Article

Body Shape Phenotypes and Breast Cancer Risk: A Mendelian Randomization Analysis

Laia Peruchet-Noray [1,2], Niki Dimou [1], Anja M. Sedlmeier [3], Béatrice Fervers [4], Isabelle Romieu [5], Vivian Viallon [1], Pietro Ferrari [1], Marc J. Gunter [1], Robert Carreras-Torres [6] and Heinz Freisling [1,*]

[1] International Agency for Research on Cancer (IARC/WHO), Nutrition and Metabolism Branch, CEDEX 08, 69372 Lyon, France
[2] Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, 08007 Barcelona, Spain
[3] Department of Epidemiology and Preventive Medicine, University of Regensburg, Franz-Josef-Strauss-Allee 11, 93053 Regensburg, Germany
[4] Département Prévention Cancer Environnement, Centre Léon Bérard, CEDEX 08, 69373 Lyon, France
[5] National Institute of Public Health, Cuernavaca 62100, Morelos, Mexico
[6] Digestive Diseases and Microbiota Group, Institut d'Investigació Biomèdica de Girona Dr. Josep Trueta (IDIBGI), 17190 Salt, Spain
* Correspondence: freislingh@iarc.who.int

Simple Summary: Investigating body shapes may offer novel insights into the role of adiposity and body size in breast cancer development. We derived three distinct body shapes from combinations of height, weight, body mass index (BMI), waist and hip circumference, and waist-to-hip ratio (WHR) by a principal component analysis. We then used genetic variants linked to these body shapes and investigated associations with breast cancer risk. A body shape characterizing general adiposity was associated with a lower breast cancer risk, however this is likely because the genetic variants predicting this body shape reflect body fatness during childhood and adolescence rather than during later adulthood. A body shape characterizing tall women with a low WHR was weakly associated with a higher breast cancer risk. A body shape characterizing tall women with a large WHR was not associated with breast cancer risk. These findings can spur wider application in cancer research and possibly risk stratification.

Abstract: Observational and genetic studies have linked different anthropometric traits to breast cancer (BC) risk, with inconsistent results. We aimed to investigate the association between body shape defined by a principal component (PC) analysis of anthropometric traits (body mass index [BMI], height, weight, waist-to-hip ratio [WHR], and waist and hip circumference) and overall BC risk and by tumor sub-type (luminal A, luminal B, HER2+, triple negative, and luminal B/HER2 negative). We performed two-sample Mendelian randomization analyses to assess the association between 188 genetic variants robustly linked to the first three PCs and BC (133,384 cases/113,789 controls from the Breast Cancer Association Consortium (BCAC)). PC1 (general adiposity) was inversely associated with overall BC risk (0.89 per 1 SD [95% CI: 0.81–0.98]; p-value = 0.016). PC2 (tall women with low WHR) was weakly positively associated with overall BC risk (1.05 [95% CI: 0.98–1.12]; p-value = 0.135), but with a confidence interval including the null. PC3 (tall women with large WHR) was not associated with overall BC risk. Some of these associations differed by BC sub-types. For instance, PC2 was positively associated with a risk of luminal A BC sub-type (1.09 [95% CI: 1.01–1.18]; p-value = 0.02). To clarify the inverse association of PC1 with breast cancer risk, future studies should examine independent risk associations of this body shape during childhood/adolescence and adulthood.

Keywords: body shape; breast cancer; height; Mendelian randomization; obesity

1. Introduction

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]. Mammography screening reduces breast cancer mortality effectively by detecting breast cancer at an early and curable stage [2]. In addition, mutations in high penetrance genes (e.g., BRCA1/2) can help to identify individuals at a higher risk of developing breast cancer [3]. However, the breast cancer screening participation rate varies largely across countries and can be as low as 16% [4] and the known genetic predisposing genes account for 5–10% of the total breast cancer burden [5]. The identification of breast cancer risk factors is crucial to reduce incidence and associated mortality, and to risk-stratify the population to optimize screening participation [2].

Established risk factors for female breast cancer include age, family history, alcohol consumption, ionizing 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 [6].

There is inconsistent evidence from prospective observational and genetic studies linking body fatness, mostly assessed by body mass index (BMI), to breast cancer incidence [6,7]. Mostly from observational studies, BMI has been reported to be either positively or inversely associated with breast cancer depending on menopausal status and the molecular breast cancer sub-type [6]. However, genetically predicted BMI is inversely associated with breast cancer risk, independently of the individual's menopausal status [7]. Additionally, other single anthropometric traits have also been assessed as risk factors for breast cancer risk. For instance, 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]. Finally, height has been positively associated with a risk of overall and hormone-receptor-positive breast cancer [6,8].

While a correlation between different anthropometric variables exists, individuals with similar values in one trait (e.g., BMI) may differ considerably 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. In 2016, 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. We have previously shown that these body shapes were differentially associated with breast cancer risk in an observational analysis [10]. For example, a body shape characterizing overall adiposity was positively associated with post-menopausal breast cancer risk, but not associated with pre-menopausal breast cancer risk, while a body shape contrasting tall women with large WHR and short women with low WHR was not associated with either pre- or post-menopausal breast cancer [10]. However, considerable uncertainty remains about the potential 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 the causal inference in exposure–outcome associations in the absence of pleiotropic effects (i.e., a single genetic locus influences multiple phenotypes). 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 [11].

