Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Body composition and the breast cancer risk associated with dense breasts: a prospective cohort study of 106,813 predominantly premenopausal Korean women

Yun Tae Kim
Kangbuk Samsung Hospital
Eun Young Kim
Kangbuk Samsung Hospital
Chan Heun Park
Kangbuk Samsung Hospital
Yong Lai Park
Kangbuk Samsung Hospital
Ji-Sub Yun
Kangbuk Samsung Hospital
Soo Ha Lee
Kangbuk Samsung Hospital
Kwan Ho Lee
lkhdotcom@naver.com

Kangbuk Samsung Hospital

Research Article

Keywords: Mammographic density, Breast density, Body composition, Skeletal muscle, Body fat, Causal mediation analysis, Breast cancer, Cohort study, Effect modification, Menopause

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11323564/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Title

Body composition and the breast cancer risk associated with dense breasts: a prospective cohort study of 106,813 predominantly premenopausal Korean women

Yun Tae Kim, PhD[1,†], Eun Young Kim, MD, PhD[2], Chan Heun Park, MD, PhD[2], Yong Lai Park, MD, PhD[2], Ji-Sub Yun, MD[2], Soo Ha Lee, MD[2], Kwan Ho Lee, MD, PhD[2,*,†]

[1] Division of Biostatistics, Department of Academic Research, Kangbuk Samsung Hospital, Seoul, Republic of Korea

[2] Department of Surgery, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[†] These authors contributed equally to this work.

[*] Corresponding author: Kwan Ho Lee, MD, PhD, Department of Surgery, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, 29 Saemunan-ro, Jongno-gu, Seoul 03181, Republic of Korea. E-mail: lkhdotcom@naver.com

Abstract

Background Mammographic density is a strong breast cancer risk factor, yet higher body mass index is associated with lower density but higher postmenopausal risk. Because body mass index combines fat and lean mass, it is unclear which compartment underlies these opposing associations. It is also unknown whether a dense breast carries less risk when its density reflects

little fat rather than abundant fibroglandular tissue, a question relevant to East Asian women, who often have dense breasts at low body mass index.

Methods We followed 106,813 Korean women in a health-examination cohort by registry linkage (86% premenopausal; mean 6.5 years; 891 invasive breast cancers). Fat and skeletal muscle mass were measured by bioelectrical impedance and modelled jointly against an extremely dense breast (BI-RADS D versus A–C; recorded for 76,551 women, 713 cancers) and breast cancer. We tested whether the density–breast cancer association, adjusted for age, body mass index, waist circumference, and reproductive and lifestyle factors, varied across quartiles of fat mass or skeletal muscle percentage (muscle mass/body weight), or by menopausal status.

Results Per SD, fat mass was strongly associated with lower odds of an extremely dense breast (odds ratio 0.41; 95% CI 0.40–0.42) and skeletal muscle mass weakly with higher odds (1.03; 1.01–1.05). Neither was significantly associated with breast cancer (hazard ratio [HR] 0.98; 0.90–1.05 and 1.07; 1.00–1.16). An extremely dense breast was associated with breast cancer (HR 1.46; 1.23–1.73). If density merely reflected little fat, the association should be weakest in women with the least fat and most muscle; instead, it was strongest in the lowest fat mass quartile (HR 1.74) and highest skeletal muscle percentage quartile (1.67), although not significantly (interaction P = 0.15 and 0.13). The association was stronger after than before menopause (HR 2.93 versus 1.38; interaction HR 2.66, 1.54–4.62; 69 postmenopausal cancers).

Conclusions In this predominantly premenopausal cohort, body fat was inversely associated with density but not with breast cancer, and an extremely dense breast carried no less risk when reflecting low body fat. Although a modest modification cannot be excluded, density appears a valid risk marker across body habitus in East Asian women.

Keywords Mammographic density, Breast density, Body composition, Skeletal muscle, Body fat, Causal mediation analysis, Breast cancer, Cohort study, Effect modification, Menopause

Background

Mammographic density is among the strongest and most prevalent risk factors for breast cancer, with a two- to sixfold gradient in risk across density categories [1, 2]. Its clinical footprint is widening. Since September 2024, mammography reports in the United States must state whether a woman's breasts are dense [3]; density is embedded in contemporary risk prediction models; and change in density is used as a response biomarker in prevention trials, where the reduction achieved with tamoxifen tracks the reduction in subsequent breast cancer [4], so that trials of risk-reducing agents now enrol healthy women with density reduction as the primary endpoint [5]. The first two uses treat density as a marker and require only that it predict risk. The third, and any proposal to lower risk by lowering density, carries a stronger assumption: that density lies on the causal path from its determinants to breast cancer, so that a factor which moves density moves risk in proportion. That assumption has rarely been tested directly.

The clearest reason to doubt it is the body mass index paradox. Body mass index is inversely associated with mammographic density when density is expressed as a proportion of the breast or as a BI-RADS category, largely because adipose tissue enlarges the non-dense compartment against which fibroglandular tissue is seen [6–8], yet positively associated with postmenopausal breast cancer [9] and inversely associated with premenopausal breast cancer [10, 11]. Body mass index, however, is not a single tissue. It aggregates an adipose and a lean compartment that may act on the breast in opposite directions, and a measure that sums them cannot distinguish their contributions. We have previously shown, cross-sectionally in 143,456 women, that skeletal

muscle percentage was the strongest body mass index-independent predictor of an extremely dense breast on BI-RADS classification [12]. Whether the two compartments that body mass index conflates also diverge in their relation to breast cancer itself has not been examined.

This raises two questions. The first is whether the factors that determine density relate to breast cancer in the same way that density does, and how much of their effect actually passes through density. Mediation studies of body size have given different answers depending on the exposure: body size in early life appears to act on breast cancer largely through density [13], whereas adult body shape shows more modest mediation [14]. The second question is whether a dense breast carries the same risk regardless of what made it dense. A breast may look dense because there is little fat to dilute the fibroglandular tissue, or because fibroglandular tissue is abundant, and the two need not carry the same risk. If they do not, the association between density and breast cancer should differ according to the body composition that produced the density. This prediction can be tested, but to our knowledge it has not been.

We addressed both questions in a prospective cohort of 106,813 Korean women with bioelectrical impedance body composition measurement and BI-RADS density classification. We modelled fat mass and skeletal muscle mass simultaneously against both dense breast and incident breast cancer; we used causal mediation analysis to quantify how much of the body-composition effect is transmitted through density; and we tested formally whether the density–cancer hazard is modified by the body composition that produced it and by menopausal status.

Methods

Study population

The Kangbuk Samsung Health Study is a prospective cohort of adults who undergo annual or biennial comprehensive health examinations at the Kangbuk Samsung Hospital Total Healthcare Centers in Seoul and Suwon, Republic of Korea [15]; participants are employees of various companies and local government organisations and their spouses, whose examinations are provided under employer health programmes or purchased individually; the cohort is not a breast-screening cohort, and mammography is one component of a comprehensive examination package. We analysed female participants whose first eligible examination (baseline) took place between 2012 and 2022 and who were followed through linkage to the national cancer registry. Participants consent to linkage of their examination records with the national cancer registry, and linkage is renewed periodically for those who attend and renew that consent, so each woman was followed until the last linkage round that included her: 31 December 2020 for 54,399 women and 31 December 2022 for 52,414 women. Follow-up was administratively censored at that date. Of 191,503 women aged 20 years or older who met initial eligibility, 4,614 had no bioelectrical impedance measurement, 6,673 had a history of cancer, and 14,682 were followed for less than 6 months, an exclusion intended to limit the influence of prevalent but undiagnosed cancer at baseline, leaving 165,534 women.

