

Article

Is Mammographic Breast Density an Endophenotype for Breast Cancer?

Ellie Darcey ¹, Nina McCarthy ², Eric K. Moses ³, Christobel Saunders ⁴, Gemma Cadby ^{1,†}
and Jennifer Stone ^{1,*,†} 

¹ Genetic Epidemiology Group, School of Population and Global Health, University of Western Australia, Perth, WA 6009, Australia; ellie.darcey@uwa.edu.au (E.D.); gemma.cadby@uwa.edu.au (G.C.)

² School of Biomedical Science, University of Western Australia, Perth, WA 6009, Australia; nina.mccarthy@uwa.edu.au

³ Menzies Institute for Medical Research, The University of Tasmania, Hobart, TAS 7000, Australia; eric.moses@utas.edu.au

⁴ School of Medicine, University of Western Australia, Perth, WA 6009, Australia; christobel.saunders@uwa.edu.au

* Correspondence: jennifer.stone@uwa.edu.au; Tel.: +61-(0)8-6488-1603

† Joint Senior Author.

Simple Summary: Candidate endophenotypes should be systematically assessed against five criteria: (i) the endophenotype is associated with disease in the population; (ii) the endophenotype is heritable; (iii) within families, endophenotype and disease co-segregate; (iv) the endophenotype found in affected family members is found in non-affected family members at a higher rate than in the general population and (v) the endophenotype is primarily state independent (manifests in an individual whether or not disease is active). This study assesses the suitability of mammographic breast density as an endophenotype for breast cancer. Formally establishing a trait as a disease endophenotype confirms that the trait and endophenotype share a biological basis, thereby enabling genetic dissection of an endophenotype to inform disease risk. As breast density can be measured for any woman who has had a mammogram, studies investigating the genetic architecture of breast density could identify breast cancer risk variants that act through effects on this trait.

Abstract: Mammographic breast density (MBD) is a strong and highly heritable predictor of breast cancer risk and a biomarker for the disease. This study systematically assesses MBD as an endophenotype for breast cancer—a quantitative trait that is heritable and genetically correlated with disease risk. Using data from the family-based kConFab Study and the 1994/1995 cross-sectional Busselton Health Study, participants were divided into three status groups—cases, relatives of cases and controls. Participant’s mammograms were used to measure absolute dense area (DA) and percentage dense area (PDA). To address each endophenotype criterion, linear mixed models and heritability analysis were conducted. Both measures of MBD were significantly associated with breast cancer risk in two independent samples. These measures were also highly heritable. Meta-analyses of both studies showed that MBD measures were higher in cases compared to relatives ($\beta = 0.48$, 95% CI = 0.10, 0.86 and $\beta = 0.41$, 95% CI = 0.06, 0.78 for DA and PDA, respectively) and in relatives compared to controls ($\beta = 0.16$, 95% CI = -0.24 , 0.56 and $\beta = 0.16$, 95% CI = -0.21 , 0.53 for DA and PDA, respectively). This study formally demonstrates, for the first time, that MBD is an endophenotype for breast cancer.

Keywords: mammographic breast density; breast cancer; risk factor; endophenotype



Citation: Darcey, E.; McCarthy, N.; Moses, E.K.; Saunders, C.; Cadby, G.; Stone, J. Is Mammographic Breast Density an Endophenotype for Breast Cancer? *Cancers* **2021**, *13*, 3916.

<https://doi.org/10.3390/cancers13153916>

Academic Editor: Fabio Puglisi

Received: 14 June 2021

Accepted: 28 July 2021

Published: 3 August 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breast cancer is the second most commonly diagnosed cancer in Australia and the most commonly diagnosed cancer in females, with a lifetime prevalence of 1 in 7 [1]. Women with an affected first-degree female relative are at approximately two-fold greater risk of

developing breast cancer than women from the general population. Rare, highly penetrant mutations in genes such as *BRCA1/2*, identified by linkage/positional cloning in breast cancer families more than 20 years ago, remain the single largest known genetic risk factor for breast cancer, accounting for ~30% of excess familial risk [2]. In recent years, large genome-wide association studies (GWAS) in unrelated individuals have identified many common, low-risk alleles that together account for an additional ~18% of familial risk [3]. New study designs are needed to identify the remaining “missing heritability” and causal risk variants that also contribute to inter-individual differences in breast cancer susceptibility.

Mammographic breast density (MBD) is a strong [4,5] and highly heritable [6,7] predictor of breast cancer risk and considered a strong biomarker for the disease. MBD is the white appearance of epithelial and stromal tissue on a mammogram, in contrast to adipose (fatty) breast tissue which appears dark. Dense breast tissue is quite common, with ~43% of screen-aged women estimated to have heterogeneously or extremely dense breasts [8]. MBD is a modifiable risk factor [9] and it has been shown that reducing MBD by medical interventions such as tamoxifen is associated with significantly reduced breast cancer risk [10]. Previous investigations of the associations between known common breast cancer-susceptibility variants and MBD has demonstrated significant evidence of a shared genetic basis between MBD and breast cancer risk with ~18% overlap of genetic associations [11]. However, MBD has never been formally examined as an endophenotype for breast cancer—a quantitative trait that is heritable and genetically correlated with disease risk. Formally establishing a trait as a disease endophenotype confirms that the trait and endophenotype share a biological basis, thereby enabling genetic dissection of an endophenotype to inform disease risk. As MBD can be measured for any woman who has had a mammogram, study designs that quantitatively examine the genetic architecture of MBD could significantly help identify risk variants for breast cancer that act through effects on this trait. As MBD is a modifiable risk factor, analysis of the genetic overlap could also help identify other possible less-invasive interventions that could be used to target women at high risk of breast cancer and thereby aid prevention of the disease.