In this study, a two-sample MR analysis was conducted to investigate the potential causal associations between three body shape phenotypes, as derived in Ried et al. [9], and the risk of breast cancer using summary level data including 133,384 breast cancer cases and 113,789 controls from the Breast Cancer Association Consortium (BCAC) [12]. We also investigated associations by hormonal receptor status breast cancer sub-types.

2. Materials and Methods

2.1. 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]. Each PC represents a specific composition of the six anthropometric traits and thus captures a specific aspect of body shape. The first three PCs explained 96.7% of the variation of these six anthropometric traits [9]. PC1 explained on average 64.4% of the variation in all traits and had equally high positive loadings on all traits, except for height, characterizing individuals with general adiposity. PC2 explained 18.5% of the variation and 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 explained 13.8% of the variation and 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 S1.

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, the results obtained are not reported in this manuscript due to their imprecision and the impossibility of building solid conclusions.

The genome-wide association study (GWAS) to identify genetic variants robustly associated with the body shape phenotypes included more than 170,000 individuals of European descent from 65 studies from the GIANT Consortium. All study participants gave written informed consent and ethic committees approved all studies. The results included 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) (Supplementary File S1).

2.2. Breast Cancer Risk Data

Genetic variant summary statistics for associations with the risk of overall breast cancer and 5 sub-types 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) [12]. The overall breast cancer analyses were based on mostly middle-aged women of European ancestry from 82 BCAC studies from over 20 countries including controls, cases of invasive breast cancer, cases of carcinoma in situ and cases of unknown invasiveness. In the sub-type analyses, 81 BCAC studies were included, restricted to controls and cases of invasive breast cancer. The study comprised data on luminal A (7325 cases), luminal B (1682 cases), HER2+ (718 cases), triple negative (2006 cases) and luminal B/HER2 (1779 cases) invasive breast cancer sub-types, and 20,815 controls. More detailed information on sample size and tumor heterogeneity is provided in Supplementary Table S2. All participating studies were approved by their appropriate ethics or institutional review board and all participants provided informed consent.

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.

2.3. 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 (Figure 1).
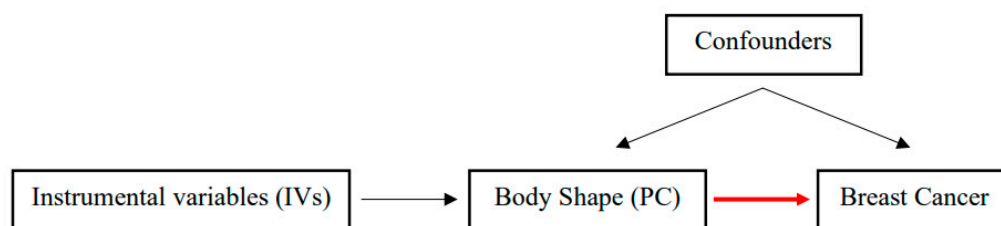
Figure 1. Causal directed acyclic graph (DAG) of Mendelian Randomization analysis. The red arrow highlights the causal association between the body shape (exposure) and breast cancer (outcome) assessed. PC, principal component.

The MR analysis was performed and reported in line with the STROBE-MR guidelines [13,14]. The STROBE-MR guidelines checklist can be found in Supplementary File S2. 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 [15,16]. 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].

We used the fixed-effects inverse-variance weighted (IVW) method as the primary approach [17]. This analysis can be understood as a meta-analysis of the Wald estimates for each genetic variant effects which are calculated as the ratio of the genetic variant-breast cancer over the genetic variant-body shape association [15,18]. The fixed-effects IVW model assumes absence of horizontal pleiotropy and will return an unbiased estimate provided that all variants are valid instrumental variables. We calculated the MR-derived odds ratio (OR) of breast cancer risk for a one standard deviation (SD) increase in genetically predicted body shape phenotype.

2.4. 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 [19], further analyses were performed to assess any potential violation of both assumptions.

First, we used PhenoScanner (available at http://www.phenoscanner.medschl.cam.ac.uk, accessed on 12 June 2021) and the GWAS catalog (available at https://www.ebi.ac.uk/gwas/, accessed on 15 June 2021) to investigate whether our selected genetic variants were associated with other traits in previous GWAS [19–21].