Mammography is not routinely offered in Korea before 30 years of age, and density was consequently missing for most younger participants. We therefore excluded 25,267 women aged under 30 years, and a further 75 women whose body weight was not recorded and for whom skeletal muscle percentage could not be calculated. Analyses were restricted to complete cases: after excluding 33,379 women with missing data on any covariate of the primary model (see Covariates) or on any of the three bioelectrical impedance measures used here (skeletal muscle mass, fat mass, and height), the full analysis cohort comprised 106,813 women, among whom

891 incident breast cancers occurred over 695,501 person-years (Additional file 1: Table S1). Baseline was the first examination at which eligibility was met; body composition, mammographic density, and all covariates were taken from that same examination, so that exposure, mediator, and covariates refer to one time point. Mediation analyses were restricted to the 76,551 women with a BI-RADS density category recorded at that examination (density subcohort; 713 events). BI-RADS density has been captured electronically at the study centres since the early 2010s and recorded uniformly thereafter, so the absence of a density record reflects not having undergone mammography at the centres, chiefly among younger women, rather than any characteristic of the mammogram itself. Women with and without density data are compared in Additional file 1: Table S3. The study is reported in accordance with the STROBE recommendations [16] (Additional file 2) and, for the mediation analyses, the AGReMA guideline [17] (Additional file 3).

Body composition

Body composition was measured with multifrequency bioelectrical impedance analysis using eight-point tactile electrodes (InBody 720, Biospace, Seoul, Republic of Korea). The device passes alternating currents at several frequencies through each limb and the trunk, estimates segmental body water and fat-free mass from the measured impedance, and derives skeletal muscle mass and fat mass from these estimates; measurements were taken after an overnight fast, and the method has been validated against dual-energy X-ray absorptiometry across the adult age range [18, 19]. The device reports skeletal muscle mass and fat-free mass as separate outputs, skeletal muscle mass being a component of fat-free mass; both are reported in Table 1 and are distinguished throughout.

Three exposures were analysed. Fat mass (kg) and skeletal muscle mass (kg) were used for the joint models. Skeletal muscle percentage was calculated as skeletal muscle mass (kg) divided by body weight (kg) multiplied by 100, following Janssen et al. [20] and our previous report [12], and was analysed per 1%, per 1 SD, and in quartiles (< 35.47%, 35.47 to 37.76%, 37.76 to 40.08%, and > 40.08%). We abbreviate it SMI, the term used by Janssen et al. and in our previous report; it should not be confused with the height-normalised skeletal muscle mass index, which we abbreviate SMMI. Because absolute skeletal muscle mass carries body size, and because bioelectrical impedance predicts muscle mass from an impedance index in which height is embedded, we repeated the muscle-mass analyses with additional adjustment for height and, separately, using the SMMI (skeletal muscle mass divided by height squared, kg/m^2), which applies the height-squared normalisation used by the European and Asian sarcopenia consensus groups for appendicular muscle mass [21, 22].

Mammographic density

Mammograms were interpreted by experienced radiologists using the BI-RADS density classification, fifth edition [23], which grades the proportion of the breast occupied by fibroglandular tissue and therefore behaves as a percent-density-type measure rather than a measure of absolute dense area; each mammogram was read by a single radiologist as part of routine reporting. Because heterogeneously dense breasts (category C) predominate in East Asian women [24], the primary definition of a dense breast was category D versus categories A, B, and C, as in our previous report [12]. The alternative dichotomy (categories C and D versus A and B) left only 8,045 women (10.5%) in the non-dense group, which precluded stable estimation within strata; it was therefore used only in sensitivity analyses.

Outcome ascertainment

Incident invasive breast cancer (ICD-10 code C50) was identified by direct linkage to the Korea Central Cancer Registry, for which registration completeness exceeds 98% [25]. Carcinoma in situ (D05) was not included as an outcome: in situ lesions are detected almost exclusively by screening mammography, whose sensitivity falls as density rises, so that including them would couple ascertainment of the outcome to the mediator. Follow-up ran from the baseline examination to the date of breast cancer diagnosis, death, or the woman's administrative censoring date (the end of her most recent registry linkage), whichever came first; participants remain under examination, but events after that date are not captured by the registry.

Covariates

Three nested models were specified. Model 1 adjusted for age. Model 2 additionally adjusted for examination year, centre, body mass index, waist circumference, smoking status, alcohol intake, and physical activity measured by the International Physical Activity Questionnaire. Model 3, the primary model, additionally adjusted for menopausal status, parity, age at menarche, hormone replacement therapy use (ever versus never), family history of breast cancer, diabetes mellitus, and hypertension. Diabetes mellitus was defined as a fasting serum glucose of 126 mg/dL (7.0 mmol/L) or higher, a glycated haemoglobin of 6.5% or higher, or current use of glucose-lowering medication; hypertension as a systolic blood pressure of 140 mmHg or higher, a diastolic blood pressure of 90 mmHg or higher, or current use of antihypertensive medication.

In the models that entered fat mass and skeletal muscle mass simultaneously, body mass index was omitted because it is a deterministic function of the two exposures, and waist circumference was omitted because fat mass measures the same construct directly. Retaining waist

circumference raised the variance inflation factor of fat mass from 1.34 to 4.40, with a value of 5.34 for waist circumference itself; that specification is reported as a sensitivity analysis. Variance inflation factors for all models are given in Additional file 1: Table S2.

Statistical analysis

The association of each exposure with a dense breast was estimated by multivariable logistic regression, and its association with incident breast cancer by Cox proportional hazards regression with time since the baseline examination as the time scale. Unless stated otherwise, confidence intervals for these regression estimates are Wald intervals.

Causal mediation analysis followed the approach of Valeri and VanderWeele, which allows the exposure and the mediator to interact [26, 27]. Two models were fitted. The mediator model was a logistic regression of dense breast on the exposure and covariates. The outcome model was a Cox regression of breast cancer on the exposure, dense breast, their product term, and covariates. From these two models the total effect of the exposure on breast cancer was split into the part that passes through density (the natural indirect effect) and the part that does not (the natural direct effect). The exposure was centred, so the hazard ratio for dense breast (the b-path) is reported at the mean exposure of the analysis sample (SMI 37.72%) for continuous scales and at the first quartile for the quartile contrast. Confidence intervals for all quantities from the mediation models, including the b-path, were nonparametric bootstrap percentile intervals from 1,000 resamples [28]; such intervals need not be symmetric on the log scale. The analysis was run in R with the survival and boot packages; code is available from the corresponding author. Because hazard ratios are non-collapsible (a hazard ratio conditional on covariates differs from the marginal one even without confounding), all effects were also estimated on the risk ratio

scale; the two scales agreed closely and both are reported in Additional file 1: Tables S5, S6 and S8. We do not report the proportion mediated for SMI, because when the total effect is close to 1 and the indirect effect is above 1 this proportion has no meaningful value. Mediation analyses were repeated separately in premenopausal and postmenopausal women.

Effect modification of the density–breast cancer association was examined in two ways. First, the b-path was estimated separately within quartiles of SMI and within quartiles of fat mass, and separately in premenopausal and postmenopausal women. Second, product terms of dense breast with SMI, with fat mass, and with menopausal status were tested in the pooled sample by likelihood ratio test. Because the stratum-specific estimates and the product-term estimates come from different models, the ratio of two stratum-specific hazard ratios need not equal the interaction hazard ratio.

Null findings were interpreted in light of what the study could have detected, rather than reported as absence of an effect. For each null estimate we calculated the minimum detectable effect at 80% power and a two-sided alpha of 0.05 from the observed standard error, and we calculated E-values for the point estimate and for the confidence limit closest to the null [29].

The proportional hazards assumption was assessed by Schoenfeld residual tests for every Cox model, globally and for each term. Where a covariate showed evidence of non-proportionality the model was refitted with the baseline hazard stratified by that covariate. Prespecified secondary analyses comprised a minimally adjusted model that removed body mass index, waist circumference, physical activity, and diabetes, because each may lie on a causal pathway from body composition to breast cancer, while retaining the remaining Model 3 covariates; and the height specifications described above. Analyses were performed in R version 4.3 or later

(packages survival, boot). Two-sided P values below 0.05 were considered statistically significant.

