A new look at breast density and breast cancer risk

Gerco Haars

A new look at breast density and breast cancer risk

Een nieuwe kijk op borst densiteit en het risico op borstkanker
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag
van de rector magnificus, prof. dr. J.C. Stoof, ingevolge het besluit van het
college voor promoties in het openbaar te verdedigen op
dinsdag 9 september 2008 des middags te 12.45 uur

door

Gerco Haars

geboren op 13 januari 1976, te Lopik

Promotoren: Prof.dr. D.E. Grobbee
Prof.dr. P.H.M. Peeters

Co-promotoren: Dr. P.A.H. van Noord
Dr. C.H. van Gils

Manuscripts based on this thesis

Chapter 2.1
Haars G, van Noord PAH, van Gils CH, Peeters PHM, Grobbee DE. Heritable aspects of dysplastic breast glandular tissue (DY). Breast Cancer Res Treat 2004 Sep;87:149-156.

Chapter 2.2
Baksh MF, Haars G, Todd S, Van Noord PAH and Whitehead J. Comparing correlations of continuous observations from two independent populations using a sequential approach. Statist Med 2006; 25:4293–4310

Chapter 3
Haars G, van Noord PA, van Gils CH, Grobbee DE, Peeters PH. Measurements of breast density: no ratio for a ratio. Cancer Epidemiol Biomarkers Prev 2005 Nov;14(11 Pt 1):2634-40.

Chapter 4
Haars G, van Gils CH, Elias SG, Peeters PHM, Grobbee DE, van Noord PAH. The influence of a period of caloric restriction due to the Dutch Famine on breast density.

Chapter 5
Haars G, van Gils CH, Elias SG, Peeters PHM, Grobbee DE, van Noord PAH. The influence of dense and non-dense tissue on breast cancer risk.

Chapter 6
Haars G, Grobbee DE. The next phase in breast density studies.

Contents

1

A brief history of breast density

In 1976 John N Wolfe published a position paper that introduced a classification of breast patterns visible on a mammogram[1]. These patterns are caused by the difference in radio density of the various tissue types present in the breast and are called parenchymal patterns. Wolfe's classification consists of four categories of increasing radio density. The breast pattern is assessed by a radiologist as the percentage of the breast occupied by dense tissue and the appearance of the dense tissue in either prominent duct patterns (PDP) or homogeneous density referred to as dysplasia (DY). Dysplasia was regarded to be non-pathological hyperplasia of the ductal tissue, but this term has been largely discarded in current literature. The four Wolfe classifications are;

N1 (Normal) for breast with little or no dense tissue (<5%)

P1 for limited, dense PDP (5-25%)

P2 for extensive, dense PDP (25-75%)

DY for widespread, homogeneous density (>75%)

In his position paper, Wolfe reported a higher risk of breast cancer associated with higher levels of density. For years, the Wolfe classification was the gold standard in mammogram evaluation and the association of high mammographic density with increased risk of breast cancer has been repeatedly confirmed[2]. A slightly modified classification that focuses more on extent of tissue density and disregards the qualitative terminology of 'PDP' and 'DY' is included in BI-RADS (Breast Imaging Reporting and Data System)[3]. This classification system of the American College of Radiology is not specifically aimed at categorising breast density, but is widely used in routine care and also allows for assessment of breast cancer risk[4, 5].

With the widespread application of automation, computerised techniques were also developed that enabled more objective measurement of the dense tissue of mammograms on a continuous and absolute scale[6-8]. Although these techniques measure amounts of total breast tissue and dense/non-dense tissue, the results were nearly always transformed into percentages again, staying in line with the Wolfe classification and making comparison possible[2, 9]. The most widely used continuous measure that is currently used was developed by Yaffe and Byng[6, 8]. Studies using such a continuous measurement again confirmed the higher risk of breast cancer associated with higher percentages of breast density.

McCormack and Dos Santos Silva recently performed a meta-analysis, including 42 studies[2]. They assessed the risk of women aged 30-75+ with a Wolfe DY score at 1.7-4.0 times the risk of women with an N1 score. For women with breasts with >75% percent density a 3.7-4.6 times higher risk was found, compared to women with <5% percent density.

Pathological studies have established that breast density reflects the stromal and epithelial cells of the breast and that the concentration of cells is higher in more dense tissue[10-12]. Albanes and Winick proposed that the actual number of cells vulnerable to carcinogenesis has a direct reflection on the chance of developing cancer[13]. Trichopoulos and Lipman extended this hypothesis to breast cancer by postulating that the amount of dense tissue reflects the amount of target tissue at risk for breast cancer[14]. These hypotheses are based purely on the stochastical notion that every cell has the chance to step out of line and the

more cells one has, the higher the chance that one actually will. Although this is a very logical, straightforward theory, it has not yet been confirmed or definitively disproved. Percent density would not be an ideal measure to study this hypothesis, as equal percent densities may reflect different amounts of target tissue (absolute dense tissue) due to different breast sizes.

Epidemiological studies have identified numerous (possible) determinants of breast density, of which some have also been related with breast cancer risk. These determinants include, but are not limited to, life events such as parity and the occurrence of the menopause and environmental influences such as nutrition, use of hormone (replacement) therapy (HRT), oral contraceptives and alcohol and tobacco use. The literature on these various determinants is vast, but key publications by Boyd[15], Byrne[16] and Vachon[17] are among the most influential.

In general, breast development during puberty due to proliferation of ducti will result in relatively high amounts of dense breast tissue. During life and aging, on average the amount of dense tissue will decrease. The developmental phase is difficult to study, because it is unethical and risky to subject healthy, young women to x-ray mammography. There is however evidence, that women before the age of 30 already show breasts patterns with a mean percent density over 40% and individual densities up to over 90%[18].

Early cross-sectional studies, although based on group statistics rather than individual changes, have made clear that, on average, the relative density decreases with age, with a slight modification by factors such as BMI and HRT[16, 17, 19, 20]. A first childbirth and going through menopause cause a substantial decrease in density, possibly due to final differentiation of the breast tissue. Further childbirths also lead to some decrease, but less than the first[21]. These effects were recently confirmed in one of the first longitudinal studies[22].

For breast cancer development, endogenous sex-hormones play an important role[23, 24]. Yet a recent overview of studies between hormones and breast density summarises the lack of support for involvement of breast density in this relation[25]. Concomitant influences might obscure the causal pathway between hormones, density and cancer, but it is more likely that hormones and breast density have separate causal relationships with cancer, as proposed by Tamimi et al[26].

The factor that probably has the most complex relation with breast density and cancer is BMI. First inferences of the role of body size were made by De Waard et al., who showed differences in cancer risk with varying weight[27]. Shortly thereafter, the Body Mass Index was introduced and applied in most studies. Higher BMI has since been related to an increased risk of postmenopausal breast cancer and a lower risk of pre-menopausal breast cancer[28-30]. A higher BMI, on the other side, on average leads to an increased amount of fat (non-dense tissue), subsequently to larger breasts and thus to lower relative density.

It is the aim of this thesis to provide further knowledge about determinants of breast density, measures of breast density and factors that modify the relation between breast

Box 1 – Interpretation of heritability

The term 'heritability' implies that it is a direct measure of the genetic similarity between people. This is in fact the way that geneticists use and interpret the result. A gene, chromosome or entire genome is mapped with one or more markers. The correspondence in these markers, measured as the correlation, is then interpreted as the genetic heritability. Whereas this heritability reflects a hard measure (markers), it does not directly indicate whether the shared genes actually illicit the same effect. The expression of the gene or the functionality of the transcribed protein may differ between people or the gene might transcribe a signal protein (such as oestrogen) for which the two people have different receptors. By assessing the genetic heritability within a selected population with a certain phenotype (e.g. a disease), the inference between the phenotype and the genetic heritability can safely be made.

Social biologists however apply 'heritability' directly to the correspondence in phenotype. The correspondence between people is measured in the amount of similarity in for example height, weight, eye colour or IQ as a combined effect of genes, expression and functionality[33, 34]. As this measure includes the actual translation of the genes into a phenotype, it is also subject to non-genetic influences and the results therefore have a different implication and need to be interpreted with caution[35].

Reference List

1. Falconer DS, MacKay TFC. Introduction to Quantative Genetics. 4th Edition ed. New York: Longman Inc.; 1996.
2. Rice TK, Borecki IB. Familial resemblance and heritability. Adv Genet 2001; 42:35-44.
3. van Asselt KM, Kok HS, van der Schouw YT, Peeters PH, Pearson PL, Grobbee DE. Role of genetic analyses in cardiology: part II: heritability estimation for gene searching in multifactorial diseases. Circulation 2006; 113(8):1136-1139.

Box 2 – Sequential analysis methods

Sequential analysis is a mathematical methodology aimed at reducing the sample size necessary to achieve a reliable outcome[36]. This may help conserve valuable resources such as rare blood samples or it may reduce costs and duration of research[37]. On the other side, there is a threat that sequential analyses will need a larger final sample size to obtain a definitive result. In practice, the sample size is reduced in most studies.

The sequential approach applies a series of interim analyses based on the amount of information at hand. At each interval, the measure of effect is set off against the amount of available information. A statistically significant result is obtained when the difference between exposed and unexposed is large enough given the amount of cases so far and the error rates that were set for the chance of a false positive (α) or a false negative result (β). So if all events were occurring in the exposed group, a result would be obtained quickly, whereas it would take more interim analyses if the difference between groups was smaller and events were also occurring in the unexposed group.

Reference List

1. Whitehead J. The Design and Analysis of Sequential Clinical Trials (Revised Second Ed.). Chichester: John Wiley and Sons; 1997.
2. Van der Tweel I, van Noord PA. Sequential analysis of matched dichotomous data from prospective case-control studies. Stat Med 2000; 19(24):3449-3464.

density, breast cancer and determinants of either. The key approach will be to investigating influences of the absolute and relative amounts of breast density and the differences between these influences. This segregation of influences of the established relative density and its tissue components may give additional insight into the role and value of breast density in the aetiology of breast cancer and may provide new approaches to investigate the association between breast density and breast cancer.

In Chapter 2.1 we address heritability of breast patterns. We set out to estimate the heritability of breast patterns as an indication of the extent to which breast density is determined by in-born factors. The calculated heritability is an estimate of the variance in breast density between women that can be explained by differences in heritable factors (see Box 1). The most often used approach to calculate heritability is to measure the correlation between genetically identical persons (monozygotic twins) and then to compare the results with the correlation in people that are only 50% identical on average (dizygotic twins or siblings). For our study we used monozygotic twins, dizygotic twins and sisters identified within the DOM ('Diagnostisch Onderzoek Mammacarcinoom') cohorts[31, 32].

In Chapter 2.2 we used the study-population from which we estimated the heritability (chapter 2.1) to study the applicability of sequential analyses. The aim of this study is to further develop this methodology as it offers great potential to reduce sample sizes (see Box 2). In addition, by using the continuous, absolute outcome measurement, we wanted to confirm and extend the results from chapter 2.1.

In Chapter 3 we assess the effect of using percent density, the established relative score, as an outcome measure, as opposed the 'new' absolute component measures. We perform an analysis of the relation of some known determinants of breast cancer and/or relative density with measures of total breast size, amount of dense tissue and amount of non-dense tissue. The associations are compared to associations of the same determinants with relative density in a population of 418 postmenopausal women.

In Chapter 4 we study the effect of the Dutch famine on breast patterns. The Dutch famine in the winter of 1944-45 was a terrible event for the people involved, but has provided a unique scientific situation. The severe under-nutrition that a lot of people suffered can be regarded as a 'natural experiment'. Early developmental influences have long been proposed as determinants of later cancer risk and Elias et al. showed that the short but severe period of caloric restriction directly led to an increased risk of breast cancer decades later[38]. Elias also found indications that sex-hormone levels and the insulin-like-growth factor system might be influenced[39, 40]. Both these systems have also been associated with breast density. The aim of this study was to investigate the role of breast density in this pathway. The

influence of the famine on breast density is studied in a sample of 1035 women from the cohorts of Elias et al.

In Chapter 5 we analyse relations between the relative and absolute measurements of breast parenchymal patterns and the risk of breast cancer. We aim to distinguish the independent roles of the components making up relative density and to assess the effect of the combination of dense tissue and the surrounding non-dense (fat) tissue. In a case-cohort study, levels of the various breast density measures in all (361) newly developed breast cancer cases are compared with the levels in a random sample of 600 women.

Chapter 6 is a critical reappraisal of our results associated with absolute measurements and other recent findings in the field of breast density, while Chapter 7 is a summary of the complete thesis.

Reference List

1. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976; 37(5):2486-2492.
2. McCormack VA, dos SS, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006; 15(6):1159-1169.
3. American College of Radiology. Breast Imaging reporting and data system, breast imaging atlas. 4th ed. Reston, VA: American College of Radiology, 2003. 2008.
4. Kerlikowske K, Ichikawa L, Miglioretti DL et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst 2007; 99(5):386-395.
5. Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004; 13(5):715-722.
6. Byng JW, Yaffe MJ, Lockwood GA, Little LE, Tritchler DL, Boyd NF. Automated analysis of mammographic densities and breast carcinoma risk. Cancer 1997; 80(1):66-74.
7. Highnam R, Jeffreys M, McCormack V, Warren R, Davey SG, Brady M. Comparing measurements of breast density. Phys Med Biol 2007; 52(19):5881-5895.
8. Yaffe MJ, Boyd NF, Byng JW et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998; 7 Suppl 1:S47-S55.
9. Boyd NF, Guo H, Martin LJ et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3):227-236.
10. Boyd NF, Jensen HM, Cooke G, Han HL. Relationship between mammographic and histological risk factors for breast cancer. J Natl Cancer Inst 1992; 84(15):1170-1179.
11. Bright RA, Morrison AS, Brisson J et al. Relationship between mammographic and histologic features of breast tissue in women with benign biopsies. Cancer 1988; 61(2):266-271.
12. Hawes D, Downey S, Pearce CL et al. Dense breast stromal tissue shows greatly increased concentration of breast epithelium but no increase in its proliferative activity. Breast Cancer Res 2006; 8(2):R24.
13. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst 1988; 80(10):772-774.
14. Trichopoulos D, Lipman RD. Mammary gland mass and breast cancer risk. Epidemiology 1992; 3(6):523-526.
15. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer 1998; 78(9):1233-1238.
16. Byrne C, Schairer C, Wolfe J et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995; 87(21):1622-1629.
17. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control 2000; 11(7):653-662.
18. Li T, Sun L, Miller N et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. Cancer Epidemiol Biomarkers Prev 2005; 14(2):343-349.
19. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. Cancer Epidemiol Biomarkers Prev 1998; 7(12):1133-1144.
20. Wolfe JN. Breast parenchymal patterns and their changes with age. Radiology 1976; 121(3 Pt. 1):545-552.
21. de Waard F, Rombach JJ, Collette HJ, Slotboom B. Breast cancer risk associated with reproductive factors and breast parenchymal patterns. J Natl Cancer Inst 1984; 72(6):1277-1282.
22. Kelemen LE, Pankratz VS, Sellers TA et al. Age-specific Trends in Mammographic Density: The Minnesota Breast Cancer Family Study. Am J Epidemiol 2008; 167(9):1027-1036.
23. Key TJ, Appleby PN, Reeves GK et al. Body mass index, serum sex hormones, and breast cancer risk in postmenopausal women. J Natl Cancer Inst 2003; 95(16):1218-1226.
24. Onland-Moret NC, Kaaks R, van Noord PA et al. Urinary endogenous sex hormone levels and the risk of postmenopausal breast cancer. Br J Cancer 2003; 88(9):1394-1399.

25. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. Breast Cancer Res 2008; 10(1):201.
26. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 2007; 99(15):1178-1187.
27. de Waard F. Breast cancer incidence and nutritional status with particular reference to body weight and height. Cancer Res 1975; 35(11 Pt. 2):3351-3356.
28. Lahmann PH, Hoffmann K, Allen N et al. Body size and breast cancer risk: findings from the European Prospective Investigation into Cancer And Nutrition (EPIC). Int J Cancer 2004; 111(5):762-771.
29. Sonnenschein E, Toniolo P, Terry MB et al. Body fat distribution and obesity in pre- and postmenopausal breast cancer. Int J Epidemiol 1999; 28(6):1026-1031.
30. Ursin G, Longnecker MP, Haile RW, Greenland S. A meta-analysis of body mass index and risk of premenopausal breast cancer. Epidemiology 1995; 6(2):137-141.
31. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984; 1(8388):1224-1226.
32. de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 1984; 37(1):1-44.
33. Falconer DS, MacKay TFC. Introduction to Quantative Genetics. 4th Edition ed. New York: Longman Inc.; 1996.
34. Rice TK, Borecki IB. Familial resemblance and heritability. Adv Genet 2001; 42:35-44.
35. van Asselt KM, Kok HS, van der Schouw YT, Peeters PH, Pearson PL, Grobbee DE. Role of genetic analyses in cardiology: part II: heritability estimation for gene searching in multifactorial diseases. Circulation 2006; 113(8):1136-1139.
36. Whitehead J. The Design and Analysis of Sequential Clinical Trials (Revised Second Ed.). Chichester: John Wiley and Sons; 1997.
37. Van der Tweel I, van Noord PA. Sequential analysis of matched dichotomous data from prospective case-control studies. Stat Med 2000; 19(24):3449-3464.
38. Elias SG, Peeters PH, Grobbee DE, van Noord PA. Breast cancer risk after caloric restriction during the 1944-1945 Dutch famine. J Natl Cancer Inst 2004; 96(7):539-546.
39. Elias SG, Keinan-Boker L, Peeters PH et al. Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. Int J Cancer 2004; 108(4):628-630.
40. Elias SG, Onland-Moret NC, Peeters PH et al. Urinary endogenous sex hormone levels in postmenopausal women after caloric restriction in young adulthood. Br J Cancer 2004; 90(1):115-117.

2

Heritable aspects of dysplastic breast glandular tissue (DY)

Gerco Haars

Paul AH van Noord

Carla H van Gils

Petra HM Peeters

Diederick E Grobbee

2.1

Abstract

Breast parenchymal patterns, as visible on mammograms, are determined by the relative amount of radio dense, glandular dysplastic tissue (DY). High percentages of DY are related to higher breast cancer risk. Previous studies reported heritable influences on DY of 32–67%, depending on the family relationship that was studied. We assessed heritability in 466 sister-, 25 dizygotic twin- and 26 monozygotic twin-pairs participating in a population-based breast cancer screening program; the DOM project (Diagnostic Investigation Mamma Carcinoma).

The heritability was estimated for non-twin sisters, dizygotic and monozygotic twins separately by computing correlations between siblings from the dichotomous DY-score (high risk versus low risk). This was done using methods based on the number of shared genes per sib type. Heritability estimates were 38, 34 and 88% for sisters, dizygotic twins and monozygotic twins respectively. Heritability estimates from models that combine monozygotic twins with dizygotic twins or sisters indicated that dominant gene effects, genetic interactions or gene–environment effects could be involved. Parity appeared to have an effect on the heritable influence with estimates ranging from 90% in sisters that were both nulliparous, to 2% in sister pairs discordant for nulliparity. These results indicate a substantial genetic influence on DY, but with a possible modifying ability of other factors, such as parity.

Introduction

Breast dysplasia (DY) has been the topic of much research over the last decades. It has been related to proliferated connective and stromal tissue and[1, 2] the relative amount of DY in the breast determines the breast parenchymal pattern, as visible on a mammogram because of the relatively high radio-density of DY[3]. Since the first classification of parenchymal patterns by Wolfe[4], DY has been identified as a

risk factor for breast cancer and has been proposed as an intermediate endpoint for breast cancer studies[5, 6]. Several factors have been established to affect breast density, including age, weight, body mass index (BMI), parity, age at fist childbirth, menopausal status, causation of menopause (natural versus induced), use of hormone replacement therapy (HRT) and nutrition[7-13].

Nonetheless, these factors explain only a limited fraction of the variance in DY among women and much is still unclear about the origin of DY, the underlying mechanisms by which it is influenced and the direct pathway by which DY affects breast cancer risk. Genetic factors may also play a role, as was previously reported by Pankow et al.[14] in a study in breast cancer cases and their relatives. They estimated a heritability, which is defined as the amount of variance in a trait that can be ascribed to genetic variation, of DY in sisters of 32–54%. Recently, Boyd et al.[15] reported the heritability of breast dysplasia in a population of recruited monozygotic and dizygotic twins from Canada and Australia. This study estimated the heritability of DY at 60–67%. The aim of the present study was to determine the heritable influence on mammographic high risk patterns from a general population. This population was represented in the cohort of the DOM (an acronym translated as 'Diagnostic Investigation Mammacancer') breast cancer screening program from Utrecht, The Netherlands[16]. Contrary to the study by Boyd, in which self-selected twin pairs were enrolled in the study and in contrast to Pankow, who specifically selected relatives of breast cancer cases, we included both sisters and twin pairs from a population-based cohort.