Second, we used Cochran's Q test to evaluate the heterogeneity of causal effects for each variant, with a *p*-value < 0.05 indicating statistical significance [22]. In the case of heterogeneity, likely indicating pleiotropy [23], a random-effects IVW MR analysis was used [24]. This method relaxes the third assumption as the total pleiotropic effect of a single genetic variant no longer needs to be null but assumes a zero mean between all the genetic variants (i.e., balanced pleiotropy) [24]. In addition, we implemented MR-Egger regression [25] and the weighted median approach [26]. The MR-Egger methodology provides an unbiased causal effect estimate even if the third MR assumption is violated and all the variants are invalid instruments. However, the Instrument Strength Independent of Direct Effect (InSIDE) assumption needs to be held. This additional assumption is based on the independency between the horizontal pleiotropic effects and the variants-exposure effects. MR-Egger also 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 the potential violation of the NO Measurement Error (NOME) assumption and suggests the unreliability of MR-Egger inferences at values below 90% [25,27]. The weighted median estimator allows for 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 of the three basic MR assumptions) [23,26]. Both tests provide valid MR estimates in the presence

of overall directional pleiotropy but suffer from reduced power [25,26]. We also used the MR Pleiotropy RESidual Sum and Outlier (MR-PRESSO) method to identify and remove any outlying variants, applying a random-effects IVW model [28]. This method regresses genetic variant-outcome on genetic variant-exposure and uses the square of residuals to identify outliers. Finally, leave-one-genetic-variant-out analyses were used to assess whether any association was driven by specific genetic variants.

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

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 Figure S1).

All analyses were conducted with the statistical software R 4.0.4 and RStudio 1.4.1106 using the MendelianRandomization 0.6.0 and MRPRESSO 1.0 R packages [28,29].

3. Results

3.1. 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) (Figure 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 S3). Nevertheless, the 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 Figure S2A). 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 Figure S3). This genetic variant along with the rs1582874 variant (associated with height and other traits) (Supplementary File S3) 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 (ρ distortion = 0.13) (Supplementary Table S4).

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) (Figure 2), but the confidence interval included the null. We found the presence of heterogeneity across genetic variants (*p*-value < 0.001), but directional pleiotropy was not detected (Supplementary Table S3). Accordingly, the dispersion of the estimates provided by the genetic variants in the funnel plot showed balanced pleiotropy (Supplementary Figure S4). However, MR PRESSO indicated that the presence of five outlier genetic variants was potentially biasing the initial risk estimate (ρ distortion = 0.013) (Supplementary Table S4). The weighted median and IVW random MR approaches without outlier genetic variants estimations were attenuated toward the null (Supplementary Figure S2B).

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) (Figure 2). We observed evidence of heterogeneity (*p*-value < 0.001), but without the presence of overall directional pleiotropy (Supplementary Table S3). The sensitivity analysis supported a null association (Supplementary Figure S2C).

3.2. Body Shape Phenotypes and Risk of Breast Cancer Sub-Types

PC1 was inversely associated with a 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) (Figure 2). Nevertheless, all of the associations showed heterogeneity and overall directional pleiotropy (Supplementary Table S3). Sensitivity analyses accounting for unbalanced pleiotropy and outliers supported the inverse associations of the main analysis (Supplementary Figure S2A). MR-PRESSO classified genetic variants as outliers in all the sub-types but only in luminal A sub-type identified biased association estimates

(ρ distortion = 0.02) (Supplementary Table S4).Comparable to the overall breast cancer risk analysis, PC2 risk estimates included the null, except for luminal A (IVW$_{\text{random-effects}}$ OR = 1.09 [95% CI: 1.01–1.18]; *p*-value = 0.02) (Figure 2) with the presence of balanced pleiotropy (heterogeneity *p*-value luminal A < 0.001; Egger intercept *p*-value luminal A = 0.256) (Supplementary Table S3).
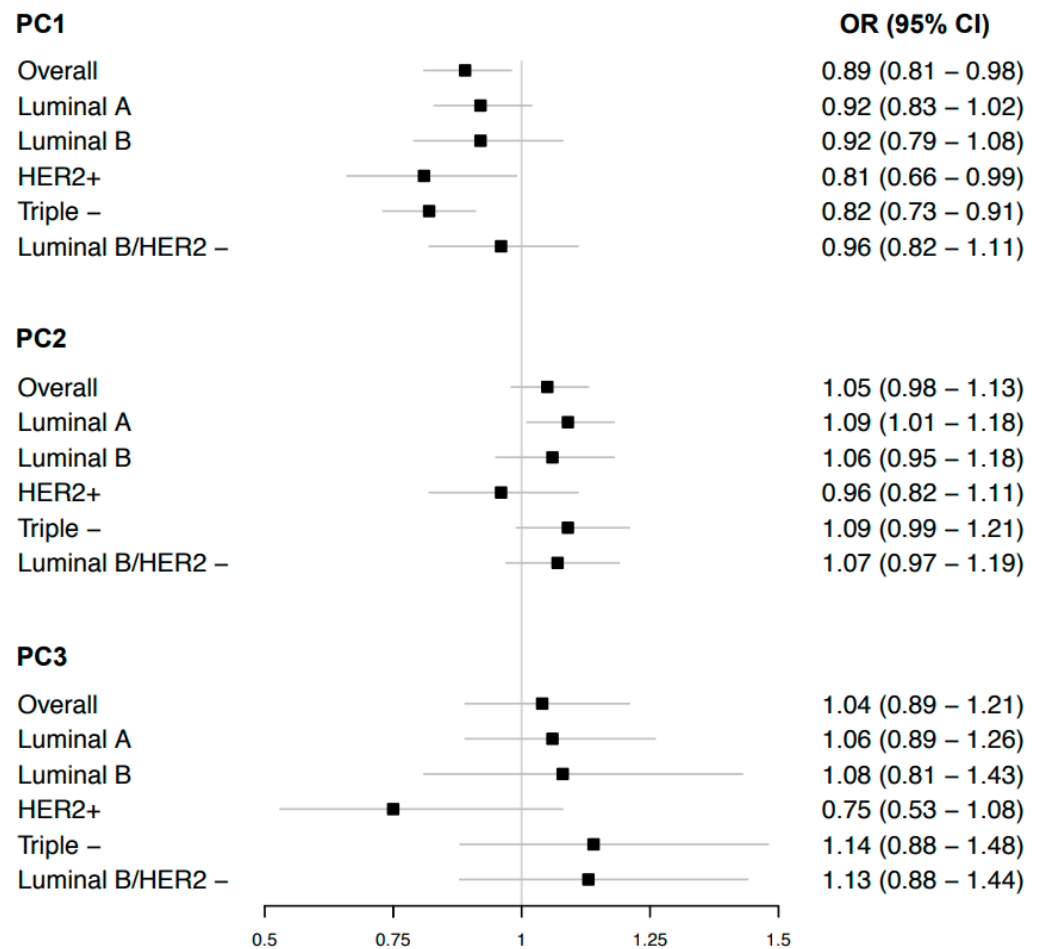
PC1   **OR (95% CI)**