Results

Study population

The full analysis cohort comprised 106,813 women with a mean age of 39.7 years (SD 8.9), of whom 86.1% were premenopausal (Table 1). Over 695,501 person-years (mean follow-up 6.5 years) 891 women developed breast cancer, a crude incidence of 1.28 per 1,000 person-years. Mean body mass index was 22.0 kg/m² (SD 3.2), mean fat mass 17.3 kg (SD 5.9), mean skeletal muscle mass 21.2 kg (SD 2.5), and mean skeletal muscle percentage 37.8% (SD 3.3).

Table 1 Baseline characteristics of the full analysis cohort and the mammographic density subcohort[a]

Characteristic	Full analysis cohort (n = 106,813)	Density subcohort (n = 76,551)
Age, years	39.68 (8.90)	41.95 (8.86)
Anthropometry		
Height, cm	160.46 (5.32)	160.08 (5.34)
Weight, kg	56.59 (8.49)	56.84 (8.42)
Body mass index, kg/m²	21.99 (3.17)	22.19 (3.17)
Waist circumference, cm	76.34 (8.38)	76.80 (8.39)
Body composition		
Fat mass, kg	17.32 (5.87)	17.47 (5.84)
Fat-free mass, kg	36.99 (3.87)	37.10 (3.88)
Skeletal muscle mass, kg	21.21 (2.45)	21.28 (2.46)
Skeletal muscle percentage (SMI), %	37.78 (3.33)	37.72 (3.34)
Skeletal muscle mass index, kg/m²	8.23 (0.77)	8.29 (0.77)
Reproductive factors		
Postmenopausal	14,816 (13.9)	13,229 (17.3)

Parity ≥ 1	88,933 (83.3)	67,531 (88.2)
Age at menarche, years	13.86 (1.64)	13.99 (1.65)
Hormone replacement therapy use	1,621 (1.5)	1,181 (1.5)
Family history of breast cancer	3,156 (3.0)	2,462 (3.2)
Lifestyle		
Smoking status		
Never	96,010 (89.9)	68,644 (89.7)
Former	8,242 (7.7)	5,989 (7.8)
Current	2,561 (2.4)	1,918 (2.5)
Alcohol intake, g/day[b]	2.00 [0.00, 5.00]	2.00 [0.00, 4.00]
Physical activity (IPAQ)		
Low	62,546 (58.6)	43,270 (56.5)
Moderate	29,802 (27.9)	21,762 (28.4)
High	14,465 (13.5)	11,519 (15.0)
Comorbidity		
Diabetes mellitus	2,252 (2.1)	2,031 (2.7)
Hypertension	11,130 (10.4)	10,061 (13.1)
Examination and follow-up		
Year of examination	2014.89 (2.79)	2014.72 (2.74)
Follow-up, years	6.51 (2.83)	6.64 (2.79)
Incident breast cancer	891 (0.8)	713 (0.9)
Mammographic density (BI-RADS)		
A, almost entirely fatty[c]	—	659 (0.9)
B, scattered fibroglandular	—	7,386 (9.6)
C, heterogeneously dense	—	35,036 (45.8)
D, extremely dense	—	33,470 (43.7)

[a] Values are mean (standard deviation) or number (percentage) unless stated otherwise. Women were aged 30 years or older at baseline and had complete data on all covariates and on bioelectrical impedance measures of skeletal muscle mass, fat mass, and height.

[b] Median [interquartile range].

[c] Mammographic density was recorded only in the density subcohort; the remaining 30,262 women had no BI-RADS classification, predominantly because they had not undergone mammography at the study centres. A comparison of women with and without density data is given in Additional file 1.

BI-RADS Breast Imaging Reporting and Data System, IPAQ International Physical Activity Questionnaire, SMI skeletal muscle percentage (skeletal muscle mass / body weight × 100)

A BI-RADS density category was recorded for 76,551 women (71.7%), among whom 713 breast cancers occurred. Extremely dense breasts (category D) accounted for 43.7% and heterogeneously dense breasts (category C) for a further 45.8%, so that only 8,045 women (10.5%) fell into categories A or B. Women without a density record were considerably younger than those with one (mean age 33.9 versus 42.0 years; standardised mean difference 1.06) but differed little in body composition (standardised mean differences 0.062 for skeletal muscle percentage, 0.094 for skeletal muscle mass, and 0.090 for fat mass; Additional file 1: Table S3).

Body fat and skeletal muscle in relation to density and to breast cancer

When fat mass and skeletal muscle mass were entered into the same model, they were associated with mammographic density in opposite directions (Table 2). Each 1 kg of fat mass was associated with 14.1% lower odds of a dense breast (odds ratio 0.859; 95% CI 0.856–0.863) and each 1 kg of skeletal muscle mass with 1.1% higher odds (1.011; 1.003–1.020). Placed on a common scale, 1 SD of fat mass corresponded to an odds ratio of 0.412 (0.403–0.423) and 1 SD of skeletal muscle mass to 1.027 (1.007–1.050).

Table 2 Body fat and skeletal muscle mass in relation to mammographic density and breast cancer

Exposure	Dense breast (BI-RADS D)[a] OR (95% CI)	Incident breast cancer[b] HR (95% CI)
Primary model[c]		
Fat mass		
Per 1 kg	0.859 (0.856–0.863)	0.996 (0.982–1.009)
Per 1 SD (5.84 kg)[e]	0.412 (0.403–0.423)	0.977 (0.899–1.054)
Skeletal muscle mass		
Per 1 kg	1.011 (1.003–1.020)	1.029 (0.999–1.061)

Per 1 SD (2.46 kg)[e]	1.027 (1.007–1.050)	1.073 (0.998–1.157)
Sensitivity: additionally adjusted for waist circumference[d]		
Fat mass		
Per 1 kg	0.918 (0.912–0.924)	1.019 (0.995–1.044)
Per 1 SD (5.84 kg)[e]	0.607 (0.584–0.630)	1.116 (0.971–1.286)
Skeletal muscle mass		
Per 1 kg	1.057 (1.048–1.067)	1.046 (1.012–1.081)
Per 1 SD (2.46 kg)[e]	1.146 (1.122–1.173)	1.117 (1.030–1.211)

[a] Multivariable logistic regression of dense breast (BI-RADS D versus A/B/C) in the mammographic density subcohort (n = 76,551).

[b] Cox proportional hazards regression in the full analysis cohort (n = 106,813; 891 incident breast cancers).

[c] Fat mass and skeletal muscle mass were entered simultaneously and adjusted for age, examination year, centre, smoking status, alcohol intake, physical activity, menopausal status, parity, age at menarche, hormone replacement therapy use, family history of breast cancer, diabetes, and hypertension. Body mass index was omitted because it is a deterministic function of the two exposures, and waist circumference was omitted because fat mass is a direct measure of the same construct. Variance inflation factors in this model did not exceed 2.69.

[d] The same model with waist circumference added. The variance inflation factor of fat mass rose from 1.34 in the primary model to 4.40, and that of waist circumference was 5.34. The fat mass estimates were similar in both models, whereas the association of skeletal muscle mass with dense breast was larger when waist circumference was included.

[e] Estimates per 1 SD were obtained by rescaling the per-unit log estimates; SD values are those of the mammographic density subcohort.

Note. The proportional hazards assumption, assessed in the specification including waist circumference, was satisfied for both exposures (P = 0.799 for fat mass and P = 0.172 for skeletal muscle mass) and globally (P = 0.079). Age showed evidence of non-proportionality (P = 0.002); with the baseline hazard stratified by age group the estimates were essentially unchanged (fat mass 1.015, 95% CI 0.991–1.039; skeletal muscle mass 1.034, 1.000–1.069, per 1 kg).