Candidate endophenotypes should be systematically assessed against five endophenotype criteria: [12] (i) the endophenotype is associated with disease in the population; (ii) the endophenotype is heritable; (iii) within families, endophenotype and disease co-segregate; (iv) the endophenotype found in affected family members is found in non-affected family members at a higher rate than in the general population and (v) the endophenotype is primarily state independent (manifests in an individual whether or not illness is active).

This study aims to assess the suitability of MBD measures as endophenotypes for breast cancer. As it is accepted that MBD is state independent of breast cancer (i.e., breasts can be dense both with and without the presence of breast cancer) [13], we investigate the suitability of MBD against the remaining four endophenotype criteria using supporting data from two epidemiological studies—the kConFab Consortium and the Busselton Health Study.

2. Materials and Methods

2.1. Study Participants

Two study populations were used to assess the suitability of mammographic density measures as endophenotypes for breast cancer—kConFab and the Busselton Health Study (BHS).

2.1.1. kConFab

kConFab (The Kathleen Cuninghame Foundation Consortium for research into Familial Breast cancer) has been collecting genetic, epidemiological, medical and psychosocial data from families with a strong history of breast cancer since 1997 and has accumulated data on more than 1400 multigenerational, multicase kindreds [14]. The consortium makes data and biospecimens widely available to researchers for use in peer-reviewed, ethically-approved research. Methods for participant recruitment and data collection are described in detail

elsewhere [14]. Via the Western Australian Department of Health Data Linkage Branch, we linked all kConFab participants residing in Western Australia (WA) with BreastScreen WA to obtain and measure their mammographic images. This was possible for 426 Western Australian kConFab participants from 197 families with more than 200 cases of breast cancer. For the current study, we selected families who were known not to carry the BRCA1/2 genes and who had members with available mammograms from Breast Screen Western Australia ($n = 405$ participants from 183 families (plus friends) with 114 cancers).

Ethics approval was obtained from the Western Australian Department of Health Human Research Ethics Committee (#RGS0000002834) and the University of Western Australia Human Research Ethics Office (#RA/4/1/9183).

2.1.2. Busselton Health Study (BHS)

Busselton is a rural, historically stable community ~230 km south of Perth, WA; predominantly of British (Anglo-Saxon) expatriate origin. The BHS is one of the longest-running international epidemiological research programs, with repeated cross-sectional surveys of adults undertaken between 1966 and 2007. The recruitment and data collection of participants from the BHS have previously been described in detail [15]. In 1994/1995, a follow-up survey was conducted of all surviving participants previously surveyed with approximately 5700 individuals attending. High-density single-nucleotide polymorphism (SNP) genotyping data are available for 4671 of these 1994/1995 BHS participants using either an Illumina 660 W or 610 W genome-wide association chip [16]. Other available data include obesity-related markers (measured at time of appointment by a research nurse), reproductive history, and exogenous hormone use. Via the Western Australian Department of Health Data Linkage Branch, we linked all 1994/1995 BHS participants with BreastScreen WA to obtain and measure their mammographic images, and the WA Cancer Registry to obtain all breast cancer diagnoses from 1980 onwards.

Informed consent was granted from all participants in the 1994/1995 survey and ethics was obtained by the University of Western Australia Human Research Ethics Committee (#RA4/1/6694). The current study was approved by the Western Australian Department of Health Human Research Ethics Committee (#RGS0000002801).

2.2. Status Allocation

The participants for both studies were divided into three status groups: cases, unaffected relatives of cases (henceforward relatives) and controls. For the kConFab participants, information collected about each participant's families was used to create pedigrees and assign each participant to a status group. For BHS participants, allocation to a status group was determined using the genetic relatedness matrix (GRM) generated from genome-wide SNP genotype data. BHS participants were assigned either as a case if they had a WA Cancer Registry-confirmed breast cancer, as a relative if they were not a case and the GRM estimated them to have a relatedness of greater than 0.0875 to a case (i.e., proportion of relatedness to a case to capture 1st cousins or greater, where relatedness of 0.125 indicates third-degree relatives and 0.5 indicates first-degree relatives), or finally as a control if they were not related to a case and were not a case themselves [17]. Within both studies, all relatives were breast cancer free and if a case was related to another case, they were assigned to the case group; however, their relationship to the case was captured through the inclusion of the genetic relatedness matrix in all analyses.

2.3. Mammogram Selection

Cranio-caudal film mammograms were retrieved from BreastScreen WA, digitized and measured by author JS using the Cumulus software (Sunnybrook Health Sciences Centre, Toronto, ON, Canada). Where mammograms from multiple screening visits were available, the pre-diagnosis mammogram closest to the diagnosis date was selected for cases. For relatives and controls, the mammogram closest to the epidemiological data

collection date was selected. If a relative or control had no epidemiological data, the earliest mammogram available was selected.

The MBD measurements included absolute dense area and percentage dense area. Percentage dense area could not be measured for six participant's mammograms from the kConFab study as the images had bad edges and thus total breast size could not be measured. Mammograms were measured twice for 10% of participants to assess reliability. Intraclass correlation coefficients for absolute dense area and percentage dense area were 0.98 in the kConFab Study and 0.99 in the BHS.