Materials and methods

Subjects

The DOM project was initiated in Utrecht, The Netherlands, in 1974 as a population-based breast cancer screening program with a scientific orientation[16]. Between 1974 and 1986 55,519 women, aged 39–68 at recruitment, attended one or more screening rounds. At each round, mammograms were made of both breasts, anthropometric indices were measured and an extensive questionnaire was administered. Through these questionnaires, data have been collected on, among others, demography, family history of breast cancer and reproductive and menstrual history. In the course of the project, women were invited to the screening in 4 sub-cohorts (DOM I–IV). Between these sub-cohorts, age at recruitment differed and varying questionnaires were used to collect baseline characteristics.

This resulted in a varying availability of information and subsequently missing data for some characteristics for some participants. We identified sister- and twin-pairs, of which both siblings participated in DOM, using probabilistic matching[17]. All data used for matching had

been collected at the time of recruitment in DOM. Sister pairs were identified by matching participants for family-name, mothers date of birth and number of siblings. Additionally, the resulting linked participants were checked for agreement in other available family-based variables, such as age of death of mother, number of sisters and/or number of brothers, date of birth within natural range and date of birth in accordance with birth-rank. Sister pairs that were fully concordant for the matching variables and not conflicting in any other family-based variable (for which a value was available for both siblings) were selected. This resulted in a total of 466 sister pairs that were considered eligible for analyses.

Twin pairs were identified by first selecting women for self reported twin-ship and self reported type of twin-ship. The women in the resulting group were subsequently matched for family-name and date of birth. Additionally, the remaining linked twin pairs were checked for accordance in other family-based variables such as date of birth of mother, number of sisters and/or brothers and age of death of mother. Twin pairs that were fully concordant for the matching variables and not conflicting in any other family-based variable were selected, which resulted in a total of 26 dizygotic twin pairs and 25 monozygotic twin pairs that were considered eligible for analyses.

Scoring of parenchymal patterns

As part of the original screening examination, mammograms were evaluated by a trained radiologist and the parenchymal pattern of each breast was classified[18]. All siblings were evaluated completely independent since no attention was paid to familial relations at this time. During the first screening rounds, the evaluation was done according to the Wolfe-classification, which distinguishes four categories; N for normal, non-dysplastic breasts (<25% DY), P1 for low-dysplastic breast (25–50% DY), P2 for breast with prominent ductal patterns and medium-dysplastic breast (50–75% DY) and DY for highly dysplastic breast (>75% DY)[4]. In later rounds, this scoring was reduced to a two level score for low and high risk patterns, which relate to combinations of N/P1 and P2/DY respectively. Therefore, the classification scores in our study were all collapsed into a dichotomous score of DY- versus DY+, which also relate to N/P1 and P2/DY respectively. DY values were available for all the identified siblings.

Data analysis

The dichotomous DY scores were cross-tabulated for concordance or discordance in DY between the two siblings within a pair (DY1+/- versus DY2 +/-). This was done separately for sisters, monozygotic twins and dizygotic twins. For each of the three populations, a correlation coefficient was calculated that is based on a comparison of the expected and observed number of concordant and discordant pairs (tetrachoric correlation, SAS version 8). As is described and utilised extensively in genetic epidemiology, the heritability (H^2) of a factor is the percentage of variance in a trait that can be ascribed to variation in genetics. Heritability is usually estimated through several Mendelian based approaches[19, 20], depending on the family relationship that is being studied.

(a) for monozygotic twins, the heritability equals the correlation;

(b) for dizygotic twins, the heritability equals two times the correlation;

(c) for sisters, the heritability equals two times the correlation;

(d) two times the difference between monozygotic twins and dizygotic twins;

(e) two times the difference between monozygotic twins and sisters.

These methods are based on the assumption that monozygotic twins share 100% and dizygotic twins and sisters on average share 50% of their genes. Consequently, if genetic effects are completely additive, then the correlation in monozygotic twins will be twice as high as the correlation between dizygotic twins or sisters. All five approaches we used to estimate the heritability (a–e) will then yield equal result and the combined models (d, e) are then limited to 100%. A result in the combined models (d, e) which exceeds 100% indicates that the data does not fit the additive model and that more complex genetic effects are present.

To evaluate whether non-genetic determinants of breast patterns were distributed equally according to DY-concordance, we compared the average within-pair difference for a number of characteristics between DY-concordant sister pairs and DY-discordant sister pairs. Differences in means were tested by ANOVA or a Chi-square test. In case of significant differences ($p_{2-sided} < 0.05$), heritability-estimates were calculated for a homogeneous sub-set of pairs. For example; DY-concordant and DY-discordant sib pairs may differ in the prevalence of nulliparity. Nulliparity is known to be related to a higher frequency of DY. So if one member of a sib pair is nulliparous and the other is parous, heritability could be masked due to the influence of parity on the DY patterns. We therefore estimated the heritability in DY in sub-sets of sister pairs in which either both sisters were parous or were both nulliparous.

Results

The characteristics of the study population, with regard to a number of factors that are known to influence the breast parenchymal pattern and/or the breast cancer risk, are shown in Table 1. A total of 466 sister pairs; 932 subjects, 26 dizygotic twin-pairs; 52 subjects and 25 monozygotic twin pairs; 50 subjects, were available for the study. Due to the use of different questionnaires during the screening program, information was not available for all subjects for all characteristics. Nonetheless, the characteristics of the study-population as they were found are representative for the cohort from which the population was drawn.

Of 466 sister pairs, 61% (285 pairs) was concordant as to their DY pattern score. For dizygotic and monozygotic twins, these percentages were 58% (15 of 26) and 88% (22 of 25) respectively (Table 2).

The estimated heritability of dysplastic pattern was 38% (r = 0.19, 95%CI 0.04–0. 34) for sisters (2 x $r_{sisters}$), 34% (r = 0.17, 95%CI -0.45–0.79) for dizygotic twins (2 x $r_{dizygotic\ twins}$) and 88% (r = 0.88, 95%CI 0.63–1.00) for monozygotic twins ($r_{monozygotic\ twins}$). Through model d, which combines the estimates of monozygotic twins and dizygotic twins, the overall heritability of the breast dysplastic pattern is estimated at 142% (2 x ($r_{t,monozygotic\ twins}$ -

$r_{t,dizygotic\ twins}$)). When using sisters as a proxy for dizygotic twins (model e), this estimate is 138% (2 x ($r_{t,monozygotic\ twins}$ - $r_{t,sisters}$)). These values over 100% show that the additive models are not sufficient to fit the data and are indicative of additional, non-additive effects.

Table 1 Characteristics of sisters, dizygotic (DZ) twins and monozygotic (MZ) twins.

Characteristic	Sisters		DZ Twins		MZ Twins	
	N[a]	Mean $\pm$ SE	N[a]	Mean $\pm$ SE	N[a]	Mean $\pm$ SE
Age at Examination (yr)	932	50.0 $\pm$ 0.21	52	51.7 $\pm$ 0.56	50	52.9 $\pm$ 0.72
Height (m)	932	1.63 $\pm$ 0.002	52	1.62 $\pm$ 0.009	50	1.61 $\pm$ 0.007
Weight (kg)	932	67.4 $\pm$ 0.35	52	66.7 $\pm$ 1.55	50	64.9 $\pm$ 1.43
Body Mass Index (kg/m^2)	932	25.1 $\pm$ 0.12	52	25.5 $\pm$ 0.53	50	25.1 $\pm$ 0.59
Age at Menarche (yr)	927	13.8 $\pm$ 0,05	35	13.5 $\pm$ 0.23	29	13.7 $\pm$ 0.27
Age at Natural Menopause (yr)	274	49.4 $\pm$ 0.25	16	48.1 $\pm$ 0.81	21	50.1 $\pm$ 0.78
Time Since Natural Menopause (yr)	274	7.2 $\pm$ 0.32	16	7.1 $\pm$ 1.06	21	4.8 $\pm$ 0.53
Children (average number)[b]	761	3.5 $\pm$ 0.06	46	3.3 $\pm$ 0.21	33	3.4 $\pm$ 0.33
Age at First Childbirth[b] (yr)	761	26.0 $\pm$ 0.15	46	27.0 $\pm$ 0.55	33	26.2 $\pm$ 0.87
Parity (% nulliparous)	932	18%	52	12%	50	34%
Menopause (% natural)	353	78%	29	55%	31	68%

[a] Number of subjects for which data was available. Values were not available for all women due to the use of different questionnaires in the course of the screening program.
[b] Only parous women included.

Table 2 Distribution of concordant pairs, correlation coefficients and heritability estimates of dysplastic glandular tissue patterns in sisters, dizygotic (DZ) twins and monozygotic (MZ) twins.

	N	DY-conc./DY-disc.[a]	r_t (95% CI)[b]	Heritability (H^2)[c]	
Sister pairs	466	285/181	0.19 (0.04 - 0.34)	2 x $r_{sisters}$ x 100% =	38%
DZ Twin pairs	26	15/11	0.17 (-0.45 - 0.79)	2 x $r_{DZ\ twins}$ x 100% =	34%
MZ Twin pairs	25	22/3	0.88 (0.63 – 1.00)	$r_{MZ\ twins}$ x 100% =	88%
Combined estimate from MZ and DZ twins				2 x (r_{MZ} – r_{DZ}) x 100% =	142%
Combined estimate from MZ twins and sisters				2 x (r_{MZ} – r_{Sister}) x 100% =	138%

[a] Dichotomized DY patterns according to Wolfe[4].
[b] Correlation coefficient.
[c] Formula to compute heritability based on literature[19, 20].

Table 3 shows the distribution of within-pair differences of known determinants of DY patterns in the sister population. DY-concordant sister pairs and DY-discordant sister pairs differed significantly ($p_{2\text{-sided}}$ < 0.05) in the average amount of within-pair difference in height, BMI and parity status. All sister pairs were subsequently stratified into quartiles of the within pair difference for height and BMI separately. For both variables, pairs in the first quartile were considered homogeneous. When we repeated the calculation of the

heritability, including only sister pairs that were homogeneous for BMI or height the estimated heritability rates were 60 and 68% respectively (Table 4).

For pairs homogeneous for parity (stratified for general parity-status as well as regarding the number of children), the estimated heritability rates were 46% for pairs with a similar parity status and 2% for pairs for who the two sisters differed in their parity status. Pairs with two nulliparous sisters had an estimated heritability of 90% and pairs with two parous sisters lead to an estimate of 26%, ranging from 30% in the substratum of pairs with an equal number of children to 24% in the sub-strata of pairs with an unequal number of children.

Discussion

The results of this study show a substantial influence of genes on the parenchymal pattern of the female breast. The estimates of the heritable effect range from 34 to 88% in different types of sibs. This study also shows that complex genetic effects, such as dominant genes or gene–environment interaction may be involved. Analyses of a number of non-genetic factors indicate that some of these factors, especially parity, have a substantial influence on the heritability estimates.

Our study allowed for the analyses of sisters as well as dizygotic and monozygotic twins from the same population. The results enable a high generalisation since they are based on a population-based screening cohort that was selected independent of breast cancer predisposition with a participation rate of over 80%. The estimates in our study differed only slightly from those previously reported and are of a similar direction and magnitude.

Since this study was begun in the early days of mammography where mammographic resolution may have been poor and breast density high, particularly in pre-menopausal women, we analysed the data in this regard. We obtained heritability from sisters with mammography in the 1970s (first portion of the recruitment period) and compared this to the estimate obtained from sisters screened in the 1980s. These estimates were 42% (r = 0.21, 95%CI -0.05–0.47) and 24% (r = 0.12, 95%CI -0.07–0.31) respectively and do not significantly differ from each other or the grouped estimate.

In a previous study, Wolfe et al. included 122 sister pairs of which both sisters had been referred for diagnostic work-up in a breast clinic[21]. Although the original paper had a different set-up, their data can be restructured into a fourfold table of DY-scores, consisting of 88 DY-concordant pairs (63 DY+/DY+ and 25 DY-/ DY-) and 34 DY-discordant pairs. This results in an estimated heritability of 78%, which is even substantially higher than our estimate in sisters.

However, since the women in that study were selected on the basis of their presence in the hospital files and had been referred because of breast symptoms, this may have caused an artificially high prevalence of DY positive scores. Women with a high DY score are known to have a higher risk of developing breast cancer and are also more often referred to the clinic because of a mammogram that is more difficult to interpret. In the population of Wolfe, 553 of 928 women (60%) had a P2 or DY score. This corresponds to DY+ in our population, which

Table 3. Mean difference in characteristic between sisters within a pair, according to pairs concordant and discordant for their DY score.

	DY-Concordant sister pairs		DY-Discordant sister pairs		
			Within-Pair Differences		
Characteristic	N[a]	Mean Δ	N[a]	Mean Δ	p
Age at Examination (yr)	285	3.39	181	3.30	0.70
Height (m)	285	0.04	181	0.05	0.03*
Weight (kg)	285	9.96	181	9.08	0.29
Body Mass Index (kg/m^2)	285	3.60	181	3.03	0.04*
Age at Menarche (yr)	283	1.45	180	1.50	0.64
Age at Natural Menopause (yr)	86	4.42	52	4.54	0.59
Time Since Natural Menopause (yr)	86	5.12	52	5.66	0.59
Children [b] (average number)	211	1.41	118	1.23	0.26
Age at First Childbirth (yr)	211	3.63	118	3.81	0.62
Parity[c] (discordant pairs)[d]	285	52	181	50	0.02*
Natural Menopause[c] (discordant pairs)[e]	86	28	52	10	0.12

[a] Only pairs with a valid result for both siblings were included.

[b] Only parous women included.

[c] Chi-square test. All other variables compared by ANOVA.

[d] 52 of 285 DY-concordant and 50 of 181 DY-discordant sib pairs include one women who has had at least one child, while her sister has not.

[e] 28 of 86 DY-concordant and 10 of 52 DY-discordant sib pairs include one women who has had a natural menopause, while her sister has not.

* Significant difference between groups; $p_{2-sided} < 0.05$

Table 4 Heritability rates for homogeneous sister pairs.

	N	DY-conc./DY-disc.	r_t (95%CI)	H^2
BMI[a]				
	116	70/46	0.30 (0.02 - 0.57)	60%
Height[b]				
	116	75/41	0.34 (0.07 - 0.62)	68%
Parity				
Concordant parity within pair	363	233/130	0.23 (0.05-0.40)	46%
Both sisters nulliparous	34	22/12	0.45 (-0.01-0.91)	90%
Both sisters parous	329	211/118	0.13 (-0.06-0.32)	26%
Equal number of children	104	65/39	0.15 (-0.18-0.49)	30%
Unequal number of children	225	146/79	0.12 (-0.12-0.36)	24%
Discordant parity within pair	103	52/51	0.01 (-0.30-0.31)	2 %

[a] Homogeneous pairs: pairs with a within pair BMI difference belonging to the lowest quartile of within pair BMI differences. The mean within pair BMI difference was 0.51 kg/m^2; range 0.01-1.08 kg/m^2

[b] Homogeneous pairs: pairs with a within pair Height difference belonging to the lowest quartile of within pair Height differences. The mean within pair Height difference was 0.81 cm; range 0 −1.70 cm

was found in 333 of 1034 women (32%). This high prevalence of DY+ scores in the study by Wolfe may have lead to a higher prevalence of DY-concordant sister pairs and thereby an overestimation of the heritability, if interpreted to reflect the general population.

Recently, Boyd et al. (2002) reported an estimated heritability of 63% in a cohort of monozygotic and dizygotic twins, based on continuously measured percent DY scores adjusted for age, BMI, age at menarche, cessation of menstruation, parity and (in parous women) number of live births and age at first birth[15]. Prior to this, in a study in sisters belonging to breast cancer families, Pankow et al. also reported a statistically significant heritability (32–54%) in continuously measured percent breast density adjusted for several factors[14].

As it is known that certain determinants influence the dysplastic appearance of the breast, we determined the within-pair difference in these determinants in our population of sisters. The average within-pair difference in these determinants between the sister-pairs with a concordant DY score and those with a discordant DY score was compared with ANOVA or a chi-square test. This indicated that the sister-pairs with a discordant DY score, on average, had a significantly ($p_{2-sided} < 0.05$) higher difference in BMI, height and parity-status.

In sister pairs that were homogeneous for BMI (116 pairs in the lowest quartile of within pair difference), the estimated heritability was 60% (r = 0.30, 95%CI 0.02–0.57). For sisters homogeneous for their height (116 pairs), the estimated heritability was 68% (r = 0.34, 95% CI 0.07–0.62).

For all sister pairs with a similar parity-status (both sisters parous or both sisters nulliparous), the estimated heritability was 46% (r = 0.23, 95% CI 0.05–0.40). Estimates in parous pairs were similar and not affected by a difference in the number of children. In nulliparous pairs, however, the estimated heritability was 90% (r = 0.45, 95% CI -0.01–0.91). Although this estimate is based on only 34 observations, it is in sharp contrast to the estimated heritability in sister pairs with a dissimilar parity status, which was only 2% (r = 0.01, 95% CI -0.30–0.31).

These results indicate that there may be a very strong, genetically determined tendency for siblings to develop a similar parenchymal pattern, but that this heritability may be obscured in a heterogeneous population. As it appears, changes in a women's parenchymal pattern, caused by a first childbirth, are so pronounced that they can completely mask the heritable influence. The similarity apparently partially returns if both sibs are parous, while it is only marginally affected by the additional number of children. In addition to an important role for genes, this clearly shows important roles for non-genetic factors (in this case we assume that nulliparity is not genetically determined). Although the confidence intervals of the estimates in the various strata are wide and are overlapping, these results indicate that non-genetic influences, especially parity, may have a strong influence on the heritability of parenchymal patterns.

In our study, heritability in the various types of sibs was highest in the monozygotic twins and lower in dizygotic twins and sisters. The heritability in monozygotic twins is thought to be overestimated due to a higher concordance in other, non-genetic, circumstances. This

influence of other factors is indicated by the analysis in selected homogeneous sister pairs. An assumption that is made in twin studies is that environmental factors are the only cause of phenotypic variance in monozygotic twins. In the estimation of heritability, shared environmental influences can, however, not be fully excluded.

This argument is supported by our estimation of the heritability from the combined approaches. In these approaches, the heritability is estimated by two times the difference between the correlation in monozygotic twins minus the correlation in dizygotic twins or sisters (H^2 = 2 x ($r_{monozygotic}$ - $r_{dizygotic}$) or H^2 = 2 x ($r_{monozygotic}$ - $r_{sisters}$)). It is based on the number of genes that is shared by the different types of siblings. In the case of a purely additive genetic affect, the estimated heritability from monozygotic twins, who share 100% of their genes, should be double the effect found in dizygotic twins or sisters who only share half of their genes. This would result in outcomes from the combined approaches which are equal to the outcome in monozygotic twins alone[19, 20].

In our case, the combined approach results in an estimated heritability of 142% when using the dizygotic-estimate and 138% when using the sister estimate. These estimates exceed 100% and thus indicate that the 'simple' additive model does not fit the data. With regard to the small number of mono- and dizygotic twins and the indication that environmental effects have had a certain influence on the heritability estimates in sibs, these high estimates may be due to random variation. They may, however, also be an indication that the factors that affect parenchymal patterns are more than merely genetically additive and that a dominant genetic influence, shared environmental factors, genetic interaction (epistasis) or gene–environment interaction also play a role.

The heritability of DY and the mechanisms underlying it are important, because they may point to the influence of genes in relation to the occurrence of breast cancer. A study in Scandinavian twins revealed a heritability of breast cancer of 27% (95% CI 4–41%)[22]. Known genes involved in breast cancer, such as BRCA1, BRCA2 and ATM, however account for only a small proportion of this effect, leaving sufficient room for other, less penetrating genes, gene–gene interaction and/or gene–environment interaction[23]. Vachon et al. reported the identification of a region on chromosome 6 that appeared to be linked to breast density, which strengthens the evidence towards an independent genetic influence on breast density[24]. In a more recent publication, the predictive value of polymorphisms in genes involved in steroid hormone biosynthesis and metabolism were studied, but a consistent relation could not be found[25]. To our knowledge, further advances have not been reported so far.

In conclusion, genetic factors may play a substantial role in breast density, but modifying effects by other (non-genetic) factors such as parity are just as important. In view of the established relation between high risk parenchymal patterns and breast cancer risk, the identification of genes that are involved in breast density could be an important step in breast cancer research.

