directed Acyclic Graph

ER	Estrogen Receptor
GWAS	Genome-Wide Association Study
HC	Hip Circumference
IV	Instrumental Variable
IVW	Inverse Variance Weighted
MR	Mendelian Randomization
MR-PRESSO	MR Pleiotropy RESsidual Sum and Outlier
OR	Odds Ratio
PC	Principal Component
PCA	Principal Component Analysis
SD	Standard Deviation
SNP	Single-Nucleotide Polymorphism
WC	Waist Circumference
WHR	Waist-to-Hip Ratio

Declarations

Funding: This work was supported by the French National Cancer Institute (INCA_N°2019-176).

Ethic approval and consent to participate

All studies participating in BCAC were conducted in accordance with the Declaration of Helsinki. For each study, investigators satisfied the local requirements for ethical research, including obtaining informed consent from participants.

Consent for publication

This article does not contain any individual-level data.

Availability of data and materials

All relevant data are within the manuscript and its Supplementary information.

Competing interests

The authors declare no conflict of interest.

Funding

This study was supported by the French National Cancer Institute (INCA_N°2019-176).

Author's contributions

The idea for the study was developed by HF and further specified by PF, VV, and BF. HF, ND, and RCT designed the study and protocol. LPN conducted the statistical analysis and supervision of ND and RCT. LPN and HF drafted the manuscript. All authors contributed to the manuscript through critical revision of drafts and approved the final manuscript.

Acknowledgements

The breast cancer genome-wide association analyses for BCAC and CIMBA were supported by Cancer Research UK (C1287/A10118, C1287/A16563, C1287/A10710, C12292/A20861, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), The National Institutes of Health (CA128978, X01HG007492- the DRIVE consortium), the PERSPECTIVE project supported by the Government of Canada through Genome Canada and the Canadian Institutes of Health Research (grant GPH-129344) and the Ministère de l'Économie, Science et Innovation du Québec through Genome Québec and the PSRSIIRI-701 grant, the Quebec Breast Cancer Foundation, the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), the European Union's Horizon 2020 Research and Innovation Programme (634935 and 633784), the Post-Cancer GWAS initiative (U19 CA148537, CA148065 and CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer (CRN-87521), the Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund. All studies and funders are listed in Zhang H et al (Nat Genet, 2020).

Disclaimer

Where authors are identified as personnel of the International Agency for Research on Cancer / World Health Organization, the authors alone are responsible for the views expressed in this article and they do not necessarily represent the decisions, policy or views of the International Agency for Research on Cancer / World Health Organization.

References

1. Ferlay J, Ervik M, Lam F, Colombet M, Mery L, Piñeros M, et al. Global Cancer Observatory: Cancer Today. Int Agency Res Cancer. 2020;68(6).
2. WCRF International WCRF. Diet, nutrition, physical activity and breast cancer. Am Inst Cancer Res. 2018;

3. Fortner RT, Katzke V, Kühn T, Kaaks R. Obesity and breast cancer. In: *Recent Results in Cancer Research*. 2016.
4. Guo Y, Warren Andersen S, Shu XO, Michailidou K, Bolla MK, Wang Q, et al. Genetically Predicted Body Mass Index and Breast Cancer Risk: Mendelian Randomization Analyses of Data from 145,000 Women of European Descent. *PLoS Med*. 2016;13(8).
5. Rosner B, Eliassen AH, Toriola AT, Chen WY, Hankinson SE, Willett WC, et al. Weight and weight changes in early adulthood and later breast cancer risk. *Int J Cancer*. 2017;140(9).
6. Chen GC, Chen SJ, Zhang R, Hidayat K, Qin JB, Zhang YS, et al. Central obesity and risks of pre- and postmenopausal breast cancer: a dose–response meta-analysis of prospective studies. Vol. 17, *Obesity Reviews*. 2016.
7. Gao C, Patel CJ, Michailidou K, Peters U, Gong J, Schildkraut J, et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Int J Epidemiol*. 2016;45(3).
8. Zhang B, Shu XO, Delahanty RJ, Zeng C, Michailidou K, Bolla MK, et al. Height and breast cancer risk: Evidence from prospective studies and mendelian randomization. *J Natl Cancer Inst*. 2015;107(11).
9. Ried JS, Jeff JM, Chu AY, Bragg-Gresham JL, Van Dongen J, Huffman JE, et al. A principal component meta-analysis on multiple anthropometric traits identifies novel loci for body shape. *Nat Commun*. 2016;7.
10. Smith GD, Ebrahim S. “Mendelian randomization”: Can genetic epidemiology contribute to understanding environmental determinants of disease? Vol. 32, *International Journal of Epidemiology*. 2003.
11. Zhang H, Ahearn TU, Lecarpentier J, Barnes D, Beesley J, Qi G, et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat Genet*. 2020;52(6).
12. Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Smith GD. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008;27(8).
13. Didelez V, Sheehan N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat Methods Med Res*. 2007;16(4).
14. Curigliano G, Burstein HJ, Winer EP, Gnant M, Dubsy P, Loibl S, et al. De-escalating and escalating treatments for early-stage breast cancer: The St. Gallen International Expert Consensus Conference on the Primary Therapy of Early Breast Cancer 2017. *Ann Oncol*. 2017;28(8).
15. Burgess S, Scott RA, Timpson NJ, Smith GD, Thompson SG. Using published data in Mendelian randomization: A blueprint for efficient identification of causal risk factors. *Eur J Epidemiol*. 2015;30(7).
16. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7).
17. Dimou NL, Tsilidis KK. A primer in Mendelian Randomization methodology with a focus on utilizing published summary association data. In: *Methods in Molecular Biology*. 2018.