2.4. Data Analysis

Analyses were conducted in R version 3.6.3 [18] and Genome-wide Complex Trait (GCTA) [19]. Medians and interquartile ranges (IQR) were used to describe the study populations. Participants missing body mass index (BMI) information were excluded from analyses. Descriptive analyses showed those missing BMI were similarly aged at the time of their mammogram but were less dense than those who had reported BMI. The GRMs were estimated using the pedigrees deduced from the family relationship data collected during interviews for kConFab and from results of the genome-wide SNP data for BHS. The latter was estimated using Linkage Disequilibrium Adjusted Kinships (LDAK) software [20] as described previously [16]. Relatedness was set to zero for those with relatedness below 0.05 in the BHS GRM as this has been shown to reduce potential bias in heritability and genetic correlation estimates from using both closely and distantly related individuals [21]. All regression analyses were adjusted for age at mammogram, BMI, time between mammogram and when BMI was reported, and the GRM. In addition, all BHS analyses included number of live births and a menopause status variable (defined as 1 if the woman had reported her periods had stopped and was not taking hormone replacement therapy, and 0 otherwise).

2.4.1. Test of Criterion (i)—The Endophenotype Is Associated with Illness in the Population

We assessed the association of absolute dense area and percentage breast density with breast cancer by testing for differences between cases and combined relatives and controls. Generalised linear mixed models, with a binomial distribution with logit link function (R package GMMAT [22]) was used to compare the status groups adjusting for age at mammogram, BMI, time between mammogram and BMI collection (plus number of live births and menopause status for BHS), and the GRM as a random effect, using Wald tests. For some models, adjustment for the GRM resulted in the variance estimate being on the boundary of the parameter space observed. After additional testing involving the removal of highly influential observations we determined that the estimates change minimally when the GRM was removed and so have reported the model estimates which do not adjust for the GRM.

2.4.2. Test of Criterion (ii)—The Endophenotype Is Heritable

Narrow-sense heritability is the proportion of the variability of the phenotype that can be attributed to additive genetic variation. Heritability estimation for absolute dense area (h_{DA}) and percentage dense area (h_{PDA}) was performed using restricted maximum likelihood analysis using Genome-wide Complex Trait Analysis (GCTA) software [19]. Square root transformations were used for both MBD measurements to normalise distributions and analyses were adjusted for age at mammogram, BMI, time between mammogram and BMI collection, and the GRM plus number of live births and menopause status for BHS.

2.4.3. Test of Criteria (iii)—Within Families, Endophenotype and Illness Co-Segregate and (iv)—The Endophenotype Found in Affected Family Members Is Found in Non-Affected Family Members at a Higher Rate than in the General Population

Linear mixed models and estimated marginal means were used to assess whether the density measures differed across the three status groups: cases, relatives and controls.

This is akin to testing genetic correlation, and in the presence of non-traditional family structures within these samples, comparisons of means between the three status groups was considered the most statistically powerful method to assess this. A square root transformation was applied to the MBD measures to normalise the distributions and models were adjusted for age at mammogram, BMI, time between mammogram and BMI collection, and the GRM, plus number of live births and menopause status for BHS. As before, estimates on the boundary space were assessed and are presented without GRM adjustment. A meta-analysis of the results from both studies was conducted using the R library ‘meta’.

3. Results

The final number of participants within the kConFab study was 323 (88 cases, 179 relatives and 56 controls) and 1587 (92 cases, 72 relatives and 1423 controls) for the BHS. Descriptive statistics are shown in Table 1. For both study populations, the median age for cases was higher than in controls and relatives. BMI was similar across the groups with relatives having slightly lower median BMIs when compared to cases and controls for both studies. The mean relatedness of the relatives was 0.33 within the kConFab study, and 0.35 within the BHS, with approximately half of relatives in both studies either sister pairs or mother/daughter pairs.

Table 1. Population characteristics for the two studies kConFab and BHS.

Characteristic	kConFab				BHS			
	Total (n = 323)	Controls (n = 56, 17.3%)	Relatives (n = 179, 55.4%)	Cases (n = 88, 27.2%)	Total (n = 1587)	Controls (n = 1423, 89.7%)	Relatives (n = 72, 4.5%)	Cases (n = 92, 5.8%)
Median age (years) at mammogram (IQR)	51.0 (13.0)	51.0 (8.5)	50.0 (14.5)	56.0 (12.3)	52.6 (15.9)	52.8 (15.7)	49.5 (11.9)	54.0 (16.9)
Median BMI at interview (IQR) ^a	26.0 (7.5)	26.7 (6.4)	24.9 (7.6)	26.4 (7.0)	25.1 (6.0)	25.1 (6.0)	24.8 (6.0)	25.9 (5.6)
Median time (days) between mammogram and interview (IQR) ^a	−1084 (2644)	−2990 (3452)	−658 (2316)	−1200 (2230)	305 (3282)	331 (3278)	1043 (3371)	24.5 (2207)
Median dense area cm ² (IQR)	17.6 (26.3)	15.2 (28.6)	15.4 (25.3)	22.0 (26.3)	15.3 (24.4)	14.7 (23.8)	22.0 (28.6)	19.7 (24.1)
Median percentage dense area ^b % (IQR)	16.6 (26.1)	16.8 (33.8)	14.8 (24.4)	20.9 (25.8)	15.0 (28.0)	14.0 (27.5)	23.0 (32.5)	19.0 (26.0)

^a Negative values denote BMI was taken after mammogram. ^b A total of six missing from relatives of cases in kConFab due to poor mammogram quality. Abbreviations: BHS: Busselton Health Study, BMI: body mass index, IQR: interquartile range.