Reference List

1. Bland KI, Kuhns JG, Buchanan JB et al. A clinicopathologic correlation of mammographic parenchymal patterns and associated risk factors for human mammary carcinoma. Ann Surg 1982; 195(5):582-594.

2. Boyd NF, Lockwood GA, Byng JW, Tritchler DL, Yaffe MJ. Mammographic densities and breast cancer risk. Cancer Epidemiol Biomarkers Prev 1998; 7(12):1133-1144.

3. Boyd NF, Jensen HM, Cooke G, Han HL. Relationship between mammographic and histological risk factors for breast cancer. J Natl Cancer Inst 1992; 84(15):1170-1179.

4. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976; 37(5):2486-2492.

5. Boyd NF, Fishell E, Jong R et al. Mammographic densities as a criterion for entry to a clinical trial of breast cancer prevention. Br J Cancer 1995; 72(2):476-479.

6. Saftlas AF, Wolfe JN, Hoover RN et al. Mammographic parenchymal patterns as indicators of breast cancer risk. Am J Epidemiol 1989; 129(3):518-526.

7. Bergkvist L, Tabar L, Bergstrom R, Adami HO. Epidemiologic determinants of the mammographic parenchymal pattern. A population-based study within a mammographic screening program. Am J Epidemiol 1987; 126(6):1075-1081.

8. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer 1998; 78(9):1233-1238.

9. Brisson J, Merletti F, Sadowsky NL, Twaddle JA, Morrison AS, Cole P. Mammographic features of the breast and breast cancer risk. Am J Epidemiol 1982; 115(3):428-437.

10. Byrne C, Schairer C, Wolfe J et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995; 87(21):1622-1629.

11. Leinster SJ, Walsh PV, Whitehouse GH, al Sumidaie AM. Factors associated with mammographic parenchymal patterns. Clin Radiol 1988; 39(3):252-256.

12. Vachon CM, Kushi LH, Cerhan JR, Kuni CC, Sellers TA. Association of diet and mammographic breast density in the Minnesota breast cancer family cohort. Cancer Epidemiol Biomarkers Prev 2000; 9(2):151-160.

13. van Gils CH, Otten JD, Verbeek AL, Hendriks JH. Short communication: breast parenchymal patterns and their changes with age. Br J Radiol 1995; 68(814):1133-1135.

14. Pankow JS, Vachon CM, Kuni CC et al. Genetic analysis of mammographic breast density in adult women: evidence of a gene effect. J Natl Cancer Inst 1997; 89(8):549-556.

15. Boyd NF, Dite GS, Stone J et al. Heritability of mammographic density, a risk factor for breast cancer. N Engl J Med 2002; 347(12):886-894.

16. de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 1984; 37(1):1-44.

17. de Bruin JP, Bovenhuis H, van Noord PA et al. The role of genetic factors in age at natural menopause. Hum Reprod 2001; 16(9):2014-2018.

18. Rombach JJ, Collette BJ, de Waard F, Slotboom BJ. Analysis of the diagnostic performance in breast cancer screening by relative operating characteristics. Cancer 1986; 58(1):169-177.

19. Falconer DS, MacKay TFC. Introduction to Quantative Genetics. 4th Edition ed. New York: Longman Inc.; 1996.

20. Rice TK, Borecki IB. Familial resemblance and heritability. Adv Genet 2001; 42:35-44.

21. Wolfe JN, Albert S, Belle S, Salane M. Familial influences on breast parenchymal patterns. Cancer 1980; 46(11):2433-2437.

22. Lichtenstein P, Holm NV, Verkasalo PK et al. Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. N Engl J Med 2000; 343(2):78-85.

23. Hakansson S, Johannsson O, Johansson U et al. Moderate frequency of BRCA1 and BRCA2 germ-line mutations in Scandinavian familial breast cancer. Am J Hum Genet 1997; 60(5):1068-1078.

24. Vachon CM, King RA, Atwood LD, Kuni CC, Sellers TA. Preliminary sibpair linkage analysis of percent mammographic density. J Natl Cancer Inst 1999; 91(20):1778-1779.

25. Haiman CA, Bernstein L, Berg D, Ingles SA, Salane M, Ursin G. Genetic determinants of mammographic density. Breast Cancer Res 2002; 4(3):R5.

Comparing correlations of continuous observations from two independent populations using a sequential approach

M Fazil Baksh[1]

Gerco Haars

Susan Todd[2]

Paul AH van Noord

John Whitehead[2]

[1]Medical Research Council Epidemiology Unit, Strangeways Research Laboratory,
Worts Causeway, Cambridge CB1 8RN, U.K.

[2]Medical and Pharmaceutical Statistics Research Unit, The University of Reading,
P.O. Box 240, Reading RG6 6FN, U.K.

2.2

Abstract

A sequential study design generally makes more efficient use of available information than a fixed sample counterpart of equal power. This feature is gradually being exploited by researchers in genetic and epidemiological investigations that utilize banked biological resources and in studies where time, cost and ethics are prominent considerations. Recent work in this area has focussed on the sequential analysis of matched case–control studies with a dichotomous trait. In this paper, we extend the sequential approach to a comparison of the associations within two independent groups of paired continuous observations. Such a comparison is particularly relevant in familial studies of phenotypic correlation using twins. We develop a sequential twin method based on the intraclass correlation and show that use of sequential methodology can lead to a substantial reduction in the number of observations without compromising the study error rates. Additionally, our approach permits straightforward allowance for other explanatory factors in the analysis. We illustrate our method in a sequential heritability study of dysplasia that allows for the effect of body mass index and compares monozygotes with pairs of singleton sisters.

Introduction

A sequential study consists of a series of interim analyses at which accumulating data are examined with the objective of terminating data collection as soon as sufficient evidence regarding the hypothesis of interest has been observed. As a consequence, a sequential study has the potential for requiring fewer observations than the corresponding fixed sample size test without compromising the reliability of the conclusions. Researchers in genetic and epidemiological studies are gradually recognizing this advantage, particularly in studies where cost and efficiency are primary concerns: see Reference 1 for a review of the literature on sequential genetic studies. Sequential matched case–control studies with a dichotomous trait and based upon the score statistic have been proposed and evaluated by Van der Tweel and Van Noord[2] in a study on the MTHFR gene and rectal cancer, and applied by Dreyfus et al.[3] in an investigation of the relationship between antiphospholipid antibodies and preeclampsia. The use of a similar sequential approach to detect gene–environment interactions was investigated by Van der Tweel and Schipper[4]. Design features of the sequential matched case–control study with a dichotomous trait were explored by Baksh et al.[5].

In this paper we concentrate on association studies with a continuous trait. Consider two different and independent populations of pairs of related individuals and suppose that it is of interest to compare the within-pair trait associations of the two populations. Considering association in terms of the within-pair correlation, we develop and assess a sequential procedure for populations of distinguishable pairs of individuals, such as mother–daughter pairs. This procedure is then adapted to the comparison of monozygotic (identical) twins and other siblings, either dizygotic (non-identical) twins or pairs of singletons.

Our sequential correlation test is motivated by an ongoing investigation concerning the influence of genetic factors on dysplasia (a risk factor for cancer) in breast tissue of women from the DOM cohort[6], and the demand for efficient use of the available information. Efficiency issues arising from use of the DOM cohort in other genetic association studies motivated the procedures of References 2 & 4. In both instances a sequential approach was recommended. This paper extends the sequential procedure to heritability studies. Heritability analyses[7] using sisters (twins and singletons) from the DOM cohort was based on dichotomous scoring of dysplasia; the method developed in this paper is for continuous measurements and allows for other explanatory factors.

A sequential procedure is characterized by a stopping rule that determines when sufficient information regarding the hypotheses of interest has been collected and a decision rule that states what terminal conclusion regarding the hypotheses should be drawn. The null hypothesis to be tested is expressed as $H_0 : \theta = 0$, where θ is a suitably defined parameter of interest. The test can be defined in terms of two statistics Z and V, where Z is a cumulative measure of the evidence against H_0 and V, the variance of Z under H_0, represents the information about θ currently available. Two popular versions of the test are the sequential Wald test and the sequential score test. In a sequential Wald test $Z = \hat{\theta}V$ and $V = \{se(\hat{\theta})\}^{-2}$,

where $\hat{\theta}$ and $se(\hat{\theta})$ denote the maximum likelihood estimate of θ and its standard error. In a sequential score test, Z is the efficient score for θ and V is Fisher's information.

The sequential designs in this paper are based on the boundaries approach[8]. The test procedure involves comparing a plot of Z against V with certain stopping limits. Sequential designs defined through an α-spending function[9] can be expressed in the same way. The plot is called the sample path, while the increment in V between two consecutive interim analyses is called an inspection interval. The test procedure terminates as soon as the sample path crosses a boundary. The designs in this paper have limits obtained under the assumption that Z is normally distributed with mean θV, variance V and that successive increments in Z are independent. In other words we assume that the sequence of increments in Z forms points on a Brownian motion process with drift θ.

In this paper we investigate use of the triangular test[8] and discuss why this test may be appropriate in many situations. We consider both sequential Wald and sequential score tests. Using simulated data we assess the distributional assumption of the test statistic Z and show that an adaptation of Wald test is better than both the unmodified Wald test and the score test in preserving the error rates of the sequential procedure. We assess an approach for evaluating the average size of the sequential study and show that the sequential procedure is more efficient than the equivalent fixed sample procedure. Finally, application of the procedure to test for differences in the trait association of monozygotic and dizygotic siblings is described and this procedure is illustrated using data from the DOM cohort.

Breast Dysplasia and the DOM Cohort

The DOM project was initiated in 1974 as a population-based screening project for the early detection of breast cancer in residents of the city of Utrecht, its suburbs and adjacent areas. By the time it terminated in 1988 it had progressed in four cohorts and involved over 50 000 participants. These cohorts naturally included a large number of groups of relatives and thus presented the opportunity for genetic studies of breast cancer. One such study concerns using information on twins to investigate heritability of breast dysplasia. A difference in the association of dysplasia between monozygotic (MZ) sisters and between dizygotic (DZ) sisters would suggest that some women may be genetically predisposed to developing breast dysplasia.

Breast dysplasia, visible as radio dense tissue on mammograms, has been identified as a risk factor and/or an indicator of breast cancer. It is generally scored in categories related to the percentage of the breast occupied by dense tissue. Originally the mammograms in the DOM project were scored dichotomously for the presence or absence of dysplastic tissue; subsequent digitization allowed computer-aided continuous measurements of the dense tissue. The process of measuring the dysplasia involves first marking out and then counting the number of pixels in the marked areas. This is time-consuming and costly. Another critical feature of this investigation is that because of privacy considerations no records were kept on the identity of participants in the cohort and thus identification of which individuals are related, is progressing slowly. Related persons are identified using a method proposed in

Reference 10 and data either collected at visits by trained study staff or from questionnaires completed by participants in the breast cancer screening project. Only a small number of twins have so far been found. For these reasons it is very important to use the available data as efficiently as possible and a sequential study design was considered.

The Wald and Score Procedures

In this section we present the Wald and score procedures for comparing the within pair associations of two independent, bivariate sets of continuous observations. We consider populations of distinguishable pairs of individuals, such as mother–daughter and older sister–younger sister pairs; our approach for indistinguishable pairs is covered later. In the following section use of these statistics in a sequential setting is investigated and, based on properties identified in a simulation assessment, the most accurate approach is selected.

Let (x_{1j}, x_{2j}) denote the j th of n_X samples from a bivariate normal distribution with means μ_{X_1}, μ_{X_2} variances $\sigma^2_{X_1}$, $\sigma^2_{X_2}$ and correlation coefficient ρ_X and let (y_{1k}, y_{2k}) denote the k th of n_Y samples from a second bivariate normal distribution with means μ_{Y_1}, μ_{Y_2} variances $\sigma^2_{Y_1}$, $\sigma^2_{Y_2}$ and correlation coefficient ρ_Y. The test of a difference in the associations can be expressed in terms of the correlation coefficients ρ_X and ρ_Y of the two sets of observations. Because of certain distributional constraints Fisher's variance stabilizing (z) transformation is often used in the fixed sample comparison of the correlation coefficients. The traditional fixed sample procedure based on this transformation is described, for instance, in Reference 11 while the distributional constraints are discussed in Reference 12. Essentially, this procedure is concerned with inference about the parameter θ defined as the difference $\eta_Y - \eta_X$ where

$$\eta_X = \frac{1}{2} \ln\left(\frac{1+\rho_X}{1-\rho_X}\right) \text{ and } \eta_Y = \frac{1}{2} \ln\left(\frac{1+\rho_Y}{1-\rho_Y}\right) \tag{1}$$

are the Fisher's z-transformations of ρ_X and ρ_Y, respectively. We focus on inference regarding θ based on the Wald and score procedures in a sequential setting.

Although they are standard in sequential analysis, Wald and score approaches have not previously been investigated in the context considered here. An essential feature of both procedures is that statistics can be found with a limiting normal distribution, with mean related to θ and with a variance that is easily calculated. In Wald's procedure, which is based on the maximum likelihood estimate $\hat{\theta}$ of the parameter θ, the variance of $\hat{\theta}$ is often difficult to evaluate and instead is replaced by an estimate of the variance of the asymptotic distribution of $\hat{\theta}$. This approach is referred to as the W_e approach[13]. The traditional fixed sample test uses a modified form of the variance of the W_e procedure for bivariate normal samples and can therefore be considered a variant of Wald test. To distinguish this

traditional approach from the W_e approach, the former will be referred to as the W_v approach. In this paper we investigate the sequential use of the W_e, W_v and the score (denoted by W_u) statistics by comparing their respective anticipated study sizes and study error rates.

Wald's statistics

Wald's W_e approach and the score approach can be derived from the quadratic form of the Taylor's series expansion of the maximum likelihood ratio statistic. For a likelihood with parameters θ and ψ, where ψ is a vector of nuisance parameters, Taylor's series expansion, up to quadratic terms, of the maximum likelihood ratio statistic about the parameter point (θ_0, ψ) may be written as $(\hat{\theta} - \theta_0)^2 I(\theta_0, \psi)$; where θ_0 is the value of θ under the null hypothesis and $I(\theta_0, \psi)$ is the inverse of the variance in the limiting marginal normal distribution of $\hat{\theta}$. Under the null hypothesis, the limiting distribution of this quadratic form is central chi-squared with one degree of freedom, or equivalently, the distribution of $\hat{\theta} I(\theta_0, \psi)$ is normal with mean $\theta_0 I(\theta_0, \psi)$ and variance $I(\theta_0, \psi)$. Asymptotically equivalent test statistics are obtained in the W_e approach where $I(\theta_0, \psi)$ is replaced by $I(\hat{\theta}_0, \hat{\psi})$. Section 9.3 of Reference 13 provides a thorough discussion of the results mentioned above.

Using the invariance property of maximum likelihood estimators, $\hat{\theta} = \hat{\eta}_Y - \hat{\eta}_X$, where $\hat{\eta}_Y$ and $\hat{\eta}_X$ are Fisher's z-transformations of the maximum likelihood estimates $\hat{\rho}_Y$ and $\hat{\rho}_X$, respectively (see Appendix A), and we are interested in testing the null hypothesis that θ is zero. In the presence of nuisance parameters ψ comprising the means and variances of the two bivariate normal distributions and the sum $\phi = \hat{\eta}_Y + \hat{\eta}_X$ we obtain $I(\hat{\theta}; \hat{\phi})^{-1} = n_X^{-1} + n_Y^{-1}$ and the asymptotic null distribution of $(n_X^{-1} + n_Y^{-1})^{-1} \hat{\theta}$ is normal with mean 0 and variance $(n_X^{-1} + n_Y^{-1})^{-1}$. Thus the sequential implementation of the W_e procedure in the boundaries approach uses $Z = (n_X^{-1} + n_Y^{-1})^{-1} \hat{\theta}$ and $V = (n_X^{-1} + n_Y^{-1})^{-1}$ at each interim analysis. The expected Fisher's information $I(\theta; \phi)$ is derived in Appendix A.

The fixed sample W_v test procedure is similar to the W_e approach except that, in both Z and V, n_X is replaced by $n_X - 3$ and n_Y by $n_Y - 3$ so as to improve the approximation to normality of the distribution of the test statistic[12]. From the literature it would appear that this is an empirical result based on the asymptotic distribution of the maximum likelihood estimate of Fisher's z-transformation, rather than being derived from any general principle. Using Fisher's results for the asymptotic distribution of the correlation coefficient[14], the distribution of the maximum likelihood estimate of Fisher's z-transformation is asymptotically normal with variance equal to the inverse of the number of observed pairs less two, rather than three.

An important feature of a sequential test procedure is its potential for requiring fewer observations than the equivalent fixed sample counterpart design. The fixed sample test based on θ for no difference in pairwise associations is most powerful when the number of pairs of observations in each sample are equal. In this case, assuming that $\hat{\theta}$ is normal with mean θ and variance $1/I(\theta;\phi)$, the number of pairs $n_f(W_e)$ required in each group for the fixed sample W_e test with one-sided significance level α and power 1-βto detect is $\theta = \theta_1$ is

$$n_f(W_e) = 2\left(\frac{u_\alpha + u_\beta}{\theta_1}\right)^2 \tag{2}$$

where u_γ is the 100$(1-\gamma)$ percentage point of the standard normal distribution. The number of pairs, $n_f(W_v)$, for the equivalent W_v test is $n_f(W_v) = n_f(W_e) + 3$. This estimate of sample size is used in the comparison of the fixed sample and sequential study requirements.

The score test statistics

The score (W_u) procedure is an asymptotic equivalent of Wald procedure. In the absence of nuisance parameters the efficient score and observed Fisher's information about θ are, respectively, the first derivative and the negative of the second derivative of the log-likelihood with respect to θ. When, as in our case, nuisance parameters are present, a common approach[8] uses the profile likelihood instead of the likelihood in deriving the score and observed Fisher's information. A profile likelihood is the likelihood with the nuisance parameters replaced by their respective maximum likelihood estimates.

The log-likelihood ℓ of samples from two independent bivariate normal distributions is simply the sum $\ell_X + \ell_Y$, where ℓ_X and ℓ_Y denote the respective log-likelihoods of the two samples. Using the profile likelihood approach, the efficient score Z and observed Fisher's information V about θ, evaluated under the null hypothesis $\theta = 0$, are shown in Appendix A to be given by $Z = (Z_Y - Z_X)/2$ and $V^{-1} = V_Y^{-1} + V_X^{-1}$ where Z_X, V_X and Z_Y, V_Y are the first and negative of the second derivatives of the profile log-likelihoods for ℓ_X and ℓ_Y, respectively. Under the null hypothesis the correlation coefficients ρ_X and ρ_Y are equal to some common value ρ (say) and we have

$$Z_X = n_X \frac{\hat{\rho}_X - \hat{\rho}}{1 - \hat{\rho}\hat{\rho}_X} \quad \text{and} \quad V_X = n_X\left[1 - \left(\frac{\hat{\rho}_X - \hat{\rho}}{1 - \hat{\rho}\hat{\rho}_X}\right)^2\right] \tag{3}$$

where $\hat{\rho}_X$ denotes the maximum likelihood estimate of ρ_X using only the likelihood ℓ_X and $\hat{\rho}$ is the maximum likelihood estimate of the common value. Similar expressions are obtained for Z_Y and V_Y.

The estimate $\hat{\rho}$ of the common correlation coefficient ρ of two independent bivariate normal samples has been shown by Brandner[15] to be

$$\frac{(n_X+n_Y)(1+\hat{\rho}_X\hat{\rho}_Y)-[(n_X+n_Y)^2(1-\hat{\rho}_X\hat{\rho}_Y)^2-4n_Xn_Y(\hat{\rho}_Y-\hat{\rho}_X)^2]^{1/2}}{2(n_X\hat{\rho}_Y+n_Y\hat{\rho}_X)} \quad (4)$$

In particular when $n_X = n_Y = n$ (say) this estimate reduces to $[(1+\hat{\rho}_X\hat{\rho}_Y)-\sqrt{(1-\hat{\rho}_X^2)(1-\hat{\rho}_Y^2)}]/(\hat{\rho}_X+\hat{\rho}_Y)$ and the efficient score and Fisher's information now become $Z=n\tanh(\hat{\theta}/2)$ and $V=n[1-\tanh^2(\hat{\theta}/2]/2$, where $\hat{\theta}$ is the maximum likelihood estimate of θ.