Overall	0.89 (0.81 – 0.98)
Luminal A	0.92 (0.83 – 1.02)
Luminal B	0.92 (0.79 – 1.08)
HER2+	0.81 (0.66 – 0.99)
Triple −	0.82 (0.73 – 0.91)
Luminal B/HER2 −	0.96 (0.82 – 1.11)

PC2

Overall	1.05 (0.98 – 1.13)
Luminal A	1.09 (1.01 – 1.18)
Luminal B	1.06 (0.95 – 1.18)
HER2+	0.96 (0.82 – 1.11)
Triple −	1.09 (0.99 – 1.21)
Luminal B/HER2 −	1.07 (0.97 – 1.19)

PC3

Overall	1.04 (0.89 – 1.21)
Luminal A	1.06 (0.89 – 1.26)
Luminal B	1.08 (0.81 – 1.43)
HER2+	0.75 (0.53 – 1.08)
Triple −	1.14 (0.88 – 1.48)
Luminal B/HER2 −	1.13 (0.88 – 1.44)

0.5   0.75   1   1.25   1.5

Figure 2. Mendelian randomization estimates between body shape phenotype PC1, PC2, and PC3 and overall breast cancer risk and sub-types. CI, confidence interval; OR, odds ratio; PC, principal component. We report random-effect IVW results due to heterogeneity across the genetic instruments.

PC3 was suggestively inversely associated with the risk of the HER2+ sub-type breast cancer (IVW$_{\text{random-effects}}$ OR = 0.75 [95% CI: 0.53–1.08]; *p*-value = 0.122) (Figure 2), in contrast with other sub-types; triple negative and luminal B/HER2 negative sub-types of breast cancer, although risk estimates were imprecise (OR$_{\text{triple negative}}$ = 1.14 [95% CI 0.88–1.48]; OR$_{\text{luminal B/HER2 negative}}$ = 1.13 [95% CI 0.88–1.44]) (Figure 2). The sensitivity analysis showed comparable results except for the Egger test which presented unreliable results (I$^2_{\text{GX}}$~0%) similar to the overall breast cancer analysis (Supplementary Figure S2C).

4. Discussion

We performed a large-scale MR analysis to estimate potential causal associations between three distinct body shape phenotypes and the risk of overall breast cancer and five breast cancer tumor sub-types. A body shape capturing general adiposity (PC1) was inversely associated with the risk of overall breast cancer. The inverse association was stronger for ER− subtypes as compared to ER+ sub-types. A body shape contrasting tall women with a low WHR to short women with a large WHR (PC2) was weakly positively associated with overall breast cancer risk and most sub-types. There was no evidence of an

association between PC3, a body shape characterizing tall women with a large WHR, and the 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).

General adiposity has been traditionally assessed using BMI, and in the absence of MR studies using PC1, BMI is an indicator that is closest to the PC1 body shape in our study. Previous MR analyses showed an inverse association between BMI and pre- and post-menopausal breast cancer risk [7,30–33], which is congruent with our findings for PC1. Specifically, a 2019 MR analysis found that BMI was inversely associated with the risk of overall breast cancer, and ER+ (luminal A, luminal B and luminal B/HER2 negative) and ER− (HER2+ and triple negative) sub-types, using a similar dataset—although with less cancer cases—to the BCAC study [31]. However, the associations observed for the ER+ and ER− sub-groups slightly differed from our results. While our results suggest a 10% and 20% reduced risk, in ER+ and ER− related sub-types of 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 [31].