3.1. Test of Criterion (i)—MBD Is Significantly Associated with Breast Cancer

Table 2 shows the associations between the MBD measures and breast cancer risk for both the kConFab and BHS studies. Within the kConFab study, an increase in dense area of 1 cm² was associated with an increased odds of breast cancer of 1.015 (95% CI = 1.002, 1.029), compared to women with no breast cancer (relatives and controls). A smaller effect size (OR = 1.009; 95% CI = 0.998, 1.019) was observed within the BHS. Similarly, an increase of 1% in percent dense area was associated with an increased odds of breast cancer within kConFab (OR = 1.019, 95% CI = 1.002, 1.037), with a smaller increase in the BHS (OR = 1.011, 95% CI = 0.998, 1.025).

Table 3 shows the linear regression estimates of the association between MBD measures and case-relative-control status for both kConFab and BHS studies. Within the kConFab study, cases had higher absolute dense area ($\beta = 6.93$, 95% CI = 0.174, 13.68) and percentage dense area ($\beta = 0.818$, 95% CI = 0.177, 1.46) than controls. Similarly, within the BHS, cases also had higher absolute dense area ($\beta = 0.429$, 95% CI = 0.009, 0.850) and percentage dense area ($\beta = 0.391$, 95% CI = 0.007, 0.790). When compared to controls, however, the effect sizes were slightly smaller. These higher MBD measure estimates among cases compared to controls remained after meta-analysing both studies results.

Table 2. Logistic regression estimates (odds ratios (OR) and 95% confidence intervals (CI)) showing the associations between MBD measures and breast cancer risk for both the kConFab and BHS studies.

	kConFab ¹						BHS ²					
	Cases (n = 88)/ Relatives + Controls (n = 235 ³)		Cases (n = 88)/ Relatives (n = 179 ³)		Cases (n = 88)/ Controls (n = 56)		Cases (n = 92)/ Relatives + Controls (n = 1495)		Cases (n = 92)/ Relatives (n = 72)		Cases (n = 92)/ Controls (n = 1423)	
	OR (95% CI)	p-Value ⁴	OR (95% CI)	p-Value ⁴	OR (95%CI)	p-Value ⁴	OR (95% CI)	p-Value ⁴	OR (95% CI)	p-Value ⁴	OR (95%CI)	p-Value ⁴
Dense area (cm ²)	1.015 (1.002, 1.029)	0.029	1.015 (1.000, 1.030)	0.043	1.023 (0.9998, 1.046)	0.050	1.009 (0.998, 1.019)	0.094	1.004 (0.984, 1.025)	0.692	1.009 (0.998, 1.019)	0.101
Percentage dense area (%)	1.019 (1.002, 1.037)	0.030	1.020 (1.002, 1.039)	0.031	1.022 (0.996, 1.049)	0.098	1.011 (0.998, 1.025)	0.102	0.999 (0.978, 1.021)	0.934	1.012 (0.998, 1.025)	0.097

¹ kConFab models adjusted for age, BMI, time between BMI measurement and mammogram and the GRM. ² BHS models adjusted for age, BMI and time between BMI measurement and mammogram, number of live births and menopause status. Additionally adjustment for the GRM resulted in the variance estimate being on the boundary of the parameter space observed. After additional testing conducted involving removal of highly influential observations we determined that the estimates change minimally when the GRM is removed and so have reported the models which do not adjust for the GRM. ³ 6 relatives are missing percentage dense area measurements due to poor mammogram quality. ⁴ p value calculated using Wald test. Abbreviations: BHS: Busselton Health Study, BMI: body mass index, OR: odds ratio, and CI: confidence interval.

Table 3. Linear regression estimates (β) of the associations between the MBD measures and case-relative-control status.

	$\beta_{\text{cases vs. relatives}}$ (CI)	$\beta_{\text{cases vs. controls}}$ (CI)	$\beta_{\text{relatives vs. controls}}$ (CI)
kConFab			
Dense area (cm ²)	0.649 (0.165, 1.13)	6.93 (0.174, 13.68) ¹	0.253 (−0.404, 0.909)
Percentage dense area (%)	0.593 (0.155, 1.03)	0.818 (0.177, 1.46)	0.106 (−0.503, 0.716)
BHS			
Dense area (cm ²)	0.197 (−0.428, 0.822)	0.429 (0.009, 0.850)	0.106 (−0.394, 0.605)
Percentage dense area (%)	0.186 (−0.508, 0.685)	0.391 (−0.007, 0.790)	0.201 (−0.269, 0.672)
Meta-analysed ²			
Dense area (cm ²)	0.478 (0.098, 0.859)	0.596 (0.240, 0.951)	0.161 (−0.236, 0.556)
Percentage dense area (%)	0.413 (0.060, 0.7767)	0.511 (0.173, 0.848)	0.162 (−0.210, 0.534)

Bold type indicates statistical significance at $\alpha < 0.05$. Dense area and percentage dense area were square root transformed and all models were adjusted for age, BMI, time between BMI measurement and mammogram and the GRM unless otherwise stated. In addition, BHS models included adjustment for number of live births and menopause status. ¹ The variance estimate was on the boundary of the parameter space observed so was unable to fit model with square root transformed outcome which included the GRM. This is the effect estimate with dense area not transformed. The effect estimate with dense area square root transformed and no GRM included is: 1.022 (0.327, 1.717) (SE: 0.351). ² Meta analysis results presented are from a fixed effect meta analysis (all tests of heterogeneity $p > 0.10$). Abbreviations: BHS: Busselton Health Study, BMI: body mass index, and CI: confidence interval.