18. MacArthur J, Bowler E, Cerezo M, Gil L, Hall P, Hastings E, et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* 2017;45(D1).
19. Staley JR, Blackshaw J, Kamat MA, Ellis S, Surendran P, Sun BB, et al. PhenoScanner: A database of human genotype-phenotype associations. *Bioinformatics.* 2016;32(20).
20. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Vol. 327, *British Medical Journal.* 2003.
21. Pagoni P, Dimou NL, Murphy N, Stergiakouli E. Using Mendelian randomisation to assess causality in observational studies. Vol. 22, *Evidence-Based Mental Health.* 2019.
22. Bowden J, Del Greco M F, Minelli C, Davey Smith G, Sheehan N, Thompson J. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat Med.* 2017;36(11).
23. Bowden J, Smith GD, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol.* 2015;44(2).
24. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet Epidemiol.* 2016;40(4).
25. Bowden J, Fabiola Del Greco M, Minelli C, Smith GD, Sheehan NA, Thompson JR. Assessing the suitability of summary data for two-sample mendelian randomization analyses using MR-Egger regression: The role of the I² statistic. *Int J Epidemiol.* 2016;45(6).
26. 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).
27. Yavorska OO, Burgess S. MendelianRandomization: An R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol.* 2017;46(6).
28. Ooi BNS, Loh H, Ho PJ, Milne RL, Giles G, Gao C, et al. The genetic interplay between body mass index, breast size and breast cancer risk: A Mendelian randomization analysis. *Int J Epidemiol.* 2019;48(3).
29. Shu X, Wu L, Khankari NK, Shu XO, Wang TJ, Michailidou K, et al. Associations of obesity and circulating insulin and glucose with breast cancer risk: A Mendelian randomization analysis. *Int J Epidemiol.* 2019;48(3).
30. Lai FY, Nath M, Hamby SE, Thompson JR, Nelson CP, Samani NJ. Adult height and risk of 50 diseases: A combined epidemiological and genetic analysis. *BMC Med.* 2018;16(1).
31. Kabat GC, Xue X, Kamensky V, Lane D, Bea JW, Chen C, et al. Risk of breast, endometrial, colorectal, and renal cancers in postmenopausal women in association with a body shape index and other anthropometric measures. *Cancer Causes Control.* 2015;26(2).
32. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet.* 2014;384(9945).

33. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *Lancet*. 2008;371(9612).

34. Liu K, Zhang W, Dai Z, Wang M, Tian T, Liu X, et al. Association between body mass index and breast cancer risk: Evidence based on a dose–response meta-analysis. *Cancer Manag Res*. 2018;10.

35. Picon-Ruiz M, Morata-Tarifa C, Valle-Goffin JJ, Friedman ER, Slingerland JM. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *CA Cancer J Clin*. 2017;67(5).

36. Fichman V, Costa R de SS da, Miglioli TC, Marinheiro LPF. Association of obesity and anovulatory infertility. *Einstein (Sao Paulo)*. 2020;18.

37. Ménard S, Balsari A, Casalini P, Tagliabue E, Campiglio M, Bufalino R, et al. HER-2-positive breast carcinomas as a particular subset with peculiar clinical behaviors. *Clin Cancer Res*. 2002;8(2).

38. Balsari A, Casalini P, Bufalino R, Berrino F, Ménard S. Role of hormonal risk factors in HER2-positive breast carcinomas. *Br J Cancer*. 2003;88(7).

39. Pingault JB, O'Reilly PF, Schoeler T, Ploubidis GB, Rijdsdijk F, Dudbridge F. Using genetic data to strengthen causal inference in observational research. Vol. 19, *Nature Reviews Genetics*. 2018.

40. Brumpton B, Sanderson E, Heilbron K, Hartwig FP, Harrison S, Vie GÅ, et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat Commun*. 2020;11(1).

Figures

Figure 1

Causal directed acyclic graph (DAG) of Mendelian Randomization analysis. The red arrow highlights the causal association between the body shape (exposure) and breast cancer (outcome) assessed.

Figure 2

Mendelian randomization estimates between body shape phenotype PC1, PC2 and PC3 and overall and sub-types breast cancer risk. CI, confidence interval; OR, odds ratio. We report random-effect IVW results due to heterogeneity across the genetic instruments.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFiguresandTables.pdf](#)
- [SupplementaryFile1.pdf](#)