CI confidence interval, HR hazard ratio, OR odds ratio, SD standard deviation

Neither association was mirrored in breast cancer. Per 1 SD, the hazard ratio was 0.977 (0.899–1.054) for fat mass and 1.073 (0.998–1.157) for skeletal muscle mass; the latter lay at the margin of statistical significance and, as shown below, absolute skeletal muscle mass carries body size in this model. The compartment that moved density most therefore moved risk least, and did so in the direction opposite to the one its density association would predict. Figure 1 shows fat mass,

skeletal muscle mass, and skeletal muscle percentage positioned by their association with density on one axis and with breast cancer on the other, drawn to the same scale: the three indices span a wide range on the density axis, from 0.412 to 1.413, but are confined to a narrow band close to unity on the breast cancer axis.

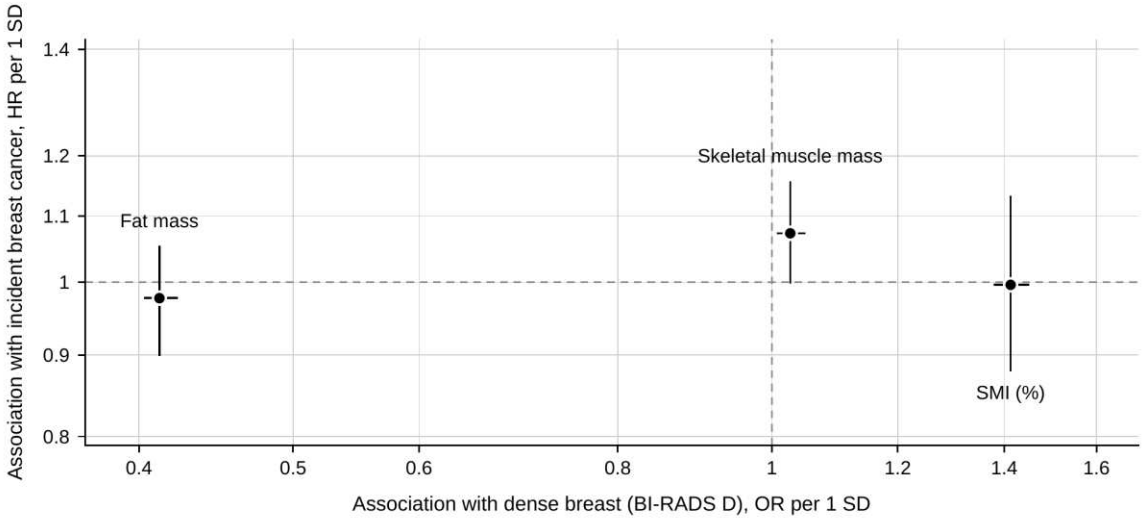


Fig. 1 Associations of three body-composition indices with an extremely dense breast and with incident breast cancer, per 1 SD, drawn to the same logarithmic scale on both axes. Horizontal position: odds ratio for BI-RADS D versus A/B/C from logistic regression in the density subcohort (n = 76,551). Vertical position: hazard ratio for incident breast cancer. Fat mass and skeletal muscle mass are from the joint model of Table 2 (both exposures entered simultaneously; body mass index and waist circumference omitted), with the hazard ratio estimated in the full cohort (n = 106,813; 891 events); SMI is from the mediation models of Table 3 (total effect, density subcohort), which additionally adjust for body mass index and waist circumference. SD values are those of the density subcohort (fat mass 5.84 kg, skeletal muscle mass 2.46 kg, SMI 3.34%). Error bars are 95% confidence intervals (Wald for fat mass and skeletal muscle mass; bootstrap percentile for SMI). Dashed lines mark the null on each axis.

Adding waist circumference to this model, which raised its variance inflation factor to 5.34 and that of fat mass to 4.40, left the fat mass estimates similar (odds ratio 0.918; 0.912–0.924 for dense breast and hazard ratio 1.019; 0.995–1.044 for breast cancer) but increased the association of skeletal muscle mass with dense breast to 1.057 (1.048–1.067) and with breast cancer to 1.046

(1.012–1.081) per 1 kg (Table 2). The direction of both associations was the same in the two specifications; the magnitude of the skeletal muscle estimates was not.

Density-mediated and direct effects of skeletal muscle percentage

Skeletal muscle percentage was strongly associated with mammographic density. The crude prevalence of an extremely dense breast rose from 22.8% in the lowest quartile to 37.3%, 49.3%, and 65.5% in the successive quartiles (Additional file 1: Table S4). After adjustment for body mass index, waist circumference, and the other Model 3 covariates, which fixes overall body size so that the contrast is between muscle and fat at a given body mass index, the odds ratio was 2.410 (2.267–2.583) for the highest versus the lowest quartile and 1.109 (1.101–1.118) per 1% (Table 3).

Table 3 Effect of skeletal muscle percentage on breast cancer decomposed through mammographic density[a,b]

Effect	Per 1% SMI	Q4 versus Q1
SMI to dense breast (a-path), OR	1.109 (1.101–1.118)	2.410 (2.267–2.583)
Dense breast to breast cancer (b-path), HR[c]	1.459 (1.230–1.734)	1.242 (0.918–1.689)
Total effect, HR[d]	0.999 (0.962–1.038)	0.982 (0.708–1.368)
Natural direct effect, HR	0.989 (0.951–1.028)	0.879 (0.615–1.238)
Natural indirect effect, HR	1.010 (1.006–1.015)	1.117 (1.060–1.181)

[a] Mammographic density subcohort (n = 76,551; 713 incident breast cancers). Causal mediation analysis using the Valeri–VanderWeele estimator with an exposure–mediator interaction term. Mediator model: logistic regression of dense breast (BI-RADS D versus A/B/C). Outcome model: Cox proportional hazards regression. Confidence intervals are nonparametric bootstrap percentile intervals with 1,000 resamples.

[b] Both models adjusted for age, examination year, centre, body mass index, waist circumference, smoking status, alcohol intake, physical activity, menopausal status, parity, age at menarche, hormone replacement therapy use, family history of breast cancer, diabetes, and hypertension.

[c] Evaluated at the mean exposure of the analysis sample (SMI 37.72%) for the continuous scale and at the first quartile for the quartile contrast.

[d] The minimum total effect detectable at 80% power with a two-sided alpha of 0.05 was 1.056 per 1% SMI and 1.601 for Q4 versus Q1. Total effects smaller than these values could not be distinguished from the null in this sample.

Note. The proportional hazards assumption was satisfied for the exposure (P = 0.748), the mediator (P = 0.216) and their product term (P = 0.259), and globally (P = 0.059). With the baseline hazard stratified by age group the estimates were essentially unchanged (SMI 0.992, 95% CI 0.959–1.027 per 1%; dense breast 1.444, 1.204–1.733). The proportion mediated is not reported because the total effect approximates unity while the natural indirect effect exceeds unity, which leaves the quantity undefined. Estimates per 1 SD, estimates on the risk ratio scale, and results stratified by menopausal status are given in Additional file 1.

BI-RADS Breast Imaging Reporting and Data System, CI confidence interval, HR hazard ratio, OR odds ratio, SMI skeletal muscle percentage (skeletal muscle mass / body weight × 100)

A dense breast was in turn associated with breast cancer, with a hazard ratio of 1.459 (1.230–1.734) evaluated at the mean exposure and 1.242 (0.918–1.689) evaluated in the lowest SMI quartile. The product of these two steps, however, was modest. The natural indirect effect through density was 1.117 (1.060–1.181) for the highest versus the lowest quartile and 1.010 (1.006–1.015) per 1% (Figure 2). More than doubling the odds of an extremely dense breast therefore corresponded to a 12% increase in breast cancer risk transmitted through density.