In general, an asymptotic estimate of the observed Fisher's information V_f required for the fixed sample study with one-sided significance level α and power $1-\beta$ to detect $\theta = \theta_1$ based on the limiting normal distribution of the efficient score is given by $V_f = [u_\alpha + u_\beta)/\theta_1]^2$. Therefore, in studies where the number of pairs in each of the two groups are equal, the required number of pairs per group in order to achieve power $1-\beta$ is $n_f(W_u) = 2[u_\alpha + u_\beta)/\theta_1]^2 \cosh^2(\theta_1/2)$. Comparison with (2) shows that the fixed sample W_u procedure is slightly less efficient than the equivalent W_e procedure. The properties of the two methods in a sequential framework is next investigated.

The Wald and Score Statistics in a Sequential Setting

A sequential test procedure using the boundaries approach (as implemented, for example, in the PEST 4 software[16]) usually involves comparing a plot of Z against V with certain stopping limits. Let the random variable V^* denote the value of V when the sequential procedure terminates. For the sequential comparison of a difference in associations, the type I error rate α and the power $1-\beta$ to reject the null hypothesis H_0 when $\theta = \theta_1$, where $\theta_1 >0$ denotes the difference $\eta_Y - \eta_X$ of epidemiological importance, will be established at the design stage. Once a power requirement has been set, rival designs can be compared in terms of their terminal sample size properties. A design that will tend to have small samples is needed in situations where early stopping is desirable, with larger samples being permissible when ethical or cost considerations would allow this. As the terminal value V^* is approximately proportional to the final sample size, its properties can be used to select a suitable design. The PEST software provides values of the expectation $E_\theta(V^*;\theta)$, the median $MED(V^*;\theta)$ and the 90th percentile $P90(V^*;\theta)$ of this quantity for a number of designs and this approach is described in Chapter 4 of Reference 8.

A triangular test is appropriate when it is important to reject H_0 when the difference $\theta >0$, but it is less important to distinguish between $\theta =0$ and $\theta <0$. It has low power to reject H_0 when $\theta = -\theta_1$ but gains by having smaller values of $E(V^*;\theta)$, $MED(V^*;\theta)$ and $P90(V^*;\theta)$ when $\theta <0$ than when $\theta >0$. This asymmetric power is appropriate for most sibling studies as

a genetic etiology is expected to make monozygotes more alike than dizygotes while situations where dizygotes are more alike than monozygotes are rare and genetically implausible.

The stopping limits and other properties of the triangular procedure are premised on the assumption of Brownian motion with drift implying that the changes in the test statistic Z at the i th inspection interval, $\delta Z_i = Z_i - Z_{i-1}$; $i = 1, 2, \ldots$; $Z_0 = 0$, are independent and normally distributed with mean $\theta \delta V_i$ and variance δV_i, where δV_i is the corresponding change in V. In the case of both Wald and the score procedures, the normal distribution of the test statistics is a large sample result. The implications of this on the error rates of the sequential test of a difference in the pairwise associations between two independent groups are next investigated by simulations.

Error rates

The error rates of the sequential W_e, W_v and the W_u procedures were assessed using triangular tests with a one-sided type I error rate of 0.025 and power of 0.90 to detect $\theta = \theta_1$. The two straight line boundaries of the test are $Z = a + cV$ and $Z = -a + 3cV$, where the intercept $a = 12.33$ and $c = 0.12$. The constants a and c are evaluated under the Brownian motion assumption. The results of 10 000 fold simulations, over a range of values for η_Y and η_X, and under both the null ($\theta = 0$) hypothesis and when $\theta = \theta_1$ are reported in Table I. For each simulated study, interim analyses were conducted after every five pairs of observations randomly drawn from each of two independent bivariate normal distributions.

Columns α and $1 - \beta$ in Table I give the proportions of times the null hypothesis was rejected under $\theta = 0$ and θ_1, respectively, for each of the three procedures under consideration. It should be noted that for data generated under $\theta = \theta_1$, the very few false rejections of the null at the lower limits of the sequential test are not counted when estimating the power. The starred values in the table lie outside the approximate 95 per cent tolerance intervals (0.02188, 0.02812) and (0.89400, 0.90600) for the simulated type I error and power, respectively. These intervals were calculated using the normal approximation to the distribution of the proportion of times the null hypothesis was rejected.

All three tests preserve the type I error rate in designs with small θ_1. This rate is however inflated for the W_e and W_u tests as θ_1 gets larger. The achieved powers for each of the three procedures are close to the design rates when $\eta_X = 0.1$. For $\eta_X = 0.3$ or 0.5, the W_e test shows a tendency for reduced power when used with triangular tests designed to detect larger values of θ. Power is however maintained for the W_v and W_u tests. The observed differences in the error rates of the three procedures are most likely a result of variations in their rates of convergence. From this simulation assessment we conclude that the W_v test statistics will be better than both the W_u and W_e statistics in preserving the error rates of the triangular design. In the remainder of this paper we concentrate on the sequential use of the W_v statistic and note that the sequential W_u procedure will, in most cases, lead to similar

conclusions.

Table I. Achieved type I error rate α and power 1-β for triangular test procedures with one sided type I error of 0.025 and power of 0.90 to detect the difference θ_1 using Wald and score statistics for no difference in the pairwise associations of the two independent bivariate groups.

η_X	θ_1	W_e test		W_v test		W_u test	
		α	1 - β	α	1 - β	α	1 - β
0.1 (0.01)	0.10 (0.10)	0.0246	0.9031	0.0245	0.9039	0.0249	0.9036
	0.20 (0.19)	0.0251	0.9008	0.0256	0.9011	0.0245	0.9017
	0.30 (0.28)	0.0262	0.9011	0.0244	0.9017	0.0255	0.9019
	0.40 (0.36)	0.0264	0.8953	0.0231	0.9063*	0.0269	0.9000
	0.50 (0.44)	0.0317*	0.8988	0.0243	0.9062*	0.0284*	0.9035
	0.60 (0.50)	0.0342*	0.8959	0.0262	0.9060	0.0316*	0.8988
	0.70 (0.56)	0.0359*	0.8944	0.0267	0.9026	0.0294*	0.8948
0.3 (0.29)	0.10 (0.09)	0.0248	0.8946	0.0261	0.8963	0.0247	0.8950
	0.20 (0.17)	0.0258	0.9032	0.0249	0.8985	0.0263	0.9048
	0.30 (0.25)	0.0282*	0.8963	0.0260	0.8982	0.0287*	0.8919*
	0.40 (0.31)	0.0305*	0.8979	0.0248	0.9029	0.0270	0.8989
	0.50 (0.37)	0.0263	0.8943	0.0295*	0.9033	0.0295*	0.8952
	0.60 (0.42)	0.0335*	0.8883*	0.0246	0.9076*	0.0322*	0.8981
	0.70 (0.47)	0.0364*	0.8935*	0.0266	0.9069*	0.0343*	0.8941
0.5 (0.46)	0.10 (0.07)	0.0232	0.8985	0.0230	0.8959	0.0229	0.8989
	0.20 (0.14)	0.0256	0.9006	0.0263	0.8994	0.0243	0.9005
	0.30 (0.20)	0.0263	0.8972	0.0255	0.9032	0.0247	0.8993
	0.40 (0.25)	0.0289*	0.8938*	0.0273	0.8992	0.0265	0.8984
	0.50 (0.30)	0.0285*	0.8927*	0.0252	0.9048	0.0281	0.8993
	0.60 (0.34)	0.0340*	0.8899*	0.0250	0.9022	0.0321*	0.8974
	0.70 (0.37)	0.0340*	0.8905*	0.0237	0.9056	0.0313*	0.8995

* Values which lie outside the approximate 95 per cent tolerance intervals given in the text. In braces are the corresponding correlation coefficients ρ_X and differences $\rho_Y - \rho_X$.

Study size

In a sequential study with an equal number of pairs of observations in either group, let the random variable N^* denote the number of pairs per group when the procedure terminates. At any point in the sequential W_v procedure with an equal number of observations n per group we clearly have $V = (n-3)/2$. It follows therefore that the relationship between N^* and the terminal information V^* is given by

$$E_\theta(N^*) = 2E_\theta(V^*) + 3 \qquad (5)$$

where $E_\theta()$ denotes expectation under θ. At the design stage $E_\theta(V^*)$ can be computed and used in estimating the requirements of the sequential study.

Using simulations, the study size of the balanced sequential W_v procedure is evaluated for a triangular design with one-sided type I error rate of 0.025 and power of 0.90 to detect $\theta = \theta_1$. Results are listed in Table II. As in the assessment of error rates, interim analyses were conducted after every five pairs of observations and the reported results are based on 10 000 fold simulations.

Table II. Average terminal number of pairs in each group and 90th percentile from 10 000 fold simulations under the null ($\overline{N}_0^*$, $P90(N_0^*)$), the alternative ($\overline{N}_1^*$, $P90(N_1^*)$) and at $\theta_{\max}$ ($\overline{N}_{\max}^*$, $P90(N_{\max}^*)$)for triangular tests with power=0.90 to detect the difference θ_1, one sided α =0.025 and using Wald W_v statistic for no difference in the pairwise associations of two independent bivariate groups.

θ_1	$\theta_{\max}$	$E(N_0^*)$	$\overline{N}_0^*$	$P90(N_0^*)$	$E(N_1^*)$	$\overline{N}_1^*$	$P90(N_1^*)$	$E(N_{\max}^*)$	$\overline{N}_{\max}^*$	$P90(N_{\max}^*)$	$n_f(W_v)$
0.10	0.059	1058.4	1060.3	1695	1275.9	1276.1	2035	1523.9	1529.2	2295	2105
0.20	0.119	267.8	267.3	430	322.1	321.1	515	384.1	385.5	580	528
0.30	0.178	121.4	121.2	195	145.5	144.1	230	173.1	182.5	260	237
0.40	0.237	70.1	69.9	110	83.7	83.4	130	99.2	99.3	150	134
0.50	0.297	46.4	46.9	70	55.1	55.1	85	65.0	65.4	95	87
0.60	0.356	33.5	33.8	50	39.6	39.4	60	46.5	46.5	70	61
0.70	0.416	25.8	25.9	40	30.2	30.1	45	35.3	35.6	50	46

Also included are the equivalent expected values and fixed sample estimates $n_f(W_v)$.

The average number of pairs of data in each group is denoted by N^*. Its values under the null ($\theta = 0$) and alternative ($\theta = \theta_1$) are subscripted in the table by 0 and 1, respectively. The average when $\theta = \theta_{\max}$ is denoted $\overline{N}_{\max}^*$, where $\theta_{\max}$ is the value for θ at which the expected sample size for the test procedure is a maximum, as determined by PEST 4. Evaluated using (5) and presented for comparative purposes are the corresponding expected values. Also included in the table are the numbers of pairs in each group $n_f(W_v)$ required by the equivalent fixed sample tests.

The results show good adherence between the averages $\overline{N}^*$ and the expectations $E(N^*)$ for each of the three values for θ considered and support the use of (5) in assessing the study requirements. In terms of efficiency relative to the fixed sample design we see that, when the null hypothesis is true, the average number of pairs of data could be reduced to almost one-half the estimated number for the equivalent fixed sample test. If the alternative is true this number could still be more than one-third less the fixed sample estimate while at $\theta_{\max}$ the average number of pairs is around one-quarter less the equivalent fixed sample test. When the null hypothesis is true, the 90th percentile of N^* is slightly more than 80 per cent of the corresponding fixed sample value. Under the alternative hypothesis these two values are similar, with the 90th percentile in the sequential case being slightly smaller. At $\theta_{\max}$, the 90th percentile exceeds the fixed sample value by a little more than 10 per cent. Overall the simulation assessment suggests that the error rates of the sequential W_v procedure will be

preserved and significant savings may be achieved from use of a sequential study design. Often it may be desirable or more practical to use a study design with $n_X \neq n_Y$, where n_X and n_Y are the number of pairs in the first and second group, respectively. In this case, by writing $n_Y = \kappa \times n_X$ for some positive constant κ, we can show that $E_\theta(N_X^*)$, the expected terminal number of pairs in the first group, is approximately equal to $[(1+\kappa)^2 E(V^*) + 3(1+\kappa^2)]/[\kappa(1+\kappa)]$. This slightly underestimates the true expected value but is nevertheless reasonable for anticipating the sequential study requirements.

Test Statistics for the Sequential Twin Study

In this section the sequential W_V procedure is adapted for a comparison of the difference between the trait associations of DZ and MZ twins. The procedure here is premised on the property that each data pair can be treated as a random sample from a bivariate distribution with identical marginal densities. In other words, rather than distinguishable pairs with two different means and two different variances, we now have pairs with common means and common variances. Let (x_{1j}, x_{2j}) denote the observed trait for the j th of n_{DZ} pairs of DZ twins and suppose that this observation is from a bivariate normal distribution with correlation coefficient ρ_{DZ}, with both means taking the same value μ_{DZ} and with both variances taking the same value σ_{DZ}^2. Similarly let (y_{1k}, y_{2k}) denote the observed trait for the k th of n_{MZ} pairs of MZ twins from a second independent bivariate normal distribution with correlation coefficient ρ_{MZ}, means μ_{MZ} and variances σ_{MZ}^2. It can then be shown that the maximum likelihood estimates of these parameters are

$$\hat{\mu}_{DZ} = \sum_{ij} x_{ij} /(2n_{DZ}), \hat{\sigma}_{DZ}^2 = \sum_{ij} (x_{ij} - \hat{\mu}_{DZ})^2 /(2n_{DZ})$$

and

$$\hat{\rho}_{DZ} = \sum_j (x_{1j} - \hat{\mu}_{DZ})(x_{2j} - \hat{\mu}_{DZ})/(n_{DZ}\hat{\sigma}^2) \tag{6}$$

with similar expressions for the maximum likelihood estimates $\hat{\mu}_{MZ}$, $\hat{\sigma}_{MZ}^2$ and $\hat{\rho}_{MZ}$. These estimates are derived in Appendix A. One noteworthy feature of the maximum likelihood estimates is that they are symmetric in the designations '1' and '2' used to distinguish twins. Using the notation introduced earlier, denote by θ the difference in Fisher's z-transformation η_{MZ} of the correlation coefficient for a MZ pair and the transformation η_{DZ} of the correlation coefficient for a DZ pair, and suppose that we are interested in the test of the null hypothesis H$_0$: θ = 0 versus the alternative hypothesis H$_1$: θ >0. Then following from the preceding section, at the i th interim analysis the sequential W_v procedure compares the statistics $Z_i = (n_{DZ}^{-1} + n_{MZ}^{-1})^{-1}\hat{\theta}$ with $V_i^{-1} = (n_{DZ} - 3)^{-1} + (n_{MZ} - 3)^{-1}$, where n_{DZ} and n_{MZ} are the respective numbers of DZ and MZ pairs observed so far and $\hat{\theta}$ is evaluated based on (6). A simulation assessment of the sequential procedure based on these statistics yielded results (not presented) that are very similar to the results obtained for the

general case in the previous section on the Wald and Score statistic.

The maximum likelihood estimate (6) of the correlation coefficient is a special case of the intraclass correlation[17] used in fixed sample genetic studies to describe trait similarity between siblings. An important advantage of using this estimate in the analysis of twin data is that estimation of the correlation coefficient does not require distinguishing between members of each pair. Our test procedure is essentially a sequential form of the method[18] for comparing the intraclass correlations of MZ and DZ twins. The difference between the two tests is the variance estimate; the method in Reference 18 uses the asymptotic variance of the maximum likelihood estimate of Fisher's z-transformation (see Section 3.1) rather than the modification for improved asymptotic normality. Simulation results (not presented) show that the sequential procedure using the asymptotic variance of the maximum likelihood estimate is not as effective as either the W_v or W_u procedures in preserving the design error rates.

The effect of any vector of covariates t can be allowed for by its inclusion in the mean. For example assuming a linear relation, the mean for the i th twin in the j th DZ pair, who has covariate value $t_{DZ_{ij}}$, becomes $\mu_{DZ_{ij}} = \mu_{DZ} + \gamma\, t_{DZ_{ij}}$, where the vector γ is used to denote the covariate effect. Similarly the mean for the i th twin in the k th MZ pair becomes $\mu_{MZ_{ij}} = \mu_{MZ} + \gamma\, t_{MZ_{ij}}$. While this is a straightforward and intuitively appealing approach, it is now much more difficult to obtain closed forms of the parameter estimates and a numerical routine is preferred to maximize the likelihood. The inclusion of covariates in this manner is in principle similar to the generally accepted fixed -sample approach[19] where the covariate effect is eliminated by performing the correlation analysis on the residuals from a regression model of trait on covariates. This approach of including any covariate effect in the mean may also be used when distinguishable pairs of relatives, such as mothers and daughters, are investigated.

Example

We illustrate the sequential twin procedure in a heritability study that allows for the effect of body mass index (BMI), known to be associated with breast dysplasia. Singleton sister pairs were used as an accepted proxy for dizygotic twins. Although this introduces possible effects of birth order and not sharing the womb, these effects, if any, are believed to be small and do not affect the validity of our sequential procedure. The actual study is ongoing and will therefore not be presented in this illustrative example of the methodology. Instead we demonstrate our sequential procedure using a total of 19MZ and 76 singleton pairs identified in the DOM cohort.

Measurements of dense tissue were obtained from the mammograms in the DOM cohort. The triangular test with one-sided α =0.025 and power of 0.90 to detect the difference θ =0.4 is used with planned inspections after every pair of MZ twins and every four pairs of singleton sisters. The effect size θ =0.4 was chosen based on results obtained in Reference 7; the hypothesized correlations for MZ twins and singletons are 0.6 and 0.3, respectively. The

inspection interval maximizes usage of the MZ twins and singletons data known to be available and is estimated in terms of information to be $\tilde{V}_{insp}$ =0.80. The expected number of observations required by the test is approximately 44 pairs of MZ twins and 176 pairs of singletons if the null hypothesis is true and 53 pairs of MZ twins and 212 pairs of singletons if the alternative is true. The equivalent fixed sample test requires 85 pairs of MZ twins and 340 pairs of singletons. Note that with inspections planned after every n_X and κn_X pairs, respectively, in the two groups, $\tilde{V}_{insp}$ is approximately $n_X \kappa /(\kappa+1)$.

Unfortunately, the current available data are insufficient to lead to a conclusion of the sequential test. So, in order to present a full worked example, the rest of the study was generated by random re-sampling from the available data. In the first example of the sequential heritability analysis, our trait of interest was the absolute measure of dense tissue. The correlation estimate between pairs of sisters was adjusted for the effect of BMI using our approach given in the preceding section. Interim analyses were conducted in the order in which the sets were identified and the resulting sequential test is pictured in Figure 1. The procedure ended after 35 MZ and 140 singleton pairs with the decision that sufficient information against the null hypothesis had been observed. This is less than one-half (0.41) the observations required by the equivalent fixed sample procedure. Using an analysis adjusted for the sequential test[16], the reported p-value of the test is 0.003. At stopping Z =15.70, V =25.94 and the median unbiased estimate for the difference θ is 0.56 with 95 per cent confidence interval (0.166, 0.956). Using this estimated difference and the BMI adjusted maximum likelihood estimate $\hat{\rho}_{DZ}$ =0.37 of the correlation coefficient of the singleton sisters, we estimate $\hat{\rho}_{MZ}$ =0.74 and get an approximate estimated heritability of $2(\hat{\rho}_{MZ} - \hat{\rho}_{DZ})$ =0.74.
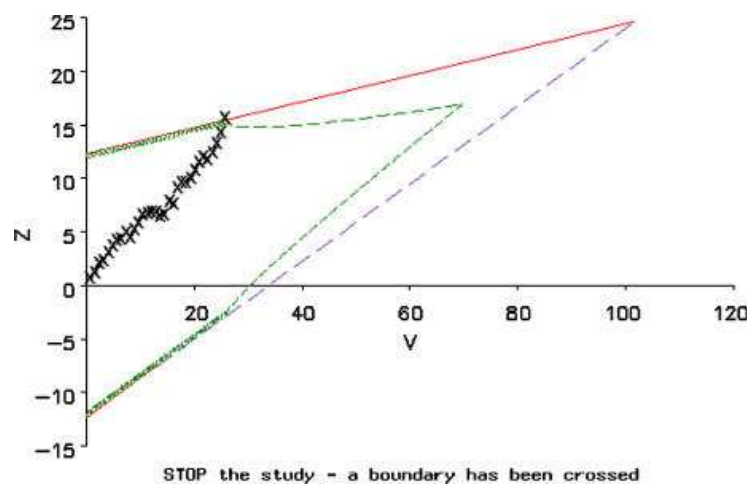

Figure 1. Sample path of the sequential Wald test of heritability of breast dysplasia, adjusted for the effect of BMI, and using monozygotic twins and pairs of sisters.
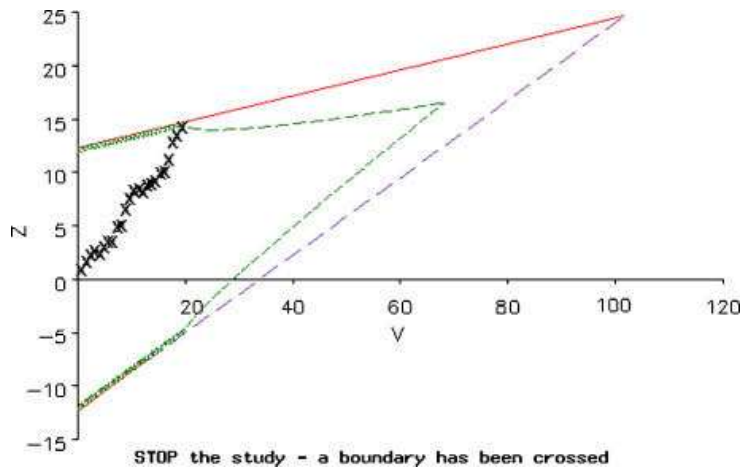
Figure 2. Sample path of the sequential Wald test of heritability of breast dysplasia based on the proportion of dense breast tissue and using monozygotic twins and pairs of sisters.