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 specific 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]. In comparison to other research in the field using anthropometric traits, such as BMI, waist circumference, or height individually, we captured body size and adiposity holistically by combining information from six anthropometric traits. The derived body shapes represent anthropometric information that is by construction orthogonal, i.e., mutually adjusted for the six traits and risk associations (here: with breast cancer) are likely less biased due to residual confounding between traits in standard multivariable regression e.g., adjusting WHR for height.

In contrast to the MR results, observational studies have shown that a high BMI throughout adulthood is convincingly associated with a higher risk of post-menopausal breast cancer [6,34]. Observational studies also showed that a high BMI prior to menopause is associated with a lower risk of pre-menopausal breast cancer [6]. Furthermore, a high BMI in young adulthood (18–30 years of age) is associated with a lower risk of pre- and post-menopausal breast cancer [6]. A 2020 MR study provided some potential insights into the apparently divergent findings between observational and MR approaches regarding post-menopausal breast cancer risk [35]. This MR study found that elevated childhood BMI (age 10 years) was associated with a decreased risk of pre- and post-menopausal breast cancer combined, whereas adult BMI (mean age 56.5 years) was suggestive of a positive association (OR = 1.08 [95% CI: 0.93–1.27]) after accounting for childhood BMI [35]. Body fatness in childhood and adolescence (indicated by BMI) has also been linked to a decreased risk of breast cancer among pre- and post-menopausal women in observational studies [36–38], which suggests a long-term effect of body fatness at a young age on breast cancer risk later in life. Mechanistic evidence from in vivo experiments in female rats supports this notion whereby adipose-tissue-derived estrogen in overweight adolescents may induce early breast differentiation and render the breast tissue less susceptible to tumorigenesis [39]. Thus, the genetically predicted PC1 exposure used in our MR analysis may reflect body fatness during childhood and adolescence rather than during later adulthood explaining the inverse association with breast cancer risk.

In post-menopausal women, adipose tissue converts adrenal androgens into estrogen by the aromatase enzyme [40]. The presence of estrogen, alone and with progestin, is directly associated with an increased post-menopausal breast cancer risk [40]. Furthermore, in vivo and in vitro models have shown that macrophages that are stimulated by an obese milieu can increase estrogen signaling in breast cancer cells through the upregulation of ERα expression [41]. Indirectly, higher levels of estrogen may upregulate leptin production, which may also be relevant for breast cancer development [40]. For example, leptin mediates

immune suppression facilitating cancer formation and progression [40]. However, in pre-menopausal women, high vs low BMI is generally associated with longer anovulatory cycles resulting in lower levels of estrogen and progestin [42], thereby potentially providing a barrier to cancer formation. Taken together, the time of onset of excess body fatness and its modulating effects on circulating levels of estrogen and progestin may partly explain the divergent associations of body fatness with breast cancer risk. 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 [30,33].

Our findings concerning PC2 (tall with low WHR) are in line with the observational analysis of a previous publication of our group, where PC2 was positively associated with pre- and post-menopausal breast cancer with risk increments of 8–10% per 1 SD increase [10]. Positive associations between height and breast cancer risk using both observational and instrumental approaches have been reported with risk increases of 7–20% per ~10 cm increase [8,43]. Our observational and MR results for PC2 (OR = 1.05 [95% CI: 0.98–0.13]) were slightly weaker than those reported in the literature for height, which can potentially be explained by the contribution of WHR into PC2. However, no clear role of WHR with cancer risk has been observed [32,44].

Regarding breast cancer sub-types, only PC2 showed a positive association with the luminal A (ER+/PR+) tumor subtype. The results of other sub-types showed a confidence interval including the null. 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 S3) when compared to Zhang et al. (168 variants) [8].

Mechanistically, breast tissue in taller women is thought to be exposed to higher levels of insulin, growth hormone and IGF-I, and thus may have undergone more cell divisions contributing to a greater potential for error during DNA replication [6]. Animal models clearly support a key role of growth hormone and its downstream signaling, including the IGF pathway, in breast cancer development [45].

For body shape PC3, indicating body height combined with small HC and large WHR, we found little evidence for an association with breast cancer risk and sub-types, congruent with observational results from our group [10]. However, HER2+ sub-type showed an inverse association trend, in contrast with other breast cancer sub-types. HER2+ sub-type showed inverse associations for all tested PCs, being only statistically significant for PC1. In the 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 [46,47].

In this study we employed different MR approaches: IVW-based, 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 and detects directional pleiotropy. The median-based method is more robust to overall pleiotropy as it allows up to 50% of invalid instruments (i.e., violating the second or third assumptions of MR) [48].

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 potential presence of unknown confounding factors. Other limitations include that our study may be slightly biased due to uncontrolled confounding from family effects, such as assortative mating, dynastic effects, and population structure [49]. We minimized this bias from population stratification by performing a two-sample MR analysis in samples composed of Caucasian individuals only. However, this may cause our findings to not be

generalizable to other ethnic groups. Referring to the body shapes, it is difficult to interpret the effect of PCs on breast cancer as it is not evident to what 1 SD corresponds with in terms of anthropometry. Finally, it is important to highlight that we were not able to differentiate between pre- and post-menopausal breast cancer with the available data.