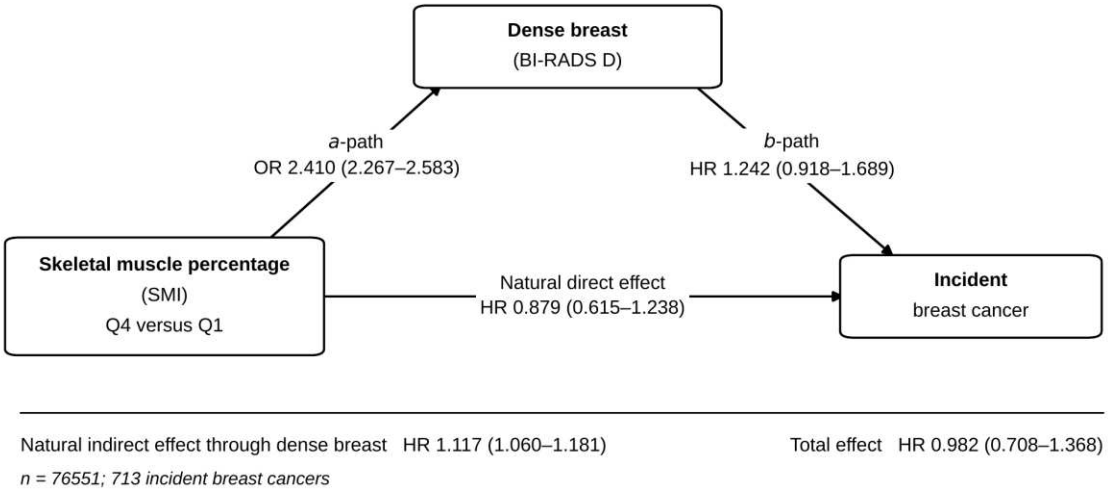


Fig. 2 Decomposition of the effect of skeletal muscle percentage (SMI) on breast cancer through mammographic density. Highest versus lowest SMI quartile in the density subcohort (n = 76,551; 713 incident invasive breast cancers). The a-path is the odds ratio for an extremely dense breast (BI-RADS D versus A/B/C) from the mediator model; the b-path is the hazard ratio for a dense breast from the outcome model, evaluated in the lowest SMI quartile; the natural direct effect, natural indirect effect, and total effect are from the Valeri–VanderWeele decomposition with an exposure–mediator interaction term. Both models were adjusted for age, examination year, centre, body mass index, waist circumference, smoking status, alcohol intake, physical activity, menopausal status, parity, age at menarche, hormone replacement therapy use, family history of breast cancer, diabetes, and hypertension. All intervals are 95% nonparametric bootstrap percentile intervals (1,000 resamples).

The natural direct effect was 0.879 (0.615–1.238) for the highest versus the lowest quartile and the total effect was 0.982 (0.708–1.368). Both intervals were wide. The smallest total effect this sample could have detected with 80% power was 1.601 for the quartile contrast and 1.056 per 1%, so the total effect is uninformative about differences of the size implied by the indirect effect, and we do not interpret it as evidence of no association. For the natural indirect effect, which was estimated precisely, the E-value was 1.48 for the point estimate and 1.31 for the lower confidence limit: an unmeasured confounder would need to be associated with both exposure and outcome by a risk ratio of at least 1.31 to move the lower limit to the null.

Modification of the density–breast cancer association

If the density produced by a high muscle-to-fat ratio were less consequential than density of other origin, the association between a dense breast and breast cancer should weaken as skeletal muscle percentage rises. It did not (Figure 3). Within quartiles of skeletal muscle percentage the hazard ratios for a dense breast were 1.299 (0.854–1.977), 1.554 (1.082–2.232), 1.358 (0.953–1.935), and 1.665 (1.165–2.380), and the formal interaction was 1.038 per 1% (0.989–1.089; P = 0.125), that is, directionally opposite to the prediction. The same analysis across quartiles of fat mass gave 1.737 (1.213–2.486), 1.208 (0.848–1.719), 1.512 (1.043–2.193), and 1.436 (0.948–2.174), with an interaction of 0.980 per 1 kg (0.954–1.007; P = 0.149). Because skeletal muscle percentage and fat mass are inversely correlated (r = −0.871), these two stratifications order the same women in nearly opposite ways, and both indicate that the hazard associated with a dense breast is if anything greater, not smaller, in women with less fat and more muscle. Under the alternative density definition the SMI interaction was 0.994 (0.891–1.109; P = 0.909).

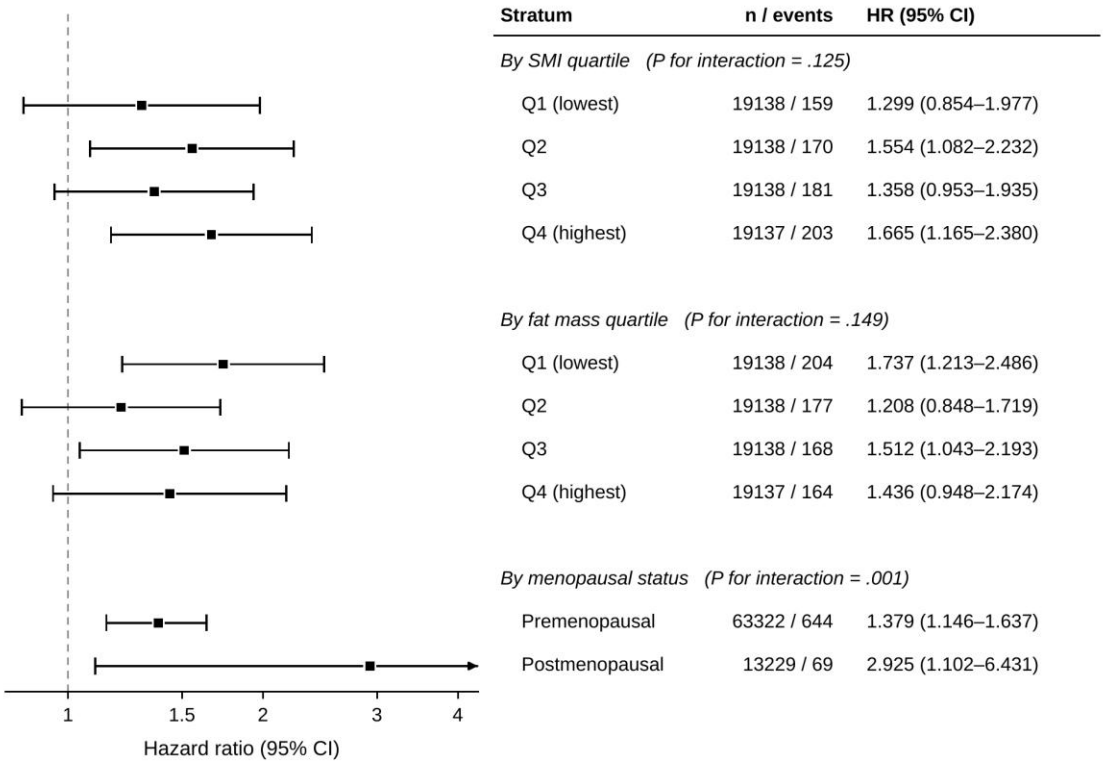


Fig. 3 Hazard ratios for an extremely dense breast (BI-RADS D versus A/B/C) and incident breast cancer within strata of skeletal muscle percentage (SMI), fat mass, and menopausal status, in the density subcohort (n = 76,551). Estimates for SMI and fat mass quartiles are hazard ratios for a dense breast from Cox models fitted within each stratum and adjusted for the covariates listed in the legend to Fig. 2 (Wald 95% confidence intervals); estimates for the menopausal strata are the b-path from mediation models fitted within each stratum (bootstrap percentile 95% confidence intervals). P values for interaction are from likelihood ratio tests of a product term between dense breast and the continuous modifier (SMI per 1%, fat mass per 1 kg) or menopausal status in the pooled sample; because the pooled and stratified models differ, the ratio of two stratum-specific hazard ratios need not equal the interaction hazard ratio. The arrow indicates that the upper confidence limit for postmenopausal women (6.431) extends beyond the axis.