Purely for comparative purposes we performed a second analysis using the proportion of dense tissue. The combination of various influences into this one measure may mask the univariate effect of dense tissue (and other factors)[20]. However the proportion is the historical and still most commonly used measure in publications on breast density. The sample path of this analysis is shown in Figure 2. The analysis ended after 27 MZ and 108 singleton pairs with a similar decision, but earlier than the above test using absolute measures. At stopping Z = 14.20, V = 19.53, the median unbiased estimate of θ is 0.71 with 95 per cent confidence interval (0.257, 1.155), $\hat{\rho}_{DZ}$ = 0.22, $\hat{\rho}_{MZ}$ = 0.73 and the estimated heritability is 1.02. This second analysis yielded a smaller estimate of the correlation between singleton pairs and a larger estimate of heritability than was obtained in the first analysis. The results obtained here are quite similar to those obtained in References 7 & 20.

Discussion

In this paper we have extended the sequential association test procedure with dichotomous observations to a comparison of the pairwise associations of two groups of continuous observations. We have shown that the sequential test can lead to a substantial reduction in the number of observations over the equivalent fixed sample procedure. This can be highly desirable when cost and efficiency are prominent concerns of the investigator. The sequential score approach is shown to be comparable to the modified Wald statistic in type II error but is slightly less accurate in type I error for larger values of the parameter of interest, given by the difference in the z-transformed correlation coefficients of the two populations. There is no obvious way to modify the score statistic to improve these results. As the score statistic is also less familiar and somewhat more complicated than the W_v statistic, we recommend using the W_v statistic in the sequential test.

The W_v test statistic Z is based on the Fisher's z-transformation of the correlation coefficient. The use of this statistic in a fixed sample setting is very well established. In particular, the form of the test statistic used to analyse twin data is related to the intraclass correlation estimate used to describe the similarity in genetic traits between two members of the same

family. Fixed sample tests for the correlation coefficient in more general settings have been well studied in the literature; in this paper we introduced the alternative to a fixed sample test of the correlation coefficient. Our sequential twin procedure is very closely related to the fixed sample method for comparing the intraclass correlations of MZ and DZ twins. For instance, the maximum likelihood estimate of the correlation coefficient from the bivariate normal distribution with common means and common variances was shown to equal the intraclass correlation estimate.

The work detailed in this paper was motivated by a need for efficient usage of available information. In this regard the triangular test is approximately optimal in that it minimizes the maximum expected terminal information among all procedures with the same significance level and power[21]. It is therefore effective in reducing the study size particularly when the true association is between the null value ($\theta = 0$) and the power value (θ_1). The efficiency of the triangular test relative to the fixed sample procedure was demonstrated in the simulations. Details on the merits of the triangular and other sequential test procedures can be found in Reference 8.

In keeping with the motivating example, the sequential test of a difference in the associations was assessed using a one-sided design. There is however no obvious reason why the results obtained here will not be equally valid for double-sided tests. In simulations, interim analyses were conducted after an equal number of observations in each group. As the example showed, in practice we can easily envisage situations where the accrual rate is not as homogeneous and where it may be more appropriate to time the interim analyses differently. While this will have some impact on the efficiency of the sequential procedure, it is not envisaged that its effect will give cause for concern.

Twin studies can usually be placed within the class of familial aggregation methods used as a first step in pursuing a possible genetic etiology. In studies of rare conditions or where phenotypic data are costly or difficult to obtain or require utilization of stored biological samples, the use of a sequential study design can lead to considerable savings in valuable resources and a dramatic decrease in the cost and time of the study[2, 3]. In this paper we have introduced a sequential method for analysing the classical twin study for quantitative traits with and without covariates. The classical twin method has been criticized for its underlying assumptions of random mating, no interactions between genes and environment, equivalent environments for MZ and DZ twins, no sibling interaction and no genetic dominance[22]. We concede that while these and other limitations exist, the extent to which they matter depends on the trait being studied and they do not obviate the usefulness of twin studies. For instance, for traits that are substantially influenced by heredity, the approximately two-fold difference in genetic similarity between the two types of twins should outweigh any complications arising from failure to satisfy the assumptions. In addition, all of these assumptions can be tested, given the proper data. A sequential study design is particularly suited in this regard as it provides the opportunity to evaluate the assumptions as the study progresses; subsequent work focuses on extending the sequential twin study to more general settings and beyond the classical twin design.

Acknowledgements

Appendix A

The log-likelihood

Suppose that we observe n_X random pairs from a bivariate normal distribution with means μ_{X_1}, μ_{X_2} variances $\sigma_{X_1}^2$, $\sigma_{X_2}^2$ and correlation coefficient ρ_X and we observe a second independent set of n_Y pairs from a bivariate normal distribution with means μ_{Y_1}, μ_{Y_2} variances $\sigma_{Y_1}^2$, $\sigma_{Y_2}^2$ and correlation coefficient ρ_Y. Expressed in terms of the Fisher's z-transformation, see equation (1), the joint log likelihood ℓ is the sum $\ell_X + \ell_Y$, where for U =X or Y,

$$
\ell_U = -n_U \ln(2\pi) + n_U \ln \cosh \eta_U - \frac{n_U}{2}[\ln(\sigma_{U_1}^2) + \ln(\sigma_{U_2}^2)]
$$

$$
-\frac{1 + \cosh 2\eta_U}{4} \sum_{j=1}^{n_U} \left[\frac{(u_{1j} - \mu_{U_1})^2}{\sigma_{U_1}^2} + \frac{(u_{2j} - \mu_{U_2})^2}{\sigma_{U_2}^2} \right] \tag{A1}
$$

$$
+\frac{\sinh 2\eta_U}{2} \sum_{j=1}^{n_U} \left[\frac{(u_{1j} - \mu_{U_1})(u_{2j} - \mu_{U_2})}{\sigma_{U_1} \sigma_{U_2}} \right]
$$

is the likelihood of the n_U pairs (u_{1j}, u_{2j}); j =1,..., n_U. Let $\theta = \eta_Y - \eta_X$ and suppose that we are interested in the test of the null hypothesis θ =0. The efficient score, Z and observed Fisher's information, V for θ evaluated under the null hypothesis are first derived. Next, the variance of Wald procedure $I(\hat{\theta}, \hat{\psi})$ is obtained using the method introduced for evaluating the observed information.

Efficient score and observed Fisher's information

To express ℓ in terms of the parameter of interest θ we define a new parameter $\phi = \eta_X + \eta_Y$. This gives the 1–1 transformation $\eta_Y = (\phi + \theta)/2$ and $\eta_X = (\phi - \theta)/2$ and leads to the log likelihood ℓ now a function of θ and the nine nuisance parameters ϕ and ξ, where ξ denotes the vector of means and variances of the likelihood. To obtain Z and V we employ a profile likelihood approach, where the nuisance parameters are replaced in the likelihood by their maximum likelihood estimates. The statistic Z is the first derivative of the profile likelihood evaluated at $\theta = 0$, or equivalently, the first partial derivative of ℓ with respect to θ evaluated at $\theta = 0$ and ($\hat{\phi}_0$, $\hat{\xi}_0$), where $\hat{\phi}_0$ and $\hat{\xi}_0$ denote the maximum likelihood estimates of ϕ and ξ, respectively, when θ is zero. Likewise V is the negative of the second derivative of the profile likelihood evaluated at $\theta = 0$, or equivalently, the negative reciprocal of the (1, 1)th element of the inverse of the matrix of second derivatives of ℓ evaluated at $\theta = 0$ and ($\hat{\phi}_0$, $\hat{\xi}_0$) [8].

For U =X, Y let $\hat{\sigma}^2_{U_v} = \sum_{j=1}^{n_U} (U_{vj} - \overline{U}_v)^2 / n_U, v = 1, 2$ denote the maximum likelihood estimates of the four variances in the likelihood and let $\hat{\rho}_U \sum_{j=1}^{n_U} (u_{1j} - \overline{u}_1)(u_{2j} - \overline{u}_2) / n_U \hat{\sigma}_{U_1} \hat{\sigma}_{U_2}$ denote the maximum likelihood estimate for ρ_U derived from ℓ_U. Using the chain rule for derivatives it can be shown that the efficient score Z is given by Z =(Z$_Y$ –Z$_X$)/2, where $Z_U = n_U (\hat{\rho}_U - \hat{\rho})/(1 - \hat{\rho}\hat{\rho}_U)$, U =X, Y, and $\hat{\rho}$ is given by (4) and is the estimate of the common correlation coefficient hypothesized under the null. Let the matrix of second derivatives of ℓ, called the hessian, be denoted by H. It can be shown that, when evaluated at $\theta = 0$ and ($\hat{\phi}_0$, $\hat{\xi}_0$), all second-order mixed derivatives of the likelihood ℓ_X involving derivatives with respect to one of its means μ_{X_1} or μ_{X_2} and one of its other parameters $\sigma^2_{X_1}$, $\sigma^2_{X_2}$ or η_X is zero, and similarly for ℓ_Y. To take advantage of this when inverting H, we partition it as

$$\begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \tag{A2}$$

where H$_{11}$ is the 6×6 matrix of all second-order derivatives of ℓ with respect to parameters from θ, ϕ, $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, $\sigma^2_{Y_1}$ and $\sigma^2_{Y_2}$; H$_{12}$ is the 6×4 matrix of all second-order mixed derivatives involving one parameter from (θ, ϕ, $\sigma^2_{X_1}$, $\sigma^2_{X_2}$, $\sigma^2_{Y_1}$, $\sigma^2_{Y_2}$) and one from (μ_{X_1} μ_{X_2}, μ_{Y_1}, μ_{Y_2}); H_{22} is the 4×4 matrix of the second-order derivatives with respect to parameters from μ_{X_1} μ_{X_2}, μ_{Y_1}, μ_{Y_2} and H_{21} is the transpose of H_{12}. As the evaluated H_{12} is a null matrix we get

$$H^{-1} = \begin{bmatrix} H_{11}^{-1} & 0 \\ 0 & H_{22}^{-1} \end{bmatrix}$$ (A3)

and, to obtain V, we are now only required to invert and evaluate H_{11} when $\theta = 0$ and at the point $(\hat{\phi}_0, \hat{\xi}_0)$.

Partitioning H_{11} as

$$\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$ (A4)

where B_{11} is the 2 × 2 matrix of all second-order derivatives of ℓ with respect to θ and ϕ; B_{12} is the 2 × 4 matrix of all second-order mixed derivatives involving one parameter from (θ, ϕ) and one from $(\sigma_{X_1}^2, \sigma_{X_2}^2, \sigma_{Y_1}^2, \sigma_{Y_2}^2)$; B_{22} is the 4 × 4 matrix of all possible second-order derivatives with respect to $(\sigma_{X_1}^2, \sigma_{X_2}^2)$ and $(\sigma_{Y_1}^2, \sigma_{Y_2}^2)$ and B_{21} is the transpose of B_{12}, its inverse now becomes

$$H_{11}^{-1} = \begin{bmatrix} (B_{11} - B_{12}B_{22}^{-1}B_{21})^{-1} & -B_{11}^{-1}B_{12}(B_{22} - B_{21}B_{11}^{-1}B_{12})^{-1} \\ -B_{22}^{-1}B_{21}(B_{11} - B_{12}B_{22}^{-1}B_{21})^{-1} & (B_{22} - B_{21}B_{11}^{-1}B_{12})^{-1} \end{bmatrix}$$ (A5)

and the derivation of V reduces to inverting $(B_{11} - B_{12}B_{22}^{-1}B_{21})$.

For U =X, Y let B_{11U} be the second-order derivatives of ℓ_U with respect to θ and ϕ and denote by r_U the ratio $(1 - \hat{\rho}^2) / (1 - \hat{\rho}\hat{\rho}_U)$. Then we can write $B_{11} = B_{11X} + B_{11Y}$ where, evaluated at $\theta = 0$ and $(\hat{\phi}_0, \hat{\xi}_0)$,

$$B_{11X} = \frac{n_X}{4}[1 - \hat{\rho}^2 - 2r_X(\cosh\hat{\phi}_0 - \hat{\rho}_X \sinh\hat{\phi}_0)]\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$ (A6)

and

$$B_{11Y} = \frac{n_Y}{4}[1 - \hat{\rho}^2 - 2r_Y(\cosh\hat{\phi}_0 - \hat{\rho}_Y \sinh\hat{\phi}_0)]\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$ (A7)

Next, consider the product $B_{12}B_{22}^{-1}B_{21}$. Its evaluation is simplified if we partition each of the matrices B_{12} and B_{22} into two components, one involving second derivatives with respect to arguments of ℓ_X only and the other involving the identical terms for ℓ_Y. The 2 × 4 matrix B_{12} can now be written as $[C_X C_Y]$ where

$$C_U = \begin{bmatrix} \partial^2 \ell_U / \partial \sigma_{U_1}^2 \partial \theta & \partial^2 \ell_U / \partial \sigma_{U_2}^2 \partial \theta \\ \partial^2 \ell_U / \partial \sigma_{U_1}^2 \partial \phi & \partial^2 \ell_U / \partial \sigma_{U_2}^2 \partial \phi \end{bmatrix}, \quad U = X, Y \tag{A8}$$

while the 4 x 4 matrix B_{22} becomes

$$B_{22} = \begin{bmatrix} D_x & 0 \\ 0 & D_Y \end{bmatrix} \tag{A9}$$

where

$$D_U = \begin{bmatrix} \partial^2 \ell_U / \partial (\sigma_{U_1}^2)^2 & \partial^2 \ell_U / \partial \sigma_{U_1}^2 \partial \sigma_{U_2}^2 \\ \partial^2 \ell_U / \partial \sigma_{U_1}^2 \partial \sigma_{U_2}^2 & \partial^2 \ell_U / \partial (\sigma_{U_2}^2)^2 \end{bmatrix}, \quad U = X, Y$$
(A10)

This partitioning yields $B_{12} B_{22}^{-1} B_{21} = C_X D_X^{-1} C_X' + C_Y D_Y^{-1} C_Y'$ where, at $\theta = 0$ and ($\hat{\phi}_0$, $\hat{\xi}_0$),

$$C_X D_X^{-1} C_X' = -\frac{n_X}{4} r_X^2 (\sinh \hat{\phi}_0 - \hat{\rho}_X \cosh \hat{\phi}_0)^2 \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$
(A11)

and

$$C_Y D_Y^{-1} C_Y' = -\frac{n_Y}{4} r_Y^2 (\sinh \hat{\phi}_0 - \hat{\rho}_Y \cosh \hat{\phi}_0)^2 \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$
(A12)

Finally subtracting this sum from B_{11} and inverting we get

$$\frac{1}{V} = \frac{1}{V_X} + \frac{1}{V_Y}$$
(A13)

where

$$V_U = n_U \left[1 - \left(\frac{\hat{\rho}_U - \hat{\rho}}{1 - \hat{\rho} \hat{\rho}_U} \right)^2 \right], \quad U = X, Y$$
(A14)

The expected Fisher's information

The evaluation of $I(\theta, \psi = \{\phi, \xi\})$, (the negative of the $(1,1)$ th element of the inverse of the expected value of the hessian H), follows a similar pattern as above. Using $E[\sum_j (U_{vj} - \mu_{Uv})] = 0$; $v = 1, 2$; U = X, Y we get $E[H_{12}]$ is null and , as before, it is only necessary to invert $E[B_{11} - B_{12}B_{22}^{-1}B_{21})]$. Next, since $E[\sum_j (U_{vj} - \mu_{U_v})^2] = n_U \sigma_{U_v}^2$; $E[\sum_j (U_{1j} - \mu_{U_1})(U_{2j} - \mu_{U_2})] = n_U \rho_U \sigma_{U_1} \sigma_{U_2}$ we now get $E[B_{11X}]$ given by (A6) with $r_X = 1$ and $\hat{\rho}^2, \hat{\rho}_X, \hat{\phi}_0$ replaced, respectively, by ρ_X^2, ρ_X, $\phi - \theta$ while $E[B_{11Y}]$ is given by (A7) with $r_Y = 1$ and $\hat{\rho}^2$, $\hat{\rho}_Y$, $\hat{\phi}_0$ replaced, respectively, by ρ_Y^2, ρ_Y and $\phi + \theta$. $E[C_X D_X^{-1} C_X]$ and $E[C_Y D_Y^{-1} C_Y]$ are similarly given by (A11) and (A12), respectively, and we get $1/I(\theta, \phi, \xi) = 1/n_X + 1/n_Y$, which do not depend on the likelihood parameters.

Maximum likelihood estimation using bivariate normal samples with common means and variances

Suppose we observe the n pairs of observations $(x_{11}, x_{21}), ..., (x_{1n}, x_{2n})$ from a bivariate normal distribution with equal means μ , equal variances σ^2 and correlation coefficient ρ . Equating to zero the derivative of the log-likelihood with respect to μ we get $(1 - \hat{\rho})/(1 - \hat{\rho}^2) \sum_{i=1}^2 \sum_{j=1}^n (x_{ij} - \hat{\mu})/\hat{\sigma}^2 = 0$ whence $\hat{\mu} = \sum_{i=1}^2 \sum_{j=1}^n x_{ij} / 2n$ provided $\hat{\rho} \neq \pm 1$. Next the derivatives of the log likelihood with respect to σ^2 and ρ equated to zero can, respectively, be written as

$$2n(1 - \hat{\rho}^2) = \sum_{i=1}^2 \sum_{j=1}^n \left(\frac{x_{ij} - \hat{\mu}}{\hat{\sigma}} \right)^2 - 2\hat{\rho} \sum_{j=1}^n \frac{(x_{1j} - \hat{\mu})(x_{2j} - \hat{\mu})}{\hat{\sigma}^2}$$

(A15)

and

$$n\hat{\rho}(1 - \hat{\rho}^2) = \hat{\rho} \sum_{i=1}^2 \sum_{j=1}^n \left(\frac{x_{ij} - \hat{\mu}}{\hat{\sigma}} \right)^2 - (1 + \hat{\rho}^2) \sum_{j=1}^n \frac{(x_{1j} - \hat{\mu})(x_{2j} - \hat{\mu})}{\hat{\sigma}^2}$$

(A16)

which solved simultaneously gives $\hat{\sigma}^2 = \sum_{i=1}^2 \sum_{j=1}^n (x_{ij} - \hat{\mu})/2n$ and $\hat{\rho} = \sum_{j=1}^n (x_{1j} - \hat{\mu})(x_{2j} - \hat{\mu})/n\hat{\sigma}^2$, provided $\hat{\rho} \neq \pm 1$.