3.2. Test of Criterion (ii)—MBD Is Heritable

The estimated heritability of absolute dense area and percentage dense area were both significant in both studies and were higher in kConFab ($h_{DA} = 0.587$, $p_{DA} = 0.002$; $h_{PDA} = 0.658$, $p_{PDA} = 0.005$) than BHS ($h_{DA} = 0.398$, $p_{DA} < 0.001$; $h_{PDA} = 0.312$, $p_{PDA} < 0.001$).

3.3. Test of Criterion (iii)—Within Families, Endophenotype and Illness Co-Segregate and (iv)—The Endophenotype Found in Affected Family Members Is Found in Non-Affected Family Members at a Higher Rate than in the General Population

From Table 3, within the kConFab study, absolute dense area ($\beta = 0.649$, 95% CI = 0.165, 1.13) and percentage dense area ($\beta = 0.593$, 95% CI = 0.155, 1.03) were higher among cases than relatives. Results from the BHS showed a similar pattern. However, evidence was weaker for both absolute dense area ($\beta = 0.197$, 95% CI = −0.428, 0.822) and percentage dense area ($\beta = 0.186$, 95% CI = −0.508, 0.685). Meta-analyses of the results from two studies found some evidence of higher absolute dense area ($\beta = 0.478$, 95% CI = 0.098, 0.859) and percentage dense area ($\beta = 0.413$, 95% CI = 0.060, 0.777) among cases when compared to relatives.

For both studies, less evidence was seen for higher MBD measures among relatives compared to controls. The meta-analysis of both study results found slightly higher absolute dense area ($\beta = 0.161$, 95% CI = −0.236, 0.556) and percentage dense area ($\beta = 0.162$,

95% CI = $-0.210, 0.534$) among relatives compared to controls. However, the evidence was weak.

Figures 1 and 2 show the estimated marginal means for case-relative-control status for each MBD measure across both studies. The estimated marginal mean of each MBD measure increases as the status changes from control to relative to case in both studies. However, the confidence intervals are overlapping.

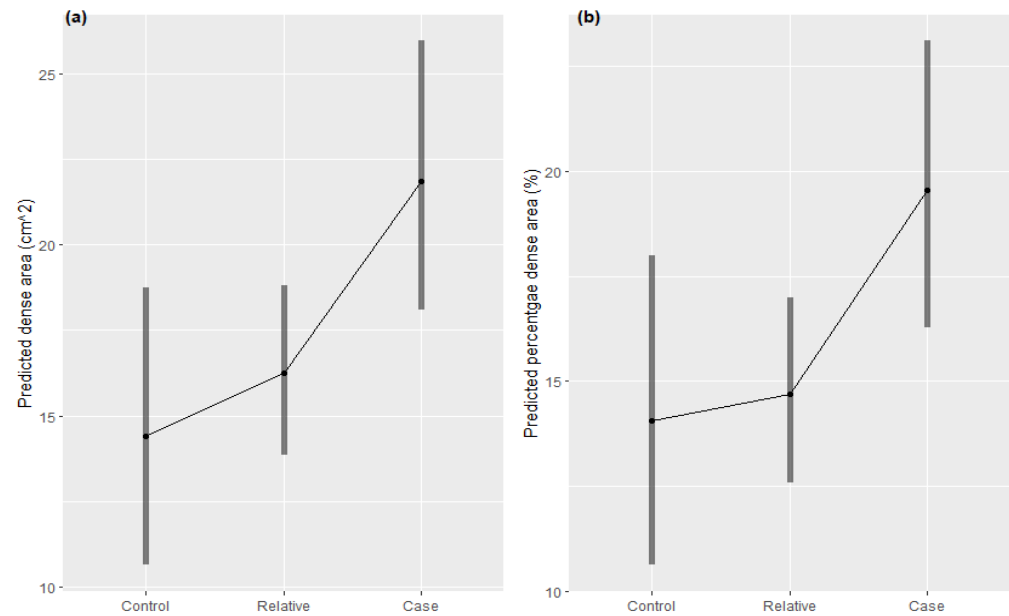


Figure 1. Estimated marginal means for dense area (a) and percentage dense area (b) adjusted for age, BMI, time between BMI measurement and mammogram and the GRM by case, relative and control status for kConFab sample.