5. Conclusions

This MR analysis provides evidence of an inverse association between a body shape that reflects general adiposity and the risk of overall breast cancer among women of European decent. The strength of association differed across sub-types of breast cancer suggesting a stronger inverse association for ER− sub-types as compared to ER+ sub-types. The most likely explanation for these inverse associations with overall breast cancer and sub-types is that the genetic instruments for this body shape indicate general adiposity during childhood/adolescence rather than during later adulthood. To test this assertion in future studies, genetic instruments for body shape during both childhood/adolescence and adulthood (age 30+ years) should be identified and their independent risk associations with breast cancer compared in a multivariable MR framework.

Furthermore, we found a weak positive association between a body shape characterizing tall women with a low WHR and overall breast cancer risk and most sub-types. This is in line with previous evidence on height and breast cancer. Our findings add to this evidence base by ruling out confounding by WHR.

A body shape characterizing tall women with a large WHR was not associated with breast cancer risk.

The current study provides complementary evidence to observational studies to interrogate multiple anthropometric measurements in relation to breast cancer risk.

Conflicts of Interest: The authors declare no conflict of interest.

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.; Znaor, A.; Soerjomataram, I.; Bray, F. *Global Cancer Observatory: Cancer Today*; International Agency for Research on Cancer: Lyon, France, 2020. Available online: https://gco.iarc.fr/today (accessed on 29 September 2021).
2. Løberg, M.; Lousdal, M.L.; Bretthauer, M.; Kalager, M. Benefits and harms of mammography screening. *Breast Cancer Res.* **2015**, *17*, 63. [CrossRef] [PubMed]
3. Thorat, M.A.; Balasubramanian, R. Breast cancer prevention in high-risk women. *Best Pract. Res. Clin. Obstet. Gynaecol.* **2020**, *65*, 18–31. [CrossRef] [PubMed]
4. Mandrik, O.; Tolma, E.; Zielonke, N.; Meheus, F.; Ordóñez-Reyes, C.; Severens, J.; Murillo, R. Systematic reviews as a "lens of evidence": Determinants of participation in breast cancer screening. *J. Med. Screen.* **2020**, *28*, 70–79. [CrossRef] [PubMed]
5. Oldenburg, R.; Meijers-Heijboer, H.; Cornelisse, C.; Devilee, P. Genetic susceptibility for breast cancer: How many more genes to be found? *Crit. Rev. Oncol.* **2007**, *63*, 125–149. [CrossRef] [PubMed]
6. World Cancer Research Fund; American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Diet, Nutrition, Physical Activity and Breast Cancer. Available online: dietandcancerreport.org (accessed on 14 September 2021).
7. Larsson, S.C.; Burgess, S. Causal role of high body mass index in multiple chronic diseases: A systematic review and meta-analysis of Mendelian randomization studies. *BMC Med.* **2021**, *19*, 320. [CrossRef]
8. Zhang, B.; Shu, X.-O.; Delahanty, R.J.; Zeng, C.; Michailidou, K.; Bolla, M.K.; Wang, Q.; Dennis, J.; Wen, W.; Long, J.; et al. Height and breast cancer risk: Evidence from prospective studies and mendelian randomization. *Gynecol. Oncol.* **2015**, *107*, djv219. [CrossRef]
9. Ried, J.S.; Jeff, M.J.; Chu, A.Y.; Bragg-Gresham, J.L.; van Dongen, J.; Huffman, J.E.; Ahluwalia, T.S.; Cadby, G.; Eklund, N.; Eriksson, J.; et al. A principal component meta-analysis on multiple anthropometric traits identifies novel loci for body shape. *Nat. Commun.* **2016**, *7*, 13357. [CrossRef]
10. Sedlmeier, A.M.; Viallon, V.; Ferrari, P.; Peruchet-Noray, L.; Fontvieille, E.; Amadou, A.; Khoei, N.S.; Weber, A.; Baurecht, H.; Heath, A.K.; et al. Body shape phenotypes of multiple anthropometric traits and cancer risk: A multi-national cohort study. *Br. J. Cancer* **2022**. [CrossRef]
11. Smith, G.D.; Ebrahim, S. "Mendelian randomization": Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int. J. Epidemiol.* **2003**, *32*, 1–22. [CrossRef]
12. Zhang, H.; Ahearn, T.U.; Lecarpentier, J.; Barnes, D.; Beesley, J.; Qi, G.; Jiang, X.; O'Mara, T.A.; Zhao, N.; Bolla, M.K.; et al. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **2020**, *52*, 572–581. [CrossRef]
13. Skrivankova, V.W.; Richmond, R.C.; Woolf, B.A.R.; Davies, N.M.; Swanson, S.A.; VanderWeele, T.J.; Timpson, N.J.; Higgins, J.P.T.; Dimou, N.; Langenberg, C.; et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomisation (STROBE-MR): Explanation and elaboration. *BMJ* **2021**, *375*, n2233. [CrossRef] [PubMed]