Reference List

1. Böddeker IR, Ziegler A. Sequential designs for genetic epidemiological linkage or association studies: a review of the literature. Biometrical Journal 2001; 43(4):501–525.
2. Van der Tweel I, van Noord PAH. Sequential analysis of matched dichotomous data from prospective case-control studies. Statistics in Medicine 2000; 19:3449–3464.
3. Dreyfus M, Hedelin G, Kutnahorsky R, Lehmann M et al. Antiphospholipid antibodies and preeclampsia: a case-control study. Obstetrics and Gynecology 2001; 97(1):29–34.
4. Van der Tweel I, Schipper M. Sequential tests for gene-environment interactions in matched case-control studies. Statistics in Medicine 2004; 23:3755–3771.
5. Baksh MF, Todd S, Whitehead J, Lucini MM. Design considerations in the sequential analysis of matched case-control data. Statistics in Medicine 2005; 24:853–867.
6. De Waard F, Collette HJA, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. Journal of Chronic Diseases 1984; 37(1):1–44.
7. Haars G, van Noord PAH, van Gils CH, Peeters PHM, Grobbee DE. Heritable aspects of dysplastic breast glandular tissue (DY). Breast Cancer Research and Treatment 2004; 87(2):149–156.
8. Whitehead J. The Design and Analysis of Sequential Clinical Trials (Revised 2nd edn). Wiley: Chichester, 1997.
9. Jennison C, Turnbull BW. Group Sequential Methods with Applications to Clinical Trials. Chapman & Hall/CRC: London, Boca Raton, FL, 2000.
10. Herings RM, Bakker A, Stricker BH, Nap G. Pharmaco-morbidity linkage: a feasibility study comparing morbidity in two pharmacy based exposure cohorts. Journal of Epidemiology and Community Health 1992; 46(2):136–140.
11. Zar JM. Biostatistical Analysis (4th edn). Prentice-Hall: Englewood Cliffs, NJ, 1999.
12. Bickel PJ, Doksum KA. Mathematical Statistics: Basic Ideas and Selected Topics. Holden-Day: California, 1977.
13. Cox D, Hinkley D. Theoretical Statistics. Chapman & Hall: London, 1974.
14. Fisher RA. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. Biometrika 1915; 10:507–521.
15. Brandner FA. A test of the significance of the difference of the correlation coefficients in normal bivariate samples. Biometrika 1933; 25(1/2):102–109.
16. MPS Research Unit. Pest 4: Operating Manual. University of Reading: Reading, 2000.
17. Armitage P. Correlation. In Biostatistical Genetics and Genetic Epidemiology, Elston R, Olston J, Palmer L (eds). Wiley: New York, 2002; 171–176.
18. Sham P. Statistics in Human Genetics. Arnold: London, 1998.
19. Thomas DC. Statistical Methods in Genetic Epidemiology. Oxford University Press: NY, 2004.
20. Haars G, van Noord PAH, van Gils CH, Grobbee DE, Peeters PHM. Measurements of breast density; no ratio for a ratio. Cancer Epidemiology, Biomarkers and Prevention 2005; 11:2634–2640.
21. Lai TL. Optimal stopping and sequential tests which minimise the maximum expected sample size. Annals of Statistics 1973; 1:659–673.
22. Neale MC. Twin Analysis. In Biostatistical Genetics and Genetic Epidemiology, Elston R, Olston J, Palmer L (eds). Wiley: New York, 2002; 743–756.

Measurements of breast density: no ratio for a ratio

Gerco Haars

Paul AH van Noord

Carla H van Gils

Diederick E Grobbee

Petra HM Peeters

3

Abstract

Breast density, as visible on mammograms, is generally assessed as the occupied percentage of the breast and is a risk factor for breast cancer. Various studies have looked into the causation and alteration of relative density but the relation of a determinant with a relative measure does not allow a direct etiologic interpretation. It was our goal to compare the effects of known determinants on relative density and the absolute amounts of dense and non-dense tissues. We measured the absolute and relative densities in a population of 418 postmenopausal women participating in a breast cancer screening program. The occupied surface area was calculated after manually tracing the contours of the tissues on digitized mammograms. Information on determinants was available through physical examination and questionnaires. Data were analyzed by multivariate linear regression. Age and parity were found to decrease the amount of dense tissue and the ages at menarche and menopause were found to increase it (R^2 = 13%). The amount of non-dense tissue was increased by higher body mass index (BMI), age, and parity (R^2 = 43%). Relative density was affected by a combination of these factors (R^2 = 29%) with directionalities of effects that are comparable to those of dense tissue. However, the magnitudes of these effects were the resultant of the effects on dense and non-dense tissues. The influence of BMI on relative density was completely due to an effect on non-dense tissue. Although relative density is a relevant prognostic factor, inferences about the aetiology of breast density should be made on the basis of absolute measures.

Introduction

The mammographic appearance of the female breast, or parenchymal pattern, is determined by the amounts of radio dense and non-dense tissues relative to each other. The radio dense tissue, which appears white on X-ray mammograms, has been shown to consist of stromal and epithelial tissues and thereby is assumed to reflect the target tissue for breast cancer[1, 2].

Wolfe[3] was the first to develop a classification of parenchymal patterns consisting of four categories: N, P1, P2, and DY. This classification was based on the overall appearance of the breast combined with the estimated percentage of the breast occupied by dense tissue with N for normal breasts with little or no dense tissue, P1 and P2 for intermediate states, and DY for a breast with predominant dense parenchymal tissue, which was referred to as dysplasia. Other classifications were subsequently developed, all using a ratio or percentage to classify the amount of dense tissue[4, 5] and with current technology the relative density is often measured on a continuous scale[6, 7].

High relative amounts of dense tissue have been consistently found to be related to elevated breast cancer risk by studies that used one or more categorical or continuous approaches[8-10]. The currently available continuous measurement of density has been found to be the most informative in this regard[11, 12]. Inspired by the established relation between relative density and breast cancer risk, many studies have tried to identify factors that cause and/or influence the parenchymal pattern. In these studies, factors including age, BMI, parity, passing through the menopause, and smoking have been shown to decrease the relative density. Factors including a late age at first birth, use of hormone therapy, and alcohol consumption have been found to increase the relative density[13-21].

Although the modern, continuous techniques determine the absolute amounts of total, dense, and non-dense tissues on the mammogram, investigators usually still work with the ratio of dense over total tissue, or in other words, the relative amount of dense tissue[13, 15, 18, 19]. The absolute amount of density was used in only a small number of studies, and of these, only the studies by Boyd et al. and Heng et al. made an inference about the difference in effects on absolute and relative densities[14, 17, 21].

The consequence of a relative measure is that for a given amount of dense tissue in a small breast (i.e., surrounded by little fat or other non-dense tissue) compared with a similar amount in a larger breast, a higher relative amount will be measured. This is a general methodological issue with ratios because they always combine the effects of the constituting measures[22]. The relative density therefore does not convey any information about the absolute amount of target tissue whereas, as hypothesized by Albanes and Winick[23] and Trichopoulos and Lipman[24], the actual amount of target cells may be a straightforward and important risk factor for cancer, although this was never conclusively shown.

The use of a ratio or percentage score thus only shows the effect of a determinant on the degree of density but obscures whether a determinant affects the absolute amount of dense tissue, the absolute amount of non-dense tissue, or both. The relative density may still be a useful and easily applicable prognostic factor as an indicator of breast cancer risk but would

not seem to be the measure of choice in etiologic research into the causes and determinants of breast density.

We set out to determine which of the established determinants of relative measurements of density indeed influence the absolute amount of dense tissue. The implications of the findings for breast cancer research are discussed.

Materials and Methods

Subjects

The Diagnostisch Onderzoek Mammacarcinoom (DOM) project is a population based breast cancer screening program in Utrecht, the Netherlands[25]. It was initiated in 1974 and up until 1986, the screening progressed in four sequential sub-cohorts, which were all of an experimental nature. At that time, 55,519 women had been recruited into the DOM study cohort. Apart from having their mammograms taken, all participants had anthropometric information collected by trained technicians and completed questionnaires on demographics, family-related information, history of disease, and reproductive history. To ensure uniformity and completeness of the follow-up, we chose only to include participants from the first sub-cohort (known as DOM-I, recruitment: 1974-1981) that had attended all screening rounds. For the final population for this study, we took a random sample of 500 women from which we excluded participants that were not definitively postmenopausal (82 women). This resulted in a population of 418 postmenopausal women for whom, at random, the right (n = 209) or left (n = 209) craniocaudal xero-print mammogram from the first screening round was retrieved. The first screening round was selected because most of the relevant other data were supplied at this point and because it again ensured the best uniformity. Xero-print mammograms preceded the currently applied X-ray film mammograms and differs by being a positive image. Although the contrast of current-film mammograms is higher than that of the older Xero-print mammograms, the latter show a great deal of detail and can easily be evaluated on the amount of radio dense tissue in a similar way as film mammograms are currently evaluated.

Measurement and Baseline Data

All xeromammograms were digitized at 100 dpi using a flatbed scanner (HP5300C Scanjet, Hewlett Packard, Palo Alto, CA). Digitized images were evaluated on the total breast size and amount of dense tissue by a trained observer (GH) who manually traced the edges of the areas by setting mouse clicks (Fig. 1). The surface of the selected area was calculated by the program Image-Xplorer (Image Sciences Institute, University Medical Center, Utrecht, the Netherlands). The calculated amounts of pixels were transformed into square centimeters on the basis of the pixel density (100 dots per inch = 10,000 dots on 6.45 cm^2). The absolute amount of non-dense tissue, which predominantly consists of fat, was calculated by subtracting the dense tissue from the total breast size. The percentage of the breast occupied by dense tissue, or relative density, was calculated by dividing the absolute amount

of dense tissue by the total breast size. Intra-reader reliability was determined by a 10% retest sample.

A number of characteristics known from the literature to be related to the relative amount of density or to breast cancer were examined for their relation with absolute and relative measures of breast density. These included anthropometric measures such as height, weight and BMI, lifestyle factors such as hormone therapy and smoking, and characteristics such as age, parity, and age at menopause. Data on some known determinants of breast density, such as race/ethnicity and insulin-like growth factor I levels, were not available in our cohort.
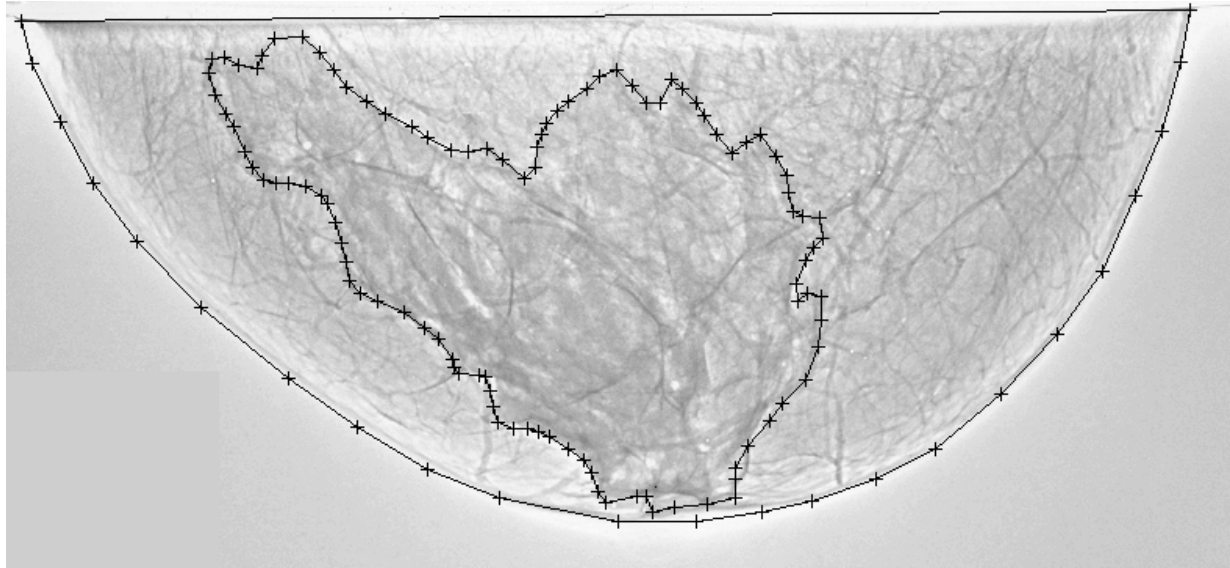

Figure 1. Manually traced contours of total breast size (outer contour) and dense tissue (inner contour) on a digitized xero-film mammogram.

Statistical Analysis

Geometric group means of dense tissue, non-dense tissue, and percent density were calculated in classes of determinants. Dichotomous classes applied to all Yes/No variables, such as family history and use of hormone therapy. Parity was used as an ordinal class, truncated at >3, and continuous variables were classified into tertiles, except for BMI, which was classified according to the generally accepted classes of underweight (<20), normal weight (20 to <25), overweight (25 to <30), and obesity (≥30). A test for linear trend was done on the geometric data through ANOVA with an F test for linearity. For use in multivariate regression analyses, the amount of dense tissue, non-dense tissue, and the percentage of dense tissue were transformed to normalize the data by taking the square root. Variables for which two or more of the measures under investigation had a clear trend over the classes were then simultaneously entered in multivariate regression models.

Mathematically related variables, such as weight and BMI or height, were never modelled simultaneously. The same applies to the age at examination and the combinations of the age at menopause and the time since the menopause and nulliparity (yes/no) and the number of children. The decision on which variables to use for multivariate modelling was made on the

basis of the P for trend and the magnitude of effect that was visible in the group means. In the multivariate analyses, variables were entered and removed on the basis of the combination of β's, p values, and overall model-fit in explained variance (R^2), but finally the models with the best fit were those with $p < 0.10$ and matching, relevant β's. The results of the final model were transformed back to the normal units.

Results

Participants selected had a mean age at the time of examination of 56.4 years (range, 49.2-65.8 years) and were all postmenopausal at the time of examination. General characteristics of the study population are given in Table 1. The median breast size in this population was 123.7 cm^2 (interquartile range, 94.7-162.6 cm^2). The median area of the breast occupied by

Table 1 Distribution of density and/or breast cancer risk factors in study population.

	Mean	SD	Range
Age at Examination (yr)	56.4	4.16	49.2 – 65.8
Weight (kg)	67.7	9.83	43.5 – 105.0
Height (m)	1.63	0.06	1.46 – 1.87
BMI (kg/m^2)	25.6	3.41	17.8 – 42.8
Age at Menarche (yr)[*]	13.5	1.57	11.0 – 18.3
Age at Menopause (yr)[§]	49.6	3.78	32. 0 – 60.0
Time since Menopause (yr)[§]	7.5	5.36	0.1 – 28.5
Parity (nr. of children)	2.4	1.91	0 – 9
Age at First Childbirth (yr)[†]	27.3	4.11	18.1 – 39.9

	N	Proportion	
Parous (yes)	332	79%	
Current OC use (yes)	8	2%	
Ever used OC (yes)	43	10%	
Used HT in last 12 months (yes)	47	11%	
Ever smoked (yes)	116	28%	
Family History of Breast Cancer (yes)[‡]	26	6%	

	Median	Interquartile range	Range
Breast size (cm^2)	123.7	94.7 – 162.6	29.1 – 345.3
Absolute amount dense tissue (cm^2)	26.4	15.2 – 46.4	0.3 – 138.0
Absolute amount non-dense tissue (cm^2)	87.8	55.6 – 130.6	4.5 – 343.8
Percent dense	23.2	11.6 – 43.6	0.2 – 90.0

[*] Data available for 416 women.
[§] Data available for 361 women.
[†] Including only parous women; n=332
[‡] Data available for 408 women.

dense breast tissue was 26.4 cm^2 (interquartile range, 15.2-46.6 cm^2) and the median area of the breast occupied by non-dense tissue was 87.8 cm^2 (interquartile range, 55.6-130.6 cm^2). The median percentage of dense tissue in the breast was 23.2% (interquartile range, 11.6-43.6%). Intra-reader reliability, as measured by the intraclass correlation coefficient, for relative density was 93%. For absolute amount of dense tissue, this was 82% and for non-dense tissue, 97%.

BMI, weight, age at the time of examination, the time since the menopause, and the number of children that a woman has had showed a negative univariate trend with the amount of dense tissue (p < 0.10). The age at which menopause occurred and having ever smoked showed a positive univariate trend with dense tissue (p < 0.10). Height and the age at menarche also showed a positive univariate trend in the group means, but with p > 0.10 (Table 2). The best multivariate fit for the absolute amount of dense tissue (R^2 = 13%, Table 3) was obtained with a model that contained the ages at time of examination (β = -0.131), occurrence of menarche (β = 0.144), occurrence of the menopause (β = 0.078), and the number of children (β = -0.284). The other factors with a univariate relationship were not found to contribute to the multivariate model.

For non-dense tissue, a positive univariate trend with p < 0.10 was found with the age at examination, weight, BMI, the time since the menopause, and the number of children (Table 2). A negative univariate trend (p < 0.10) in group means was found in having ever smoked and the age at occurrence of the menarche, although in the latter group means did not decrease in a linear way. The age at occurrence of the menopause and height were also included in the multivariate analyses based on the involvement with the other two measures. The final multivariate model (R^2 = 43%, Table 3) was composed only of BMI (β = 0.451), the age at the time of examination (β = 0.039), and the number of children (β = 0.085).

The relative amount of density showed a negative univariate relation (p < 0.10) with the age at examination, weight, BMI, the time since menopause, and the number of children (Table 2). There was a positive univariate relation (p < 0.10) with the ages at which menarche and menopause occurred and having ever smoked. There was also an apparent negative trend in grouped mean relative density with height, but with p > 0.10. The age at examination (β = -0.022), BMI (β = -0.012)), the age at occurrence of the menarche (β = 0.010), the age at which the menopause occurred (β = 0.008), and the number of children (β =-0.026) together provided the best multivariate fit (R^2 = 29%, Table 3).

Whereas 43% of the variance in non-dense tissue is explained by BMI, age at examination, and the number of children together, a model with BMI alone already has an explained variance of 40%. Addition of the other factors therefore only raises the explained variance by 3%. In contrast, absolute breast density was not related to BMI. Yet, for the relative amount of density, a univariate regression model of BMI has an explained variance of 17%, which is only raised to 29% after addition of the four other determinants.

Table 4 lists a cross-tabulation of quartiles of the absolute amount of dense tissue by the relative density to compare these two measures. There is a one-class shift in 144 of 418

Table 2 Geometric means of absolute density, absolute non-density and relative density and trend tests over groups of characteristics*.

		n	Absolute Dense (cm^2)	Absolute Non-Dense (cm^2)	Percent Dense (%)
Age at Examination (yr)	<54	139	31.3	72.0	26.1
	54-58	140	22.2	80.1	18.9
	>58	139	18.2	92.3	14.3
	p for trend	418	<0.01	<0.01	<0.01
Weight (kg)	<63	140	26.2	53.0	29.0
	63-70	143	22.6	91.3	17.6
	>70	135	21.2	114.6	13.8
	p for trend	418	0.06	<0.01	<0.01
Height (m)	<1.60	133	21.3	84.7	17.5
	1.60-1.65	151	23.6	81.0	19.4
	>1.65	134	25.2	80.2	20.8
	p for trend	418	0.15	0.51	0.18
BMI (kg/m^2)	<20	12	28.3	21.7	52.0
	20-<25	187	26.6	62.0	26.4
	25-<30	171	21.6	101.1	15.8
	≥30	41	17.3	156.1	8.9
	p for trend	418	0.07	<0.01	<0.01
Age at Menarche (yr)	<13	122	21.4	96.6	16.1
	13-14	178	23.4	72.3	21.1
	>14	116	25.5	82.8	20.5
	p for trend	416	0.16	0.07	0.08
Age at Menopause (yr)	<49	107	19.5	83.1	16.8
	49-51	139	19.9	90.9	15.8
	>51	115	27.0	77.7	22.5
	p for trend	361	0.01	0.42	0.04
Time since Menopause (yr)	<4	120	28.3	74.4	23.7
	4-9	121	21.1	81.5	18.1
	>9	120	17.4	98.5	13.5
	p for trend	361	<0.01	<0.01	<0.01
Parous	Yes	332	22.2	87.8	17.7
	No	86	28.1	62.5	26.3
	p	418	0.04	<0.01	<0.01
Age at First Childbirth (yr)	<25	110	20.8	92.6	16.2
	25-28	111	22.6	81.7	19.1
	>28	111	23.1	89.6	17.9
	p for trend	332	0.42	0.70	0.50
Parity (nr. of children)	0	86	28.1	62.5	26.3
	1	53	26.9	82.3	21.3
	2	93	28.5	77.3	23.6
	3	89	20.2	92.4	15.8
	>3	97	17.2	98.2	13.5
	p for trend	418	<0.01	<0.01	<0.01

Continued on next page

Table 2 Geometric means of absolute density, absolute non-density and relative density and trend tests over groups of characteristics*-continued

		n	Absolute Dense (cm^2)	Absolute Non-Dense (cm^2)	Percent Dense (%)
Current OC use	Yes	8	31.3	66.2	28.8
	No	410	23.2	82.2	19.1
	p	418	0.87	0.36	0.29
Ever used OC	Yes	43	28.0	78.3	23.5
	No	375	22.8	82.3	18.8
	p	418	0.18	0.64	0.19
Used HT in last 12 months	Yes	47	27.8	75.9	23.9
	No	371	22.8	82.7	18.7
	p	418	0.18	0.41	0.14
Ever smoked	Yes	116	27.4	67.5	25.1
	No	302	21.9	88.3	17.3
	p	418	0.03	<0.01	<0.01
Family History of Breast Cancer[#]	Yes	26	23.6	77.4	20.7
	No	382	23.2	82.7	19.0
	p	408	0.92	0.62	0.71

* Geometric means were calculated in tertiles of continuous variables or binary groups of dichotomous variables. For BMI, standard classes were used and for parity nominal classes, with a truncation at more than 3. Tests for linear trends were performed through ANOVA with an F-test on the geometric data. Results were transformed back to original units.
[#] Defined as one or more affected mother and/or sister.