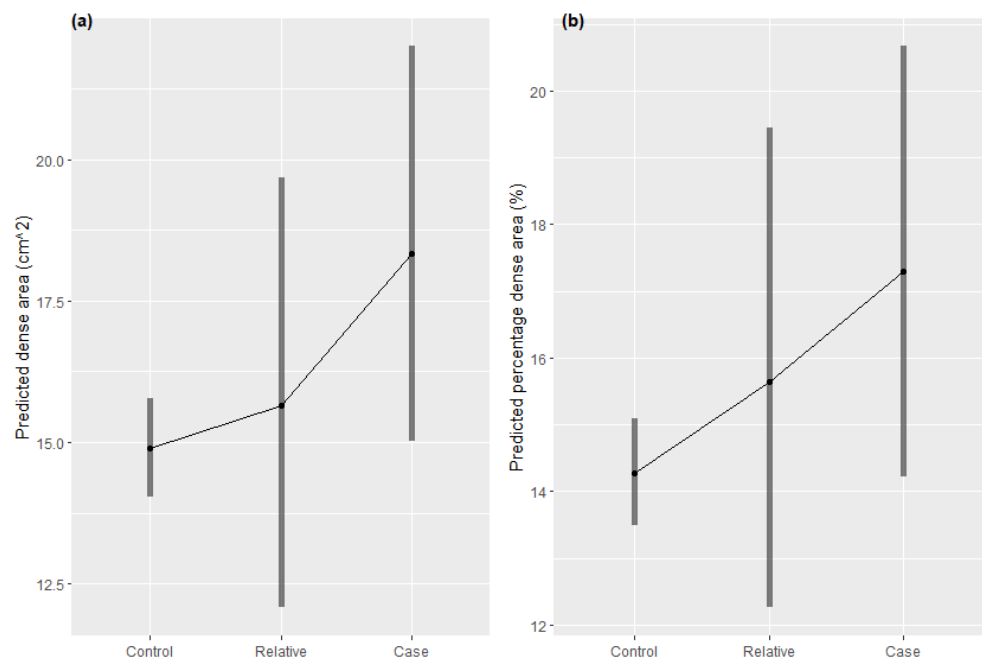


Figure 2. Estimated marginal means for dense area (a) and percentage dense area (b) adjusted for age, BMI, time between BMI measurement and mammogram, number of live births, menopause status, and the GRM by case, relative and control status for Busselton Health Study sample.

4. Discussion

We have systematically assessed, for the first time, whether breast density is an endophenotype for breast cancer using five endophenotype criteria. The results of this study show that mammographic breast density (absolute and percent dense area measures) meets most of the criteria for being an endophenotype for breast cancer. Using two independent samples with a combined sample size of 1910, we provide evidence that (i) MBD is associated with breast cancer, (ii) MBD measures are heritable, (iii) within families, MBD and breast cancer co-segregate, and (iv) MBD measures within relatives of breast cancer cases are higher than in the general population.

4.1. Criteria 1: MBD Is Associated with Breast Cancer Risk

We found that both MBD—dense area and percent dense area—were positively associated with breast cancer risk, independent of age and BMI. These findings replicate well-established knowledge that MBD is an independent risk factor for breast cancer risk [4,5], and our estimates are consistent with earlier studies investigating per unit increases in absolute and percent dense area [23,24].

4.2. Criteria 2: MBD Is Heritable

Consistent with the literature, we estimated the heritability of dense area and percent dense area to be 0.59 and 0.68, respectively, within the kConFab study. We have previously reported heritability estimates between 0.6 and 0.67 for percent dense area [6] and 0.65 for absolute dense area [7] within Australian and North American twin studies. The heritability estimates within the BHS were smaller ($h^2 = 0.39$, $h^2 = 0.30$, respectively). Heritability estimates within kConFab were calculated using known familial relationships, while estimates within the BHS were calculated using the SNP-based relatedness estimates (and therefore represent the variation due only to the SNPs). The lower estimates within the BHS compared to kConFab are therefore likely due to the fact that total heritability (due to all genetic variation) is assessed in the kConFab, whereas in the BHS the heritability estimate only reflects genetic variation captured by SNPs.

4.3. Criteria 3 and 4: Breast Cancer Segregates with Breast Density within Families, and Non-Affected Family Members Have an Intermediate (between Cases and Unrelated Controls) Breast Density

In our meta-analysis between the kConFab and BHS studies, we identified higher breast density in breast cancer cases, intermediate density in relatives of cases, and lower density in controls for both MBD measures. Evidence for these differences was strong comparing cases and relatives but was limited when comparing the relatives and controls. As the relatives of cases within the kConFab study may be more likely to have genetic variants predisposing them to breast cancer, the estimates involving relatives may be subject to selection bias. The differences in estimates involving relatives between kConFab and BHS might be due to both breast cancer and MBD having a greater genetic contribution (and lower environmental contribution) in the kConFab sample. The pooled marginal mean estimates were not significantly different (potentially due to lack of power), but did show an increase across each category, in line with cases having higher density, relatives with intermediate density and controls with the lowest density. These associations suggest that there is a genetic component in common between breast cancer and MBD measures.

This study has a number of strengths. First, we had access to genetic, mammography, and breast cancer case status from two epidemiological studies, representing 1910 women, consisting of 180 cases, 251 relatives of cases, and 1479 controls. The addition of family members within these studies allowed us to test the criteria among relatives of controls which is often not available in population-based or case-control cohorts. Second, the use of two studies also allowed us to validate our findings in an independent cohort, which is integral to genetic studies of this type. Third, we assessed the endophenotype criteria using two measures of MBD, percent and absolute dense area, and observed consistent results

for both phenotypes. This is consistent with strong genetic correlation between percent and absolute dense area (kConFab: $\rho_G = 0.938$, p -value = 0.008 and BHS: $\rho_G = 0.946$, p -value < 0.001).

However, this study has some limitations. First, we did not have *BRCA1/2* status for the BHS, and therefore were unable to exclude these women. However, the prevalence of these mutations in a population-based cohort is low (<1%; [25]), and therefore the proportion of women with these mutations in our population-based study would be small. Second, our study samples consisted mainly of women with European ancestry, and therefore our results may not be generalisable to other ethnic groups. However, previous studies have shown that MBD measures are strongly associated with breast cancer risk across different ethnic groups [26,27]. Finally, BMI measures for some participants were not available, and as BMI is a critical MDB covariate, these participants had to be excluded from analyses.