Menopausal status, by contrast, modified the association strongly. In mediation models fitted separately by menopausal status, the hazard ratio for a dense breast was 1.379 (1.146–1.637) in premenopausal women (63,322 women, 644 events) and 2.925 (1.102–6.431) in postmenopausal women (13,229 women, 69 events; bootstrap intervals). In the pooled model the interaction

hazard ratio was 2.664 (1.537–4.618; P = 0.001). The postmenopausal estimate rests on 69 events and its confidence interval is correspondingly wide, and under the alternative density definition the interaction was 1.506 (0.701–3.239; P = 0.289). The stronger downstream association after menopause did not translate into a larger density-mediated effect of skeletal muscle percentage: the natural indirect effect per 1% was 1.009 (1.004–1.014) before and 1.007 (1.001–1.013) after menopause, because the upstream association was somewhat weaker after menopause (odds ratio per 1% 1.110 before versus 1.081 after) and the prevalence of an extremely dense breast lower (Additional file 1: Table S6).

Alternative exposure specifications and model diagnostics

Absolute skeletal muscle mass was associated with breast cancer when height was not accounted for (highest versus lowest quartile 1.309; 1.057–1.712 in the density subcohort, bootstrap interval, and 1.313; 1.072–1.608 in the full cohort), but the association disappeared after adjustment for height (1.038; 0.729–1.358) and when muscle mass was expressed per height squared (SMMI, 1.041; 0.778–1.306; Additional file 1: Tables S7–S9). Skeletal muscle mass was correlated with height (r = 0.592), and mean height rose from 155.9 cm in the lowest quartile of muscle mass to 164.1 cm in the highest, whereas the height-normalised index was uncorrelated with height (r = 0.020). In a minimally adjusted model that removed body mass index, waist circumference, physical activity, and diabetes, none of the three exposures was associated with breast cancer (skeletal muscle percentage 1.006; 0.985–1.027 per 1%; skeletal muscle mass 1.024; 0.996–1.052 per 1 kg; SMMI 0.987; 0.903–1.078; Additional file 1: Table S14).

Skeletal muscle percentage was correlated with fat mass at r = −0.871 but with skeletal muscle mass at only r = 0.072. Mean skeletal muscle mass differed little across categories of physical activity (21.14, 21.29, and 21.77 kg from low to high), and the crude proportion developing breast cancer was similar across those categories (0.95%, 0.89%, and 0.94%). All estimates on the risk ratio scale agreed closely with the hazard ratio estimates (Additional file 1: Tables S5, S6 and S8). Minimum detectable effects and E-values for all null estimates are given in Additional file 1: Tables S10 and S11, and the stratum-specific and interaction estimates in Table S12.

The proportional hazards assumption was satisfied globally in all Cox models (P = 0.079 for the joint fat and muscle model, assessed in the specification including waist circumference, and P = 0.059 and 0.067 for the two mediation outcome models) and for every term of primary interest: fat mass (P = 0.799), skeletal muscle mass (P = 0.172), skeletal muscle percentage (P = 0.748 and 0.668), dense breast (P = 0.216 and 0.247), and the exposure–mediator product term (P = 0.259). Age showed evidence of non-proportionality (P = 0.002 to 0.014), as did hormone replacement therapy use (P = 0.022 to 0.062), which was recorded in 1.5% of the cohort. With the baseline hazard stratified by age group the estimates were essentially unchanged (Additional file 1: Table S13): fat mass 1.015 (0.991–1.039) and skeletal muscle mass 1.034 (1.000–1.069) per 1 kg, skeletal muscle percentage 0.992 (0.959–1.027) per 1%, and dense breast 1.444 (1.204–1.733).

Discussion

In this prospective cohort of 106,813 Korean women, body fat was strongly and inversely associated with mammographic density and skeletal muscle only weakly and positively, but

neither compartment was significantly associated with breast cancer. Because an extremely dense breast carried a modest hazard, a body-composition contrast that more than doubled the odds of extreme density corresponded to a 12% difference in risk through density. The association between an extremely dense breast and breast cancer was not detectably weaker in the women whose low body fat had produced the density, but it was modest before menopause and substantially stronger after it.

The association of body fat with density but not with breast cancer gives an arithmetic account of a long-standing anomaly. Body mass index is inversely associated with density yet positively associated with postmenopausal breast cancer [6, 7, 9]. Our data show that the two compartments body mass index sums behave oppositely on density: per 1 SD, fat mass was associated with an odds ratio of 0.412 for a dense breast and skeletal muscle mass with 1.027. When the two parts of a sum point in opposite directions, the sum cannot show which part is acting, so how body mass index relates to density tells us nothing about how it relates to breast cancer. The same reasoning applies to any composite anthropometric measure. The fat mass result was the more robust of the two, being similar whether or not waist circumference was retained, whereas the magnitude of the skeletal muscle association with density depended on that adjustment; our argument therefore rests principally on adiposity, the compartment that dominates density in this population. We also note that fat mass was not inversely associated with breast cancer despite the predominantly premenopausal composition of the cohort, in contrast to Western populations [10, 11]; adiposity here is considerably lower (mean body mass index 22.0 kg/m^2), and the protective premenopausal association may not extend to this range.

The mediation analysis shows how much of a determinant's effect can pass through density. The prevalence of an extremely dense breast rose from 22.8% in the lowest quartile of skeletal

muscle percentage to 65.5% in the highest, a larger change than most modifiable exposures produce, yet the indirect effect through density was only 1.117. This is not a paradox. The risk that reaches breast cancer through density can never be larger than the risk a dense breast itself carries, and in this cohort a dense breast carried a hazard ratio of about 1.24 to 1.46. For proposals to lower breast cancer risk by lowering density, this means that the benefit is limited by how much risk a dense breast carries, not by how much density can be changed. Comparison with other exposures suggests that which determinant changes density, and when in life it acts, matter more than the size of the change. Body size in early life appears to act on breast cancer largely through density [13], adult body shape shows more modest mediation [14], and adult body composition here transmitted little. One plausible explanation is that exposures during development change the amount of epithelium at risk, whereas adult fat mainly changes the radiographic background against which that epithelium is seen. The same distinction fits the tamoxifen data [4]. A fall in density produced by an antioestrogen reflects regression of the epithelial and stromal tissue that carries the risk; a fall in density produced by gaining fat does not. A change in density therefore predicts a change in risk only when the determinant acts on the tissue itself rather than on its radiographic background.

The effect-modification analysis addresses the most natural alternative explanation for this small transmitted effect. If a woman with little body fat has a dense breast because there is little fat to dilute the fibroglandular tissue, rather than because that tissue is abundant, the density so produced might carry less risk. We tested this directly and found no support. The hazard associated with a dense breast did not fall across quartiles of skeletal muscle percentage (1.299, 1.554, 1.358, and 1.665; interaction P = 0.125), and stratification by fat mass, which orders the same women in almost the opposite direction, gave a concordant answer (interaction P = 0.149),

with the strongest association in the lowest quartile of fat mass. Both interactions were directionally opposite to the prediction. Because the two stratifications order the same women almost inversely they are not independent tests, but their concordance shows that the null does not depend on which side of the body-composition axis is used to define the strata. The tests could detect a hazard ratio of 1.071 per 1% of skeletal muscle percentage, so we can exclude a large modification but not a modest one; and the alternative density definition was uninformative here because it left only 10.5% of women in the non-dense group.

Menopausal status modified the association strongly, and it is the premenopausal estimate that governs most of this cohort. The hazard ratio for a dense breast was 1.379 before menopause and 2.925 after it (interaction P = 0.001), so among the 86% of women who were premenopausal a dense breast conferred a modest excess. Set against this, the odds of being classified as having a dense breast varied by a factor of 2.4 across quartiles of skeletal muscle percentage, and the prevalence of an extremely dense breast rose from 22.8% to 65.5% across those quartiles. The upstream association, from body composition to density, is therefore stronger than the downstream association from density to breast cancer. Because the hazard carried by a dense breast did not depend on what produced it, density remains a valid marker of risk; but a marker whose level is substantially set by body habitus, and which conveys little of the effect of the factors that set it, is a poor lever. This distinction between predicting risk and conveying it is the practical message of these data, and it applies most directly to premenopausal women. One qualification is needed: our primary dichotomy places heterogeneously dense breasts in the non-dense group, so the modest premenopausal hazard ratio is in part a property of the contrast we chose and is not directly comparable with estimates that set categories C and D against A and B.