14. Skrivankova, V.W.; Richmond, R.C.; Woolf, B.A.R.; Yarmolinsky, J.; Davies, N.M.; Swanson, S.A.; VanderWeele, T.J.; Higgins, J.P.T.; Timpson, N.J.; Dimou, N.; et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: The STROBE-MR Statement. *JAMA* **2021**, *326*, 1614. [CrossRef] [PubMed]

15. Lawlor, D.A.; Harbord, R.M.; Sterne, J.A.C.; Timpson, N.; Smith, G.D. Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat. Med.* **2008**, *27*, 1133–1163. [CrossRef] [PubMed]

16. Didelez, V.; Sheehan, N. Mendelian randomization as an instrumental variable approach to causal inference. *Stat. Methods Med. Res.* **2007**, *16*, 309–330. [CrossRef] [PubMed]

17. Burgess, S.; Scott, R.A.; Timpson, N.J.; Smith, G.D.; Thompson, S.G.; EPIC-InterAct Consortium. Using published data in Mendelian randomization: A blueprint for efficient identification of causal risk factors. *Eur. J. Epidemiol.* **2015**, *30*, 543–552. [CrossRef]

18. Burgess, S.; Butterworth, A.; Thompson, S.G. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet. Epidemiol.* **2013**, *37*, 658–665. [CrossRef]

19. Dimou, N.L.; Tsilidis, K.K. A primer in mendelian randomization methodology with a focus on utilizing published summary association data. *Genet. Epidemiol. Methods Protoc.* **2018**, *1793*, 211–230. [CrossRef]

20. MacArthur, J.; Bowler-Barnett, E.; Cerezo, M.; Gil, L.; Hall, P.; Hastings, E.; Junkins, H.; McMahon, A.; Milano, A.; Morales, J.; et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **2016**, *45*, D896–D901. [CrossRef]

21. Staley, J.R.; Blackshaw, J.; Kamat, M.A.; Ellis, S.; Surendran, P.; Sun, B.B.; Paul, D.S.; Freitag, D.; Burgess, S.; Danesh, J.; et al. PhenoScanner: A database of human genotype–phenotype associations. *Bioinformatics* **2016**, *32*, 3207–3209. [CrossRef]

22. Higgins, J.P.T.; Thompson, S.G.; Deeks, J.J.; Altman, D.G. Measuring inconsistency in meta-analyses. *BMJ* **2003**, *327*, 557–560. [CrossRef]

23. Pagoni, P.; Dimou, N.L.; Murphy, N.; Stergiakouli, E. Using Mendelian randomisation to assess causality in observational studies. *Evid. Based Ment. Health* **2019**, *22*, 67–71. [CrossRef] [PubMed]

24. 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*, 1783–1802. [CrossRef] [PubMed]

25. Bowden, J.; Davey Smith, G.; Burgess, S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int. J. Epidemiol.* **2015**, *44*, 512–525. [CrossRef] [PubMed]

26. Bowden, J.; Smith, G.D.; Haycock, P.C.; Burgess, S. Consistent Estimation in Mendelian Randomization with Some Invalid Instruments Using a Weighted Median Estimator. *Genet. Epidemiol.* **2016**, *40*, 304–314. [CrossRef]

27. Bowden, J.; Del Greco, M.F.; Minelli, C.; Smith, G.D.; Sheehan, N.A.; Thompson, J.R. Assessing the suitability of summary data for two-sample Mendelian randomization analyses using MR-Egger regression: The role of the I^2 statistic. *Leuk. Res.* **2016**, *45*, 1961–1974. [CrossRef]

28. Verbanck, M.; Chen, C.-Y.; 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*, 693–698. [CrossRef]

29. Yavorska, O.O.; Burgess, S. MendelianRandomization: An R package for performing Mendelian randomization analyses using summarized data. *Leuk. Res.* **2017**, *46*, 1734–1739. [CrossRef]

30. Guo, Y.; Andersen, S.W.; Shu, X.-O.; Michailidou, K.; Bolla, M.K.; Wang, Q.; Garcia-Closas, M.; Milne, R.L.; Schmidt, M.K.; Chang-Claude, J.; 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*, e1002105. [CrossRef]

31. Ooi, B.N.S.; Loh, H.; Ho, P.J.; Milne, R.L.; Giles, G.; Gao, C.; Kraft, P.; John, E.M.; Swerdlow, A.; Brenner, H.; et al. The genetic interplay between body mass index, breast size and breast cancer risk: A Mendelian randomization analysis. *Leuk. Res.* **2019**, *48*, 781–794. [CrossRef]

32. Gao, C.; Patel, C.J.; Michailidou, K.; Peters, U.; Gong, J.; Schildkraut, J.; Schumacher, F.R.; Zheng, W.; Boffetta, P.; Stucker, I.; et al. Mendelian randomization study of adiposity-related traits and risk of breast, ovarian, prostate, lung and colorectal cancer. *Leuk. Res.* **2016**, *45*, 896–908. [CrossRef]