5. Conclusions

In summary, we have shown through a comprehensive assessment of endophenotype criteria that two measures of breast density—dense area and percent dense area—are endophenotypes for breast cancer. As MDB is genetically correlated with breast cancer and can be measured on any woman who has had a mammogram (regardless of breast cancer status), genetic investigations of MDB may potentially identify novel risk variants for breast cancer and help identify novel breast cancer mechanisms. Improved understanding of these genetic associations could also inform future research towards tailored screening programs and prevention strategies.

Author Contributions: Conceptualization, N.M. and G.C.; methodology, G.C.; software, G.C.; validation, E.D. and G.C.; formal analysis, E.D.; investigation, G.C. and E.D.; resources, J.S., E.K.M. and C.S.; data curation, J.S. and G.C.; writing—original draft preparation, E.D., J.S. and G.C.; writing—review and editing, N.M., C.S. and E.K.M.; visualization, E.D.; supervision, J.S. and G.C.; project administration, J.S.; funding acquisition, J.S., C.S. and E.K.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Breast Cancer Research Centre Western Australia and the Royal Perth Hospital Research Foundation. J.S. is a National Breast Cancer Foundation Research Fellow.

Institutional Review Board Statement: The kConFab study was approved by the Human Research Ethics Committee of the Western Australian Department of Health (RGS0000002834) and the University of Western Australia (#RA/4/1/9183). The Busselton Health Study was approved by the Human Research Ethics Committee of the Western Australian Department of Health (RGS0000002801) and the University of Western Australia (#RA4/1/6694).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the original studies.

Data Availability Statement: The datasets generated and analysed during the current study are not publically available but can be made available upon request (and pending approval) from Western Australia's Department of Health Data Linkage Branch and BreastScreen Western Australia.

Acknowledgments: The authors would like to acknowledge BreastScreen Western Australia and the staff at the Western Australian Data Linkage Branch and the Western Australian Cancer Registry. The Busselton Health Study (BHS) acknowledges the generous support for the 1994/1995 Busselton follow-up studies from HealthWay, the Department of Health, PathWest Laboratory Medicine of WA, The Great Wine Estates of the Margaret River region of Western Australia, the Busselton community volunteers who assisted with data collection, and the study participants from the Shire of Busselton. We wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (which has received funding from the NHMRC, the National Breast Cancer Foundation, Cancer Australia, and the National Institute of Health (USA)) for their contributions to this resource, and the many families who contribute to kConFab. kConFab is supported by a grant from the National Breast Cancer Foundation, and previously by the National Health and Medical Research Council (NHMRC), the

Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. National Breast Cancer Foundation. About Breast Cancer. Available online: <https://www.nationalbreastcancer.org/about-breast-cancer/> (accessed on 9 June 2021).
2. Michailidou, K.; Hall, P.; Gonzalez-Neira, A.; Ghoussaini, M.; Dennis, J.; Milne, R.L.; Schmidt, M.K.; Chang-Claude, J.; Bojesen, S.E.; Bolla, M.K.; et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat. Genet.* **2013**, *45*, 353–361. [[CrossRef](#)] [[PubMed](#)]
3. Michailidou, K.; Lindstrom, S.; Dennis, J.; Beesley, J.; Hui, S.; Kar, S.; Lemacon, A.; Soucy, P.; Glubb, D.; Rostamianfar, A.; et al. Association analysis identifies 65 new breast cancer risk loci. *Nature* **2017**, *551*, 92–94. [[CrossRef](#)]
4. Bond-Smith, D.; Stone, J. Methodological challenges and updated findings from a meta-analysis of the association between mammographic density and breast cancer. *Cancer Epidemiol. Biomark. Prev.* **2019**, *28*, 22–31. [[CrossRef](#)]
5. McCormack, V.A.; dos Santos Silva, I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiol. Biomark. Prev.* **2006**, *15*, 1159–1169. [[CrossRef](#)]
6. Boyd, N.F.; Dite, G.S.; Stone, J.; Gunasekara, A.; English, D.R.; McCredie, M.R.; Giles, G.G.; Tritchler, D.; Chiarelli, A.; Yaffe, M.J.; et al. Heritability of mammographic density, a risk factor for breast cancer. *N. Engl. J. Med.* **2002**, *347*, 886–894. [[CrossRef](#)] [[PubMed](#)]
7. Stone, J.; Dite, G.S.; Gunasekara, A.; English, D.R.; McCredie, M.R.; Giles, G.G.; Cawson, J.N.; Hegele, R.A.; Chiarelli, A.M.; Yaffe, M.J.; et al. The heritability of mammographically dense and nondense breast tissue. *Cancer Epidemiol. Biomark. Prev.* **2006**, *15*, 612–617. [[CrossRef](#)] [[PubMed](#)]
8. Sprague, B.L.; Gangnon, R.E.; Burt, V.; Trentham-Dietz, A.; Hampton, J.M.; Wellman, R.D.; Kerlikowske, K.; Miglioretti, D.L. Prevalence of mammographically dense breasts in the United States. *J. Natl. Cancer Inst.* **2014**, *106*, dju255. [[CrossRef](#)] [[PubMed](#)]
9. Cuzick, J.; Warwick, J.; Pinney, E.; Warren, R.M.; Duffy, S.W. Tamoxifen and breast density in women at increased risk of breast cancer. *J. Natl. Cancer Inst.* **2004**, *96*, 621–628. [[CrossRef](#)]
10. Cuzick, J.; Warwick, J.; Pinney, E.; Duffy, S.W.; Cawthorn, S.; Howell, A.; Forbes, J.F.; Warren, R.M. Tamoxifen-induced reduction in mammographic density and breast cancer risk reduction: A nested case-control study. *J. Natl. Cancer Inst.* **2011**, *103*, 744–752. [[CrossRef](#)]
11. Stone, J.; Thompson, D.J.; Dos Santos Silva, I.; Scott, C.; Tamimi, R.M.; Lindstrom, S.; Kraft, P.; Hazra, A.; Li, J.; Eriksson, L.; et al. Novel Associations between Common Breast Cancer Susceptibility Variants and Risk-Predicting Mammographic Density Measures. *Cancer Res.* **2015**, *75*, 2457–2467. [[CrossRef](#)]
12. Gottesman, I.I.; Gould, T.D. The endophenotype concept in psychiatry: Etymology and strategic intentions. *Am. J. Psychiatry* **2003**, *160*, 636–645. [[CrossRef](#)]
13. Kerlikowske, K.; Zhu, W.; Tosteson, A.N.; Sprague, B.L.; Tice, J.A.; Lehman, C.D.; Miglioretti, D.L.; Breast Cancer Surveillance, C. Identifying women with dense breasts at high risk for interval cancer: A cohort study. *Ann. Intern. Med.* **2015**, *162*, 673–681. [[CrossRef](#)]
14. Mann, G.J.; Thorne, H.; Balleine, R.L.; Butow, P.N.; Clarke, C.L.; Edkins, E.; Evans, G.M.; Fereday, S.; Haan, E.; Gattas, M.; et al. Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res.* **2006**, *8*, R12. [[CrossRef](#)] [[PubMed](#)]
15. Gregory, A.T.; Armstrong, R.M.; Grassi, T.D.; Gaut, B.; Van Der Weyden, M.B. On our selection: Australian longitudinal research studies. *Med. J. Aust.* **2008**, *189*, 650–657. [[CrossRef](#)] [[PubMed](#)]
16. Cadby, G.; Melton, P.E.; McCarthy, N.S.; Almeida, M.; Williams-Blangero, S.; Curran, J.E.; VandeBerg, J.L.; Hui, J.; Beilby, J.; Musk, A.W.; et al. Pleiotropy of cardiometabolic syndrome with obesity-related anthropometric traits determined using empirically derived kinships from the Busselton Health Study. *Hum. Genet.* **2018**, *137*, 45–53. [[CrossRef](#)]
17. Manichaikul, A.; Mychaleckyj, J.C.; Rich, S.S.; Daly, K.; Sale, M.; Chen, W.M. Robust relationship inference in genome-wide association studies. *Bioinformatics* **2010**, *26*, 2867–2873. [[CrossRef](#)] [[PubMed](#)]
18. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2014.
19. Yang, J.; Lee, S.H.; Goddard, M.E.; Visscher, P.M. GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **2011**, *88*, 76–82. [[CrossRef](#)]
20. Speed, D.; Hemani, G.; Johnson, M.R.; Balding, D.J. Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **2012**, *91*, 1011–1021. [[CrossRef](#)] [[PubMed](#)]
21. Zaitlen, N.; Kraft, P.; Patterson, N.; Pasaniuc, B.; Bhatia, G.; Pollack, S.; Price, A.L. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet.* **2013**, *9*, e1003520. [[CrossRef](#)]
22. Chen, H.; Conomos, M.P. GMMAT: Generalized Linear Mixed Model Association Tests. 2020. Available online: <https://cran.r-project.org/web/packages/GMMAT/GMMAT.pdf> (accessed on 9 June 2021).

23. Boyd, N.F.; Byng, J.W.; Jong, R.A.; Fishell, E.K.; Little, L.E.; Miller, A.B.; Lockwood, G.A.; Tritchler, D.L.; Yaffe, M.J. Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian National Breast Screening Study. *J. Natl. Cancer Inst.* **1995**, *87*, 670–675. [[CrossRef](#)] [[PubMed](#)]
24. Byrne, C.; Schairer, C.; Wolfe, J.; Parekh, N.; Salane, M.; Brinton, L.A.; Hoover, R.; Haile, R. Mammographic features and breast cancer risk: Effects with time, age, and menopause status. *J. Natl. Cancer Inst.* **1995**, *87*, 1622–1629. [[CrossRef](#)] [[PubMed](#)]
25. Risch, H.A.; McLaughlin, J.R.; Cole, D.E.; Rosen, B.; Bradley, L.; Fan, I.; Tang, J.; Li, S.; Zhang, S.; Shaw, P.A.; et al. Population BRCA1 and BRCA2 mutation frequencies and cancer penetrances: A kin-cohort study in Ontario, Canada. *J. Natl. Cancer Inst.* **2006**, *98*, 1694–1706. [[CrossRef](#)] [[PubMed](#)]
26. Darcey, E.; Lloyd, R.; Cadby, G.; Pilkington, L.; Redfern, A.; Thompson, S.C.; Saunders, C.; Wylie, E.; Stone, J. The association between mammographic density and breast cancer risk in Western Australian Aboriginal women. *Breast Cancer Res. Treat.* **2019**, *176*, 235–242. [[CrossRef](#)]
27. Maskarinec, G.; Pagano, I.; Lurie, G.; Wilkens, L.R.; Kolonel, L.N. Mammographic density and breast cancer risk: The multiethnic cohort study. *Am. J. Epidemiol.* **2005**, *162*, 743–752. [[CrossRef](#)] [[PubMed](#)]