The stronger association after menopause has a plausible basis. Breast density declines with menopause, so dense tissue persisting afterwards represents a greater departure from the expected state and may mark a different biology from the common dense breast of a premenopausal woman. This is consistent with reports that the density association is not diluted with time since mammography [30] and with the greater stromal and epithelial content of dense tissue [31]. The postmenopausal estimate rests on 69 events, the interaction was not significant under the alternative density definition, and we therefore emphasise the interaction test under the primary definition rather than the magnitude of the stratum-specific estimate. Notably, the stronger postmenopausal association did not produce a larger density-mediated effect of skeletal muscle percentage, because the upstream step was weaker and extremely dense breasts rarer after menopause; the size of an indirect effect is set by the product of the two steps, not by either alone.

Our findings also bear on what body-composition indices measure. Skeletal muscle percentage correlated with fat mass at $r = -0.871$ but with skeletal muscle mass at only $r = 0.072$, so in this population it behaves as an inverse adiposity index rather than a muscle index. Absolute skeletal muscle mass was associated with breast cancer, but the association was explained by body size: muscle mass correlated with height at $r = 0.592$, mean height rose by 8.2 cm across quartiles of muscle mass, and the pooled relative risk of 1.17 per 10 cm of height [32] predicts roughly half of the observed log hazard ratio; the whole of the association disappeared after height adjustment and when muscle mass was normalised to height squared. This is directly relevant to the report of an odds ratio of 1.60 for the highest versus lowest quintile of whole-body fat-free mass and postmenopausal breast cancer in UK Biobank [33], which did not adjust for height. Height and lean mass share determinants in the growth hormone and insulin-like growth factor axis [34, 35],

so height adjustment may be over-adjustment if that axis is the exposure of interest; we therefore report both specifications rather than choosing between them. Either way, lean mass and stature cannot be separated in these data.

This study has several strengths: a large cohort with registry-based outcome ascertainment, direct measurement of body composition rather than anthropometric proxies, measurement of exposure, mediator, and covariates at a single examination so that no time elapsed between them, simultaneous modelling of the two compartments, prespecified tests of the competing explanation for the mediation result using two inversely ordered stratifications, and the reporting of minimum detectable effects and E-values so that null findings are interpreted against their statistical resolution.

Several limitations should be considered. First, the total effect of skeletal muscle percentage was not estimable at the relevant magnitude: the minimum detectable effect was 1.601 for the quartile contrast, larger than the indirect effect it would have had to reveal, so the null total effect is not evidence of no association and we have not used it as such. Second, the natural direct effect was compatible with effects in both directions, and we make no claim of a protective direct pathway. Third, the interaction tests could exclude a large but not a modest modification of the density association by body composition. Fourth, the magnitude of the skeletal muscle association with density depended on whether waist circumference was retained, and the joint model was not adjusted for height, so the muscle coefficient contains body size; the skeletal muscle hazard ratio in that model lay at the margin of significance, and although the muscle association disappeared with height adjustment in the single-exposure analyses, a small positive association of lean mass with breast cancer cannot be formally excluded. Neither point affects the fat mass estimates on which our main argument rests. Fifth, the associations of fat mass and skeletal muscle mass with

551 breast cancer were estimated in the pooled cohort and not separately by menopausal status; a null
552 in a cohort that is 86% premenopausal cannot exclude a positive association after menopause of
553 the kind reported elsewhere. Sixth, exposure, mediator, and covariates, including hormone
554 replacement therapy use, were measured once at baseline, and non-differential misclassification
555 arising from change during follow-up would bias estimates towards the null. Seventh, a dense
556 breast lowers the sensitivity of mammography, so some cancers in dense breasts will have been
557 diagnosed later than they would otherwise have been; because the outcome was ascertained by
558 registry linkage rather than by screening at the study centres, this delays rather than removes
559 events, but it could attenuate the b-path and hence the indirect effect. Eighth, BI-RADS is a
560 categorical, single-reader assessment with known inter-reader variability, and it does not permit
561 separation of percent density, absolute dense area, and non-dense area, which recent work
562 suggests may mediate differently; reader variability unrelated to future breast cancer would
563 attenuate the mediator–outcome association and hence the indirect effect, so our findings apply
564 to density as it is reported in practice rather than to a quantitative measure. Ninth, analyses were
565 restricted to complete cases, which excluded 24% of otherwise eligible women; we did not
566 formally compare included and excluded women, and the analysis assumes that missingness was
567 unrelated to breast cancer given the covariates. Density was recorded for 71.7% of the cohort,
568 those without a record being markedly younger, although they differed little in body
569 composition, which supports the exposure-representativeness of the mediation subcohort. Tenth,
570 the cohort consists of employed adults and their spouses attending health examinations and is
571 Korean, predominantly premenopausal, and has less adiposity than Western populations; a
572 healthy-participant effect is likely, only 69 postmenopausal events were available, and the
573 findings should not be extrapolated to older or postmenopausal populations without

confirmation. Eleventh, because registry linkage is renewed only for women who continue to attend and renew consent to linkage, about half the cohort was followed to the end of 2020 and half to the end of 2022. The censoring date is fixed within each linkage round and cannot generate events, and all models adjusted for examination year and centre, but women followed for longer are by definition those who continued to attend, and we cannot exclude that their breast cancer risk differs from that of women who did not. Finally, causal mediation analysis assumes no unmeasured mediator–outcome confounding; the E-value of 1.31 for the lower confidence limit of the indirect effect indicates that moderate unmeasured confounding could account for it.

Conclusions

Body fat was strongly and inversely associated with mammographic density, whereas neither body fat nor skeletal muscle was significantly associated with breast cancer. An extremely dense breast carried no detectably lower risk when it reflected low body fat, and the association was considerably stronger after menopause than before. Because the hazard of an extremely dense breast was modest, the large differences in density that accompany body composition corresponded to small differences in breast cancer risk. Density therefore remains an informative marker of risk across body habitus, but differences in density associated with body composition should not be read as comparable differences in risk, particularly in premenopausal women.

List of abbreviations

AGReMA: A Guideline for Reporting Mediation Analyses; BI-RADS: Breast Imaging Reporting and Data System; CI: confidence interval; HR: hazard ratio; ICD-10: International Classification

of Diseases, tenth revision; IPAQ: International Physical Activity Questionnaire; OR: odds ratio; SD: standard deviation; SMI: skeletal muscle percentage (skeletal muscle mass / body weight × 100); SMMI: skeletal muscle mass index (skeletal muscle mass / height², kg/m²); STROBE: Strengthening the Reporting of Observational Studies in Epidemiology.

Declarations

Ethics approval and consent to participate

The study was approved by the Institutional Review Board of Kangbuk Samsung Hospital (KBSMC 2026-03-023). Participants had consented at their examinations to the use of their records and to linkage with the national cancer registry; the Board waived the requirement for additional study-specific informed consent because only de-identified, retrospectively retrieved data were analysed. The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Clinical trial number

Not applicable.

Availability of data and materials

The data underlying this study are available from the corresponding author on reasonable request, subject to the data-sharing agreements of the Kangbuk Samsung Health Study and the

Korea National Cancer Center and to approval by the institutional review board. Analysis code is available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

No funding was received for this study.

Authors' contributions

YTK: methodology, formal analysis, software, data curation, visualisation, writing – review and editing. EYK, CHP, YLP, JSY, SHL: investigation, resources, writing – review and editing. KHL: conceptualisation, methodology, investigation, resources, writing – original draft, writing – review and editing, supervision, project administration. All authors read and approved the final manuscript.