33. Shu, X.; Wu, L.; Khankari, N.K.; Shu, X.-O.; Wang, T.J.; Michailidou, K.; Bolla, M.K.; Wang, Q.; Dennis, J.; Milne, R.L.; et al. Associations of obesity and circulating insulin and glucose with breast cancer risk: A Mendelian randomization analysis. *Leuk. Res.* **2018**, *48*, 795–806. [CrossRef] [PubMed]

34. Recalde, M.; Davila-Batista, V.; Díaz, Y.; Leitzmann, M.; Romieu, I.; Freisling, H.; Duarte-Salles, T. Body mass index and waist circumference in relation to the risk of 26 types of cancer: A prospective cohort study of 3.5 million adults in Spain. *BMC Med.* **2021**, *19*, 10. [CrossRef] [PubMed]

35. Richardson, T.G.; Sanderson, E.; Elsworth, B.; Tilling, K.; Smith, G.D. Use of genetic variation to separate the effects of early and later life adiposity on disease risk: Mendelian randomisation study. *BMJ* **2020**, *369*, m1203. [CrossRef] [PubMed]

36. Baer, H.J.; Tworoger, S.; Hankinson, S.E.; Willett, W.C. Body fatness at young ages and risk of breast cancer throughout life. *Am. J. Epidemiol.* **2010**, *171*, 1183–1194. [CrossRef]

37. Yang, T.O.; Cairns, B.J.; Pirie, K.; Green, J.; Beral, V.; Floud, S.; Reeves, G.K. Body size in early life and the risk of postmenopausal breast cancer. *BMC Cancer* **2022**, *22*, 232. [CrossRef]

38. Baer, H.J.; Colditz, G.A.; Rosner, B.; Michels, K.B.; Rich-Edwards, J.W.; Hunter, D.J.; Willett, W.C. Body fatness during childhood and adolescence and incidence of breast cancer in premenopausal women: A prospective cohort study. *Breast Cancer Res.* **2005**, *7*, R314. [CrossRef]

39. Grubbs, C.J.; Farnell, D.R.; Hill, D.L.; McDonough, K.C. Chemoprevention of N-nitroso-N-methylurea-induced mammary cancers by pretreatment with 17β-estradiol and progesterone23. *Gynecol. Oncol.* **1985**, *74*, 927–931. [CrossRef]

40. Picon-Ruiz, M.; Morata-Tarifa, C.; Valle-Goffin, J.J.; Friedman, E.R.; Slingerland, J.M. Obesity and adverse breast cancer risk and outcome: Mechanistic insights and strategies for intervention. *CA Cancer J. Clin.* **2017**, *67*, 378–397. [CrossRef]

41. Blank, D.N.; Hermano, E.; Sonnenblick, A.; Maimon, O.; Rubinstein, A.M.; Drai, E.; Maly, B.; Vlodavsky, I.; Popovtzer, A.; Peretz, T.; et al. Macrophages upregulate estrogen receptor expression in the model of obesity-associated breast carcinoma. *Cells* **2022**, *11*, 2844. [CrossRef]

42. Fichman, V.; Costa, R.D.S.S.D.; Miglioli, T.C.; Marinheiro, L.P.F. Association of obesity and anovulatory infertility. *Einstein* **2020**, *18*, eAO5150. [CrossRef]

43. Lai, F.Y.; Nath, M.; Hamby, S.E.; Thompson, J.R.; Nelson, C.; Samani, N.J. Adult height and risk of 50 diseases: A combined epidemiological and genetic analysis. *BMC Med.* **2018**, *16*, 187. [CrossRef] [PubMed]

44. Chen, G.-C.; Chen, S.-J.; Zhang, R.; Hidayat, K.; Qin, J.-B.; Zhang, Y.-S.; Qin, L.-Q. Central obesity and risks of pre- and postmenopausal breast cancer: A dose-response meta-analysis of prospective studies. *Obes. Rev.* **2016**, *17*, 1167–1177. [CrossRef] [PubMed]

45. Subramani, R.; Nandy, S.B.; Pedroza, D.A.; Lakshmanaswamy, R. Role of growth hormone in breast cancer. *Endocrinology* **2017**, *158*, 1543–1555. [CrossRef] [PubMed]

46. Ménard, S.; Balsari, A.; Casalini, P.; Tagliabue, E.; Campiglio, M.; Bufalino, R.; Cascinelli, N. HER-2-positive breast carcinomas as a particular subset with peculiar clinical behaviors. *Clin. Cancer Res.* **2002**, *8*, 520–525.

47. 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*, 1032–1034. [CrossRef]

48. Pingault, J.-B.; O'Reilly, P.F.; Schoeler, T.; Ploubidis, G.B.; Rijsdijk, F.; Dudbridge, F. Using genetic data to strengthen causal inference in observational research. *Nat. Rev. Genet.* **2018**, *19*, 566–580. [CrossRef]

49. Brumpton, B.; Sanderson, E.; Heilbron, K.; Hartwig, F.P.; Harrison, S.; Vie, G.; Cho, Y.; Howe, L.D.; Hughes, A.; Boomsma, D.I.; et al. Avoiding dynastic, assortative mating, and population stratification biases in Mendelian randomization through within-family analyses. *Nat. Commun.* **2020**, *11*, 3519. [CrossRef]