Table 3 Multivariate determinants of absolute density, absolute non-density and relative density.

	Absolute Dense		Absolute Non-Dense		Relative Dense	
	β[*]	95% CI	β[*]	95% CI	β[*]	95% CI
BMI (kg/m^2)	-	-	0.515	0.451; 0.578	-0.022	-0.027; -0.016
Age at Examination (yr)	-0.131	-0.182; -0.079	0.091	0.039; 0.142	-0.012	-0.016; -0.008
Age at Menarche (yr)	0.144	0.008; 0.279			0.010	-0.002; 0.022
Age at Menopause (yr)	0.078	0.022; 0.135			0.008	0.003; 0.013
Parity (nr. of children)	-0.284	-0.395; -0.172	0.198	0.085; 0.312	-0.026	-0.035; -0.016
Adjusted R^2	0.13		0.43		0.29	

[*] β's from analyses based on square-root transformed measurements of density. Units of change are √cm^2 for absolute dense and non-dense and √% for relative dense.

women and a two-class shift in 12 of 418, resulting in a total level of agreement of κ = 0.50 (p < 0.01). Women in the lowest quartile of absolute amount of density had a relative density anywhere from near 0% to 33%. This was 6% to 60% for women in the second

quartile, 11% to 90% in the third quartile, and 20% to 88% in the upper quartile of absolute density.

Vice versa, the lowest quartile of relative density translated to an absolute amount of density ranging from 0.28 to 32.3 cm^2. The second and third quartiles of relative density both had a lower bound of 8.5 cm^2 with an upper bound of 65.2 cm^2 for quartile 2 and 91.5 cm^2 for quartile 3. The highest quartile of relative density had an upper bound as high as 138 cm^2 of absolute amount of dense tissue, but the lower bound of 17.9 cm^2 shows that the range in corresponding absolute amount of density of quartile 4 still overlaps the range of quartile 1.

Table 4 Crosstabulation of quartiles of absolute amount of density by quartiles of relative amount of density.

			Absolute amount of density				
		Range[#]	Q1	Q2	Q3	Q4	Total
Relative	Q1 (0-12%)	0.3-32.3	86	19	1	0	106
amount	Q2 (12-24%)	8.5-65.2	15	63	27	2	107
of	Q3 (25-42%)	8.5-91.5	3	20	43	31	97
density	Q4 (43-90%)	17.9-138	0	6	32	70	108
	Total		104	108	103	103	418
Range within quartile[*]			0.3-15.2	15.2-26.5	26.7-46.4	46.5-138	cm^2
Corresponding range[§]			0-33	6-60	11-90	20-88	%

[#] Range of the absolute amount of dense tissue (in cm^2) corresponding to the quartile of relative density.

[*] Range of the absolute amount of density (in cm^2) within each quartile.

[§] Range of the relative amount of density (in %) corresponding to the quartile of absolute amount of density.

Discussion

This study shows that the effects of determinants on the relative density are not an accurate representation of the actual relation that exists between these determinants and the dense tissue, which is considered to represent the target tissue for breast cancer. This disturbance is due to the relation that a determinant may also have with the amount of non-dense tissue. As the effect on the amounts of dense and non-dense tissues is measured in square centimeters and that on the relative density in percent, the magnitudes cannot be compared directly. However, in the case of age at examination and parity, the inverse relation of dense and non-dense tissues intuitively leads to an effect on the relative density which is an overestimation of the direct effect on dense tissue.

We can best support this by an example. If we take the ''median woman'' with median amounts of dense (26.4 cm^2) and non-dense tissues (87.8 cm^2), the relative density will be 23.1%. If we then consider a second woman who is exactly identical but 1 year older, the subsequent decrease in dense tissue of 1.38 cm^2 (calculated from β) and increase in non-dense tissue of 1.74 cm^2 (calculated from β) will indeed decrease the relative density by ~1.16%. If we now assume that there was no effect of age on the non-dense tissue, the change in relative density due to the decrease in dense tissue of 1.38 cm^2 would have only been 0.92%. This effect is most clearly seen in BMI, which is found to influence the relative

density but has no actual effect on the amount of dense tissue. Aside from the additional effect of BMI itself, the extra variable may also affect the estimates of the other variables. The relative density thus also comprises the effects on the non-target tissue, which prevents a direct, and therefore correct, etiologic interpretation of the influences.

The determinants of the relative amount of density found in our study and the explained variance (29%) are quite similar to results previously reported[14, 16, 21]. Our reliability results and the average tissue amounts are also similar to those reported in other studies. Thus, although the Xero-mammograms in this study prevented us from using the widely used, computer-assisted method for evaluating mammograms, we believe that our method gives equally reliable and comparable results.

The use of oral contraceptives (current and ever), use of hormone therapy in the last 12 months, and smoking were not found to significantly affect either the absolute or relative amount of density in the multivariate models. This may in part be due to the low prevalence of these factors in this population, which can be explained by the time period in which the participants were accrued (1974-1977) and their age at that time (>50 years). The use of oral contraceptives by postmenopausal women at the time of examination can also be ascribed to the accrual period as oral contraceptives were applied as a form of hormone therapy at the time. Despite the fact that these factors were not found to significantly contribute to the multivariate model (data not shown), the univariate influences showed an increase in the absolute amount of dense tissue and a decrease in non-dense tissue if oral contraception or hormone therapy was used (Table 2). The combined effect on the relative density is thereby amplified as the amounts of dense and non-dense tissues are negatively correlated. This makes more likely the finding of a significant effect of these factors on relative density as was done in some previous reports[26-28].

No data were available at all in our cohort on the ethnicity of the participants whereas ethnicity has long been an established determinant of breast patterns. However, the general population at the time of recruitment into the original DOM cohort was predominantly Caucasian. The results presented here should therefore be seen to reflect the situation in Caucasians and may differ to some extent in other ethnic groups.

Previously, Boyd et al.[21] and Mascarinec et al.[16, 17] reported the effects of similar determinants on the absolute amount of density in pre-menopausal women and a combined cohort of pre- and postmenopausal women, respectively. In contrast to our results, they reported that the absolute amount of dense tissue was also inversely influenced by BMI in a multivariate model. In this study, the effect of BMI on the relative amount of density is entirely due to the relation between BMI and non-dense tissue. When BMI was added to the final multivariate model for the absolute amount of density described in Table 3, the R^2 remained the same and the β for BMI was -0.044 (95% confidence interval, 0.019 to -0.108). The absence of an effect of BMI in our study in postmenopausal women may reflect the findings of a study on the role of diet on breast tissue[29]. Adherence to a 2-year low-fat, high-carbohydrate diet was shown to lower both the absolute and relative densities in pre-menopausal women, but in postmenopausal women the effect was negligible. Those results

may indicate that the menopause causes changes in the dense tissue, which decrease the susceptibility to dietary and BMI-related influences.

BMI itself has been shown to be a risk factor for breast cancer and independent roles for breast density and BMI as risk factors for breast cancer have been shown[30-32]. Den Tonkelaar et al.[33] found that one in five women experiences an increase in breast size after menopause, which is mainly due to weight gain and a subsequent increase in BMI. In view of the present study, one may argue that changes in risk that are ascribed to a difference in the relative amount of density are as likely to be attributable to differences in BMI as they are to actual changes in the dense tissue. In the mathematical structure of a ratio, BMI would deflate the influence of the absolute density on breast cancer risk in the estimated influence on breast cancer risk by the relative density. However, because BMI itself influences breast cancer risk, the elevated risk attributed to the relative density will be a resultant of the two influences and may therefore be higher than, lower than, or equal to the influence of the absolute density. Each of these options has indeed been shown in one or more studies that did calculate the breast cancer risk attributable to both the absolute and relative densities[3, 34-36]. This implies that the relative density may be a better prognostic value as the indicator of overall breast cancer risk associated to a certain parenchymal pattern but it is not necessarily the best etiologic value as the indicator of the risk that can be ascribed to the dense tissue.

Apart from the mathematical conflict between density and BMI, other factors may make the relations between BMI, absolute dense tissue, and breast cancer risk quite complex. One could argue that women with the same amount of dense tissue surrounded by more fatty tissue are at a higher risk due to increased aromatization of androgens into estrogens in the surrounding fatty tissue. Higher circulating levels of endogenous estrogens translate into higher risk and these levels may be elevated even more locally in the breast in relation to the amount of fat[37]. Similarly, higher local lipid peroxidation has been suggested to influence the dense tissue and thereby cancer risk[38]. Consequently, there may be nonlinear relations between the absolute amount of dense tissue and breast cancer risk as it is modified by the presence of fat tissue.

The comparison of quartiles of absolute density with quartiles of relative density showed that 37% (156 of 418) of the women in our study are classified differently for these measures. More importantly, however, the ranges of the relative amounts of density that correspond with each quartile of absolute amount of density show that quite similar absolute amounts merit a substantial range in relative amount and vice versa, depending on the absence or presence of fat. This shows that the relative density is not a good indicator of the absolute amount of density and that a low amount of relative density may still be associated with a large amount of high-risk, dense target tissue, simply surrounded by an even larger amount of non-dense tissue. This implies that within a category of relative density, there may still be a substantial range in risk, depending on the absolute amounts of dense and non-dense tissues constituting the percentage.

The age at menarche and the age at menopause only affect the absolute amount of density and do not influence the absolute amount of non-dense tissue. Consequently, the relative amount of density is also affected by these characteristics but the effect on relative density is diluted in comparison with the direct effect on the amount of dense tissue. Previous studies that included the age at menarche in the analyses show a large diversity in the effects that were found. Whereas our findings are quite similar to those reported by McCormack et al.[18], Maskarinec et al.[16] found an inverse relation. Other studies did not find any relation[14, 15, 19]. This high variability in findings on the role of the age at menarche indicates that this early effect on breast density may be obscured or negated by other influences later in life and needs further investigation.

The age at examination and parity are inversely related to the absolute amount of density and the absolute amount of non-dense tissue. The influence of both determinants on the relative amount of density is therefore a mix of the effects on the two absolute amounts. Although the direction of the relations between the absolute and the relative amount of density with both age and parity is similar, the effect of the relative density gives an overestimation of the direct effect on dense tissue.

The results presented here underline the notion that the relative density score is reflected by the amount of dense tissue as much as by the amount of surrounding tissue. The surrounding tissue largely constitutes of fat and BMI is its main determinant[39]. The implication is that when relative measurements of dense tissues are the (intermediate) endpoints in studies, effects on and of BMI are also being studied and the effect of a determinant on the relative density therefore does not represent the actual effects on dense tissue.

Consequently, the use of the relative amount of breast density is warranted when assessing risk in prognostic research because it is a measure that combines several pathways and is likely to yield the best, total estimate of risk associated with a certain breast. If, however, etiologic inferences are to be made in a study, absolute measures should be the measure of choice and should therefore be at least included.

Reference List

1. Boyd NF, Jensen HM, Cooke G, Han HL. Relationship between mammographic and histological risk factors for breast cancer. J Natl Cancer Inst 1992; 84(15):1170-1179.
2. Bright RA, Morrison AS, Brisson J et al. Relationship between mammographic and histologic features of breast tissue in women with benign biopsies. Cancer 1988; 61(2):266-271.
3. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. Am J Roentgenol 1976; 126(6):1130-1137.
4. Geller BM, Barlow WE, Ballard-Barbash R et al. Use of the American College of Radiology BI-RADS to report on the mammographic evaluation of women with signs and symptoms of breast disease. Radiology 2002; 222(2):536-542.
5. Gram IT, Funkhouser E, Nordgard L, Tabar L, Ursin G. Oral contraceptive use and mammographic patterns. Eur J Cancer Prev 2002; 11(3):265-270.
6. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. Automated analysis of mammographic densities. Phys Med Biol 1996; 41(5):909-923.
7. Yaffe MJ, Boyd NF, Byng JW et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998; 7 Suppl 1:S47-S55.
8. Boyd NF, Byng JW, Jong RA et al. Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study. J Natl Cancer Inst 1995; 87(9):670-675.
9. Byrne C, Schairer C, Wolfe J et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995; 87(21):1622-1629.
10. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976; 37(5):2486-2492.
11. Brisson J, Diorio C, Masse B. Wolfe's parenchymal pattern and percentage of the breast with mammographic densities: redundant or complementary classifications? Cancer Epidemiol Biomarkers Prev 2003; 12(8):728-732.
12. Warner E, Lockwood G, Tritchler D, Boyd NF. The risk of breast cancer associated with mammographic parenchymal patterns: a meta-analysis of the published literature to examine the effect of method of classification. Cancer Detect Prev 1992; 16(1):67-72.
13. Gapstur SM, Lopez P, Colangelo LA, Wolfman J, Van Horn L, Hendrick RE. Associations of breast cancer risk factors with breast density in Hispanic women. Cancer Epidemiol Biomarkers Prev 2003; 12(10):1074-1080.
14. Heng D, Gao F, Jong R et al. Risk factors for breast cancer associated with mammographic features in Singaporean chinese women. Cancer Epidemiol Biomarkers Prev 2004; 13(11 Pt 1):1751-1758.
15. Jeffreys M, Warren R, Gunnell D, McCarron P, Smith GD. Life course breast cancer risk factors and adult breast density (United Kingdom). Cancer Causes Control 2004; 15(9):947-955.
16. Maskarinec G, Lyu LC, Meng L, Theriault A, Ursin G. Determinants of mammographic densities among women of Asian, Native Hawaiian, and Caucasian ancestry. Ethn Dis 2001; 11(1):44-50.
17. Maskarinec G, Nagata C, Shimizu H, Kashiki Y. Comparison of mammographic densities and their determinants in women from Japan and Hawaii. Int J Cancer 2002; 102(1):29-33.
18. McCormack VA, dos SS, I, de Stavola BL et al. Life-course body size and perimenopausal mammographic parenchymal patterns in the MRC 1946 British birth cohort. Br J Cancer 2003; 89(5):852-859.
19. Riza E, dos SS, I, De SB et al. Correlates of high-density mammographic parenchymal patterns by menopausal status in a rural population in Northern Greece. Eur J Cancer 2005; 41(4):590-600.
20. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control 2000; 11(7):653-662.
21. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer 1998; 78(9):1233-1238.
22. Jasienski M, Bazzaz FA. The fallacy of ratios and the testability of models in biology. Oikos 1999; 84(2):321-326.
23. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst 1988; 80(10):772-774.

24. Trichopoulos D, Lipman RD. Mammary gland mass and breast cancer risk. Epidemiology 1992; 3(6):523-526.
25. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984; 1(8388):1224-1226.
26. Greendale GA, Reboussin BA, Slone S, Wasilauskas C, Pike MC, Ursin G. Postmenopausal hormone therapy and change in mammographic density. J Natl Cancer Inst 2003; 95(1):30-37.
27. Spicer DV, Ursin G, Parisky YR et al. Changes in mammographic densities induced by a hormonal contraceptive designed to reduce breast cancer risk. J Natl Cancer Inst 1994; 86(6):431-436.
28. Vachon CM, Sellers TA, Vierkant RA, Wu FF, Brandt KR. Case-control study of increased mammographic breast density response to hormone replacement therapy. Cancer Epidemiol Biomarkers Prev 2002; 11(11):1382-1388.
29. Boyd NF, Greenberg C, Lockwood G et al. Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: results from a randomized trial. Canadian Diet and Breast Cancer Prevention Study Group. J Natl Cancer Inst 1997; 89(7):488-496.
30. Adami HO, Trichopoulos D. Obesity and mortality from cancer. N Engl J Med 2003; 348(17):1623-1624.
31. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003; 348(17):1625-1638.
32. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. Cancer 2000; 89(2):369-375.
33. den Tonkelaar I, Peeters PH, van Noord PA. Increase in breast size after menopause: prevalence and determinants. Maturitas 2004; 48(1):51-57.
34. Chen Z, Wu AH, Gauderman WJ et al. Does mammographic density reflect ethnic differences in breast cancer incidence rates? Am J Epidemiol 2004; 159(2):140-147.
35. Maskarinec G, Meng L. A case-control study of mammographic densities in Hawaii. Breast Cancer Res Treat 2000; 63(2):153-161.
36. Ursin G, Ma H, Wu AH et al. Mammographic density and breast cancer in three ethnic groups. Cancer Epidemiol Biomarkers Prev 2003; 12(4):332-338.
37. Boyd NF, Stone J, Martin LJ et al. The association of breast mitogens with mammographic densities. Br J Cancer 2002; 87(8):876-882.
38. Boyd NF, McGuire V. Evidence of lipid peroxidation in premenopausal women with mammographic dysplasia. Cancer Lett 1990; 50(1):31-37.
39. Beijerinck D, van Noord PA, Kemmeren JM, Seidell JC. Breast size as a determinant of breast cancer. Int J Obes Relat Metab Disord 1995; 19(3):202-205.

The influence of a period of caloric restriction due to the Dutch Famine on breast density

Gerco Haars

Carla H van Gils

Sjoerd G Elias

Petra HM Peeters

Diederick E Grobbee

Paul AH van Noord

4

Abstract

Longer term caloric restriction often reduces cancer risk in animal studies. Recently, we showed that short, intense caloric restriction due to the Dutch famine increased breast cancer risk in women and influenced the hormonal- and IGF-system. These systems may also affect breast density, which is one of the strongest risk factors for breast cancer in humans. We set out to determine the influence of the Dutch famine on breast density, using mammograms of 1035 women. Breast size, dense and non-dense tissue and the relative density were measured on a continuous scale. Mean densities were compared between three groups of ascending levels of famine-exposure. Results were adjusted for known determinants of breast density and stratified by age at exposure.

There were no overall, significant differences by exposure. Means varied from 124 cm^2 in unexposed to 121 cm^2 (p_{trend} =0.50) in severely exposed women for breast size, from 23.4 to 21.8 cm^2 (p_{trend} =0.48) for amount of dense tissue, from 87.7 to 85.4 cm^2 (p_{trend} =0.55) for non-dense tissue and from 22.8 to 22.3% (p_{trend} =0.78) for relative density. Analysed according to age at exposure, non-dense tissue was significantly lower with higher exposure before age 10, with 53.1cm^2 for severely exposed as to 77.8 cm^2 (p_{trend}=0.03) for unexposed. This group also appeared to have smaller breast with more absolute and relative density, but not statistically significant. We observed no overall effect on breast density on women exposed to a short, intense caloric restriction. However, in women exposed before puberty, non-dense area was decreased and density tended to be higher.

Introduction

Animal studies have consistently shown that long-term caloric restriction may protect against various types of cancer, including cancer of the breast[1-4]. However, in a recent study in a cohort of women who were exposed to the Dutch famine in the winter of 1944-1945, we showed that a short period of intense caloric restriction at a young age lead to an elevated risk of breast cancer later in life[5]. This effect was most prominent in women who had been exposed before breast development, at an age of 2 -9 years old.

We also found indications of an effect by the famine on endogenous sex-hormones and the IGF system compatible with an elevated breast cancer risk[6, 7]. In a small sample of nulliparous women, estrone and estradiol levels were higher with the level of exposure to the famine and in a sample of postmenopausal women, the levels of IGF-I and IGFBP-3 were also higher with ascending exposure. Both findings indicate a permanent effect of famine on a system that has been related to breast cancer risk and it is therefore possible that the period of caloric restriction has asserted its influence on breast cancer risk through these pathways[8-13].

In humans, the breast density, as visible on mammograms, has been established as one of the strongest and most consistent, known risk factors for breast cancer[14, 15]. A breast parenchymal pattern is determined by the relative amount of radio dense target tissue in contrast to surrounding non-dense tissue. Computerised techniques that calculate a continuous measure of the amount of dense tissue in the breast are available[16, 17]. Much remains unknown about systems regulating density and how density asserts its influence on breast cancer risk, but the IGF-system that was found to be influenced by the caloric restriction has also been associated with breast cancer risk and breast density[18-20]. Contrary to this, the endogenous sex-hormone system has been suggested to be an independent pathway to breast cancer[12, 20].

In view of our previous finding that caloric restriction increases breast cancer risk, the present study examines how breast density, assessed with a continuous measurement method, is affected by the short caloric restriction caused by the Dutch famine and whether this effect is dependent on age at exposure. This would support a possible mediating role of breast density in the relation between the relatively short but intense period of caloric restriction and breast cancer risk, which could be due to a common underlying pathway such as IGF.