Acknowledgements

The authors used a large language model assistant (Claude, Anthropic) during manuscript preparation for language editing, checking the internal consistency of reported statistics across sections, and formatting of submission documents. The tool was not used in study design, data collection, statistical analysis, or interpretation of results. The authors reviewed and edited all text and take full responsibility for the content of the manuscript.

References

1. Boyd NF, Guo H, Martin LJ, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med. 2007;356(3):227–36.

2. McCormack VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev. 2006;15(6):1159–69.

3. US Food and Drug Administration. Mammography Quality Standards Act; final rule. Fed Regist. 2023;88(47):15126–71 (Docket FDA-2013-N-0134; effective 10 September 2024).

4. Cuzick J, Warwick J, Pinney E, Duffy SW, Cawthorn S, Howell A, et al. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: a nested case–control study. J Natl Cancer Inst. 2011;103(9):744–52.

5. Eriksson M, Eklund M, Borgquist S, Hellgren R, Margolin S, Thoren L, et al. Low-dose tamoxifen for mammographic density reduction: a randomized controlled trial. J Clin Oncol. 2021;39(17):1899–908.

6. Pollán M, Lope V, Miranda-García J, García M, Casanova F, Sánchez-Contador C, et al. Adult weight gain, fat distribution and mammographic density in Spanish pre- and post-menopausal women (DDM-Spain). Breast Cancer Res Treat. 2012;134(2):823–38.

7. Woolcott CG, Cook LS, Courneya KS, Boyd NF, Yaffe MJ, Terry T, et al. Associations of overall and abdominal adiposity with area and volumetric mammographic measures among postmenopausal women. Int J Cancer. 2011;129(2):440–8.

8. Burton A, Maskarinec G, Perez-Gomez B, Vachon C, Miao H, Lajous M, et al. Mammographic density and ageing: a collaborative pooled analysis of cross-sectional data from 22 countries worldwide. PLoS Med. 2017;14(6):e1002335.

9. Neuhouser ML, Aragaki AK, Prentice RL, Manson JE, Chlebowski R, Carty CL, et al. Overweight, obesity, and postmenopausal invasive breast cancer risk: a secondary analysis of the Women's Health Initiative randomized clinical trials. JAMA Oncol. 2015;1(5):611–21.

10. Premenopausal Breast Cancer Collaborative Group; Schoemaker MJ, Nichols HB, Wright LB, Brook MN, Jones ME, et al. Association of body mass index and age with subsequent breast cancer risk in premenopausal women. JAMA Oncol. 2018;4(11):e181771.

11. 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;10:143–51.

12. Lee KH, Chae SW, Yun JS, Park YL, Park CH. Association between skeletal muscle mass and mammographic breast density. Sci Rep. 2021;11(1):16785.

13. Vabistsevits M, Davey Smith G, Richardson TG, Richmond RC, Sieh W, Rothstein JH, et al. Mammographic density mediates the protective effect of early-life body size on breast cancer risk. Nat Commun. 2024;15:4021.

14. Amadou A, Freisling H, Sedlmeier AM, Bohmann P, Fontvieille E, Weber A, et al. Multi-trait body shape phenotypes and breast cancer risk in postmenopausal women: a causal

673 mediation analysis in the UK Biobank cohort. J Epidemiol Glob Health. 2024;14(2):420–
674 32.

675 15. Kim BK, Chang Y, Ahn J, Jung HS, Kim CW, Yun KE, et al. Metabolic syndrome, insulin
676 resistance, and mammographic density in pre- and postmenopausal women. Breast
677 Cancer Res Treat. 2015;153(2):425–34.

678 16. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE
679 Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology
680 (STROBE) statement: guidelines for reporting observational studies. Lancet.
681 2007;370(9596):1453–7.

682 17. Lee H, Cashin AG, Lamb SE, Hopewell S, Vansteelandt S, VanderWeele TJ, et al. A
683 guideline for reporting mediation analyses of randomized trials and observational studies:
684 the AGReMA statement. JAMA. 2021;326(11):1045–56.

685 18. Malavolti M, Mussi C, Poli M, Fantuzzi AL, Salvioli G, Battistini N, et al. Cross-calibration
686 of eight-polar bioelectrical impedance analysis versus dual-energy X-ray absorptiometry
687 for the assessment of total and appendicular body composition in healthy subjects aged
688 21–82 years. Ann Hum Biol. 2003;30(4):380–91.

689 19. Tagliabue A, Andreoli A, Bertoli S, Comelli M, De Lorenzo A. Appendicular lean body
690 mass: prediction by bioelectrical impedance analysis. Ann N Y Acad Sci. 2000;904:218–
691 20.

692 20. Janssen I, Heymsfield SB, Ross R. Low relative skeletal muscle mass (sarcopenia) in older
693 persons is associated with functional impairment and physical disability. J Am Geriatr
694 Soc. 2002;50(5):889–96.

21. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019;48(1):16–31.

22. Chen LK, Woo J, Assantachai P, Auyeung TW, Chou MY, Iijima K, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc. 2020;21(3):300–7.e2.

23. American College of Radiology. ACR BI-RADS® Atlas: Breast Imaging Reporting and Data System. 5th ed. Reston, VA: American College of Radiology; 2013.

24. del Carmen MG, Halpern EF, Kopans DB, Moy B, Moore RH, Goss PE, et al. Mammographic breast density and race. AJR Am J Roentgenol. 2007;188(4):1147–50.

25. Shin HR, Won YJ, Jung KW, Kong HJ, Yim SH, Lee JK, et al. Nationwide cancer incidence in Korea, 1999–2001; first result using the national cancer incidence database. Cancer Res Treat. 2005;37(6):325–31.

26. VanderWeele TJ. Explanation in causal inference: methods for mediation and interaction. New York: Oxford University Press; 2015.

27. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013;18(2):137–50.

28. Canty A, Ripley B. boot: Bootstrap R (S-Plus) functions. R package version 1.3-28; 2022.

29. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med. 2017;167(4):268–74.

30. Yaghjyan L, Colditz GA, Rosner B, Tamimi RM. Mammographic breast density and subsequent risk of breast cancer in postmenopausal women according to the time since the mammogram. Cancer Epidemiol Biomarkers Prev. 2013;22(6):1110–7.

31. Huo CW, Chew G, Hill P, Huang D, Ingman W, Hodson L, et al. High mammographic density is associated with an increase in stromal collagen and immune cells within the mammary epithelium. Breast Cancer Res. 2015;17(1):79.

32. 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):djv219.

33. Harris BHL, Di Giovannantonio M, Zhang P, Harris DA, Lord SR, Allen NE, et al. New role of fat-free mass in cancer risk linked with genetic predisposition. Sci Rep. 2024;14:7270.

34. Le Roith D, Bondy C, Yakar S, Liu JL, Butler A. The somatomedin hypothesis: 2001. Endocr Rev. 2001;22(1):53–74.

35. Velloso CP. Regulation of muscle mass by growth hormone and IGF-I. Br J Pharmacol. 2008;154(3):557–68.

Additional files

Additional file 1. Supplementary Tables S1–S14 (DOCX): participant flow (S1); variance inflation factors (S2); baseline characteristics and crude incidence by availability of mammographic density (S3); BI-RADS distribution and prevalence of an extremely dense breast by SMI quartile (S4); mediation estimates on the hazard ratio and risk ratio scales (S5); mediation analysis by menopausal status (S6); alternative exposure specifications (S7);

mediation estimates per 1 SD (S8); height diagnostics (S9); minimum detectable effects (S10); E-values (S11); effect modification of the dense breast–breast cancer association (S12); proportional hazards tests and age-stratified estimates (S13); minimally adjusted models and physical activity diagnostics (S14).

Additional file 2. STROBE checklist (DOCX).

Additional file 3. AGReMA checklist (DOCX).