Subjects and Methods

The DOM-project (Diagnostisch Onderzoek Mammacarcinoom) is a prospective breast cancer-screening program in Utrecht, The Netherlands[21, 22]. Between 1974 and 1986 a total of 55,519 women, born between 1911 and 1945, were recruited. Participants had their mammograms taken on one or more occasion, provided data on medical, reproductive and lifestyle status and history through questionnaires and underwent physical examination from which anthropometrical data was collected.

Between 1983 and 1986, 19,732 women received questionnaires that included questions regarding the 1944-1945 Dutch famine. The answers to three questions on exposure to "hunger", "cold" and "weight loss" were combined into one famine exposure score. Women who answered "severely" on two or more of these questions were classified as "severely exposed". Women who answered "absent" on two or more questions were classified as "unexposed" and all others were classified as "moderately exposed". Such a combined famine exposure score was available for 15,396 women (89%)[5].

In a previous study we selected a random sample of 2,352 women (15%) as a representation of the total cohort for whom follow-up on survival and development of breast cancer was obtained until January 2000[5]. From this initial selection, we sampled 40% of the "moderately exposed" and "unexposed" and all of the "severely exposed" participants for whom we tried to retrieve the first mammogram taken of the right breast in the cranio-caudal view. This resulted in a final population of 452 women "unexposed" to the famine, 358 women who were "moderately exposed" and 225 who were "severely exposed".

Baseline characteristics were obtained for a number of factors that are known to influence either breast density or breast cancer risk and could thus act as confounders. These include the age at mammography, height, weight and BMI information on menopausal status and information on parity. The age of each woman at the time of the famine was calculated and the participants were classified according to a classification adapted from Bogin; 2-9 years old (early and middle childhood), 10-18 years old (later childhood and adolescence) and older than 18 (adulthood)[23].

The 1035 mammograms were Xero-print-mammograms instead of the current film-screen mammograms. They were digitised on a flatbed scanner (Hewlett-Packard, HP5200) at 100 dpi and stored without any reference to the exposure status. A trained observer (GH) evaluated the digital images with the program Image Explorer (Image Sciences Institute, UMC Utrecht) by manually tracing the contours of the total breast and the dense tissue. The amount of non-dense tissue was calculated by subtracting the dense tissue from the breast size and the percent density was calculated by dividing the amount of density by the total breast size. A retest sample of 15% (155 mammograms) was evaluated in duplicate and showed an intraclass correlation for the intra-reader reliability of >0.99 for breast size and 0.91 for both absolute and relative amount of density.

The data were transformed to achieve a normal distribution. The relative density could be normalised by taking the square root. For breast size and the absolute amounts of dense and non-dense tissue, normality was achieved by taking the double square root. The outcomes of the analyses were back-transformed to the original units.

Breast size, absolute amount of dense and non-dense tissue and relative density were each studied in relation to the level of exposure. Analyses were also done in the three strata for age during exposure to the famine. Analyses were performed using both unadjusted and adjusted ANCOVA (SPSS 11.5, SPSS Inc.). The adjusted models included age, parity, menopausal status and BMI, as these are the most influential, known determinants of beast cancer and/or breast density.

Results

The characteristics of the population for various known risk factors of breast density and breast cancer are shown in table 1 for all women and in strata of famine exposure. The age at examination increased slightly with the level of exposure, from 52 for unexposed women to 54 for severely exposed. In the severely exposed women, 81% were postmenopausal at the time of examination compared to 60% of the unexposed, but the median age at which the menopause occurred did not differ. The unexposed women have had a natural menopause more often than those who were moderately or severely exposed (75% vs. 64%). Adjusted for age at examination, parity, body mass index and menopausal status, the geometric mean breast size in all women was $123cm^2$ (95%CI 120-126cm^2). The amount of dense tissue, non-dense tissue and relative density were $23.4cm^2$ (95%CI 22.2-24.7cm^2), $86.5cm^2$ (95%CI 83.2-88.8cm^2) and 22.9% (95%CI 21.7-24.2%) respectively. Neither breast size, amount of dense tissue or amount of non-dense tissue nor the relative density showed a crude or adjusted significant trend over severity of exposure (p-values between 0.18 and 0.89, table 2). Additional ANOVA also did not show any significant non-linear difference between levels of exposure (data not shown).

Table 1; Baseline characteristics in all women, also as per Famine Exposure.

	All	By severity of famine exposure		
		Unexposed	Moderately	Severely
n	1035	452	358	225
	Median (range) or Number (percentage)			
Age (yr)	53 (41-66)	52 (41-66)	53 (41-66)	54 (43-66)
Weight (kg)	68 (44-133)	67 (44-124)	68 (44-108)	68 (47-132)
Height (m)	1.64 (1.46-1.87)	1.64 (1.50-1.78)	1.64 (1.46-1.87)	1.63 (1.46-1.82)
BMI (kg/m^2)	25 (16-46)	25 (16-46)	25 (18-40)	26 (17-46)
Postmenopausal (yes)	717 (69%)	272 (60%)	263 (74%)	182 (81%)
Natural menopause[1] yes)	512 (71%)	205 (75%)	181 (69%)	126 (69%)
Age at Menopause[1] (yr)	49 (28-60)	49 (32-57)	50 (33-60)	49 (28-58)
Nulliparous (yes)	153 (15%)	56 (12%)	67 (19%)	30 (13%)
Parity[2] (nr. of children)	3 (1-9)	3 (1-9)	3 (1-9)	3 (1-9)
Age first child[2] (yr)	26 (18-42)	26 (18-42)	27 (18-41)	26 (18-38)
Age during the Famine				
2-9 years	144 (14%)	91 (20%)	38 (11%)	15 (7%)
10-18 years	356 (34%)	179 (40%)	111 (31%)	66 (29%)
> 18 years	535 (52%)	182 (40%)	209 (58%)	144 (64%)

[1] Only postmenopausal women included (n=717, 272, 263 and 182 respectively).

[2] Only parous women included (n=882, 396, 291 and 195 respectively).

[3] Defined as at least one affected direct relative (mother and/or sister).

In analyses stratified for the age at exposure to the famine, there was only a significant trend in the non-dense tissue for women aged 2 to 9 years at exposure to the famine. In this group the amount of non-dense tissue was lower with 77.8 cm^2 in the unexposed compared to 53.1 cm^2 in the severely exposed women (p$_{trend}$=0.03). Level of exposure did not show any other significant relation with breast size, dense or non-dense tissue or relative density within any of the age-strata in either the unadjusted or the adjusted models with p values for trend in the adjusted models between 0.10 and 0.97 (table 3, crude analyses not shown). Although there is no statistically significant trend with exposure, severely exposed women do seem to have higher amounts of absolute dense tissue compared to unexposed women in this age group (23.0 vs. 19.8 cm^2; p=0.63). They also tend to have smaller breasts (92 vs. 116 cm^2, p=0.10), with a higher relative density (30.3 vs. 22.3%; p=0.19) which is largely due to the difference in non-dense tissue.

Table 2; Adjusted means of measured breast size, dense tissue size, non-dense tissue size and relative density, also according to famine exposure.

	All	Unexposed	Moderately	Severely	P$_{trend}$
	Adjusted mean (95%CI)				
Breast Size (cm^2)	123 (120-126)	124 (120-129)	124 (118-131)	121 (116-127)	0.50
Dense Tissue (cm^2)	23.4 (22.2-24.7)	23.4 (21.4-25.6)	25.2 (22.6-28.0)	21.8 (19.1-24.7)	0.48
Non Dense Tissue (cm^2)	86.5 (83.2-88.8)	87.7 (83.2-91.2)	86.5 (82.1-91.2)	85.4 (78.9-91.2)	0.55
Relative Density (%)	22.9 (21.7-24.2)	22.8 (21.0-24.6)	23.8 (21.8-25.9)	22.3 (19.9-24.8)	0.78

[1] Adjusted for age at examination (yr), parity (yes or no), menopausal status (post or pre/peri) and BMI (kg/m^2).

Discussion

The present study shows no overall effect of famine exposure on breast patterns determined by radiographic density. Our population is a representative sample of the cohort in which we previously found an effect of a period of caloric restriction on breast cancer risk that was most prominent in women exposed before 10 years old[5]. In the same cohort concomitant influences on IGF-I, IGFBP3, estrone and estradiol were found[6, 7].

For women severely exposed to the famine before the age of 10 there is a tendency for a lower amount of non-dense tissue and a higher amount of dense tissue, resulting in smaller breast with higher relative density as opposed to unexposed women. There is however, no trend with the level of exposure and the results are not statistically significant for all measures as there are only 15 observations. These results should therefore be regarded with caution.

It is, nonetheless, striking that, as in our previous studies, the most likely effect is present in women who are exposed at a time at which the breasts still have to develop. This supports a theory of long-lasting, systemic effects in the hypothalamo-pituitary-gonadal-axis. These systemic effects would lead to an altered level of production of gonadothrophines. Later in life this would lead to altered levels of hormones and growth hormones as found in previous studies and thereby to an affected development of the breast.

In view of the possible relation between IGF and breast density, the early effects of the caloric restriction on the IGF system may have lead to subsequent effects on breast density.

Table 3; Adjusted means of breast size, dense tissue size, non-dense tissue size and relative density per age during the famine and famine exposure.

	Famine Exposure			
	Low	Moderate	Severe	P_{trend}
Breast size (cm^2)	Mean (95%CI)	Mean (95%CI)	Mean (95%CI)	
2-9 Years	116 (105-127)	119 (102-137)	92 (71-119)	0.10
10-18 Years	123 (116-131)	124 (116-134)	126 (114-138)	0.76
Over 18 Years	127 (121-134)	127 (121-134)	123 (116-131)	0.29
Dense Tissue (cm^2)				
2-9 Years	19.8 (15.7-24.7)	19.1 (13.0-27.0)	23.0 (12.5-39.1)	0.63
10-18 Years	25.2 (22.2-28.5)	28.0 (23.9-32.6)	23.4 (18.7-29.0)	0.59
Over 18 Years	22.6 (19.8-25.6)	25.2 (22.2-28.5)	22.6 (19.4-26.1)	0.97
Non-Dense Tissue (cm^2)				
2-9 Years	77.8 (68.8-87.7)	87.7 (72.7-104.9)	53.1 (37.8-72.7)	0.03
10-18 Years	85.4 (78.9-92.4)	81.0 (72.7-90.0)	86.5 (75.7-98.5)	0.92
Over 18 Years	92.4 (86.5-98.5)	91.2 (85.4-97.2)	87.7 (81.0-94.8)	0.36
Relative Density (%)				
2-9 Years	22.3 (18.3-26.6)	18.9 (13.5-25.3)	30.3 (19.4-43.4)	0.19
10-18 Years	24.1 (21.4-26.9)	26.8 (23.3-30.6)	23.3 (19.2-27.9)	0.75
Over 18 Years	21.4 (19.1-23.9)	22.8 (20.5-25.3)	21.9 (19.2-24.8)	0.81

[1] Adjusted for age at examination (yr), parity (yes or no), menopausal status (post or pr-/peri) and BMI (kg/m^2).

Our results, however, do not support an intermediate role of breast density in the relation between IGF and breast cancer.

A limitation of the present study is the limited number of women who were severely exposed, especially in the stratified analyses. Yet, the Dutch famine provides a unique natural experiment to study an extraordinary influence on the human body. Moreover, the women participating in this study were followed prospectively irrespective of their exposure to the famine, adding to the strength of the data.

The classification of the exposure is based on the subjective recollection of the participants. However, the distribution of individual exposure scores was in line with the reported area of residence during the famine, for which the overall severity of the famine was well known. Furthermore, the information on the exposure was obtained independently from the breast measurement. Any misclassification would therefore have lead to an underestimation of the effects. Although this could explain the lack of significant effects in the present study, we believe that this was not a determining factor as the data was sufficient to find results in previous studies with respect to breast cancer risk[5].

While our findings on the influence of famine exposure on the amount of non-dense tissue in women exposed before age 10 should be considered with care, the possible higher amount of dense tissue and relative density are in agreement with our previous findings of a higher breast cancer risk in women severely exposed at a young age. The absence of overall significant effects of the famine exposure on breast density however indicates that the degree of dense tissue does not play a pivotal role in the elevation of breast cancer risk by caloric restriction. Effects of the caloric restriction on the IGF system may have sideways affected breast density, whereas the effects of caloric restriction on the endogenous sex-hormones are more likely to affect breast cancer risk through another pathway.

Reference List

1. Dirx MJ, Zeegers MP, Dagnelie PC, van den BT, van den Brandt PA. Energy restriction and the risk of spontaneous mammary tumors in mice: a meta-analysis. Int J Cancer 2003; 106(5):766-770.
2. Kritchevsky D. Caloric restriction and experimental mammary carcinogenesis. Breast Cancer Res Treat 1997; 46(2-3):161-167.
3. Moreschi C. Beziehungen zwischen ernährung und tumorwachsen. Zeitschrift fur Immunitätsforschung Originale 1909; Bd.II:651-675.
4. Weindruch R, Walford RL. The retardation of aging and disease by dietary restriction. Springfield, Illinois, United States of America: Charles C Thomas Publisher; 1988.
5. Elias SG, Peeters PH, Grobbee DE, van Noord PA. Breast cancer risk after caloric restriction during the 1944-1945 Dutch famine. J Natl Cancer Inst 2004; 96(7):539-546.
6. Elias SG, Keinan-Boker L, Peeters PH et al. Long term consequences of the 1944-1945 Dutch famine on the insulin-like growth factor axis. Int J Cancer 2004; 108(4):628-630.
7. Elias SG, Onland-Moret NC, Peeters PH et al. Urinary endogenous sex hormone levels in postmenopausal women after caloric restriction in young adulthood. Br J Cancer 2004; 90(1):115-117.
8. Kari FW, Dunn SE, French JE, Barrett JC. Roles for insulin-like growth factor-1 in mediating the anti-carcinogenic effects of caloric restriction. J Nutr Health Aging 1999; 3(2):92-101.
9. McCarty MF. Up-regulation of IGF binding protein-1 as an anticarcinogenic strategy: relevance to caloric restriction, exercise, and insulin sensitivity. Med Hypotheses 1997; 48(4):297-308.
10. Sell C. Caloric restriction and insulin-like growth factors in aging and cancer. Horm Metab Res 2003; 35(11-12):705-711.
11. Tamimi RM, Cox DG, Kraft P et al. Common genetic variation in IGF1, IGFBP-1, and IGFBP-3 in relation to mammographic density: a cross-sectional study. Breast Cancer Res 2007; 9(1):R18.
12. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 2007; 99(15):1178-1187.
13. van Noord PA. Breast cancer and the brain: a neurodevelopmental hypothesis to explain the opposing effects of caloric deprivation during the Dutch famine of 1944-1945 on breast cancer and its risk factors. J Nutr 2004; 134(12 Suppl):3399S-3406S.
14. Boyd NF, Guo H, Martin LJ et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3):227-236.
15. McCormack VA, dos SS, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006; 15(6):1159-1169.
16. Byng JW, Yaffe MJ, Lockwood GA, Little LE, Tritchler DL, Boyd NF. Automated analysis of mammographic densities and breast carcinoma risk. Cancer 1997; 80(1):66-74.
17. Yaffe MJ, Boyd NF, Byng JW et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998; 7 Suppl 1:S47-S55.
18. Hankinson SE, Willett WC, Colditz GA et al. Circulating concentrations of insulin-like growth factor-I and risk of breast cancer. Lancet 1998; 351(9113):1393-1396.
19. Hankinson SE, Schernhammer ES. Insulin-like growth factor and breast cancer risk: evidence from observational studies. Breast Dis 2003; 17:27-40.:27-40.
20. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. Breast Cancer Res 2008; 10(1):201.
21. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984; 1(8388):1224-1226.
22. de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 1984; 37(1):1-44.
23. Bogin B. Patterns of human growth. Cambridge, United Kingdom: Cambridge University Press; 1999.

The influence of dense and non-dense tissue on breast cancer risk

Gerco Haars

Carla H van Gils

Sjoerd G Elias

Petra HM Peeters

Diederick E Grobbee

Paul AH van Noord

5

Abstract

Relative breast density is an established risk factor for breast cancer. Yet relative density is a combined score of dense and non-dense tissue and obscures absolute amounts. We aimed to determine the effect of the component breast tissues on breast cancer risk. We used a case-cohort design within the DOM breast cancer screening program (Utrecht, The Netherlands). All newly diagnosed invasive breast cancers in 15.5 years of follow-up were determined (n=361) and follow up experience was obtained for a random cohort sample of 600 women. Mammograms were obtained, digitized and evaluated for 309 cases and the cohort sample. The risk of breast cancer was compared between tertiles of either dense tissue, non-dense tissue or relative density and in combined categories of dense tissue and non-dense tissue (low/low vs. low/high, high/low and high/high).

Breast cancer risk appeared to increase independently with increasing amounts of density (HR 1.68; 95%CI 1.22-2.46 for highest tertile vs. lowest) and non-density (HR 1.44; 95%CI 1.03-2.03 for highest tertile vs. lowest). The highest risk for breast cancer was found in the women with above median amounts of both tissues (HR 2.20; 95%CI 1.37-3.52) compared to women with below median amounts of both tissues. Non-dense tissue also plays a role in the development of breast cancer. The fatty tissue that constitutes non-dense tissue may elevate the risk of breast cancer by local production of estrogens or local lipid peroxidation. A combination of a high amount of dense tissue and a high amount of non-dense tissue may be the most hazardous breast pattern.

Introduction

One of the strongest known risk factors for breast cancer is the mammographic appearance of the breast as to the amount of glandular target tissue[1, 2]. Breast glandular tissue is visible on mammograms due to the relatively high radio density of stromal and epithelial tissue in comparison to the surrounding tissue, which predominantly consists of fat. Due to this appearance on a mammogram, these tissues are referred to as dense and non-dense tissue. Together, the dense and non-dense tissues form a mammographic pattern. Since the first classification of these patterns by Wolfe in 1976[3], breast density has predominantly been evaluated as the percentage of the breast occupied by dense tissue. This was originally done in categories as developed by Wolfe. Subsequently the Tabar and BI-RADS classification were introduced[4, 5]. All three are based on subjective scoring of the patterns on overall appearance and relative density. In recent years, more objective, continuous measurement of dense tissue has become ever more popular due to new computerized techniques[6-8]. A high proportion of mammographic density, either scored categorically or continuously, has consistently been related to an elevated risk of breast cancer with relative risks up to 4 or 5 for women with more than 75% density compared to women with less than 5% density[1, 2].

The most established influence on the relative density comes from the Body Mass Index (BMI), where an increase in BMI is associated with a decrease in relative density[9-11]. Based on the established relation between relative density and breast cancer, this would imply a decrease in breast cancer risk with an increase in BMI. Adjustment for BMI, however, tends to increase the risk estimate of relative density. This was explained in a study by Boyd et al., where BMI and mammographic density were found to be independent risk factors for postmenopausal breast cancer[12].

In a recent study we showed that the effect of BMI on relative density is mainly through an effect on the amount of non-dense tissue, which can be seen as a proxy for the amount of fat in the breast[13]. Combined with the opposing relations of BMI with relative density and breast cancer risk, this finding indicated that the relative density as a single measure can obscure the roles of its components.

It is our goal to study the independent relations of the amount of dense tissue, the amount of non-dense tissue and the relative density with breast cancer risk and to investigate the effects of combinations of amounts of dense and non-dense tissue.

Subjects and Methods

The DOM-project (Diagnostisch Onderzoek Mammacarcinoom) is a prospective breast cancer screening program in Utrecht, The Netherlands. Between 1974 and 1986, it recruited a total of 55,519 women who were born between 1911 and 1945[14, 15]. Participants had their mammograms taken on one or more occasion, provided data on lifestyle and medical, social and reproductive history through questionnaires and underwent physical examination by trained staff at which anthropometric data was collected.

For a sub cohort of 9,856 women, all new invasive breast cancer cases between start of follow-up and January 2000 were known through linkage with the DOM-registry (for those

diagnosed between 1974-1989) and the national cancer registry (for those diagnosed after 1989). This sub cohort served as the base-cohort in the present study. Additional follow-up on vital status was available for a sample of 1,556 women (15%) as a representation of this cohort (case-cohort design) and all 361 cases identified in this cohort. We retrieved the mammogram of the right breast in the cranio-caudal view for the cases and 600 of the 1,556 women (6% of the total cohort).

As a consequence of the time at recruitment of these participants, this involves Xero-print mammograms for the large majority of participants. Xero-print mammograms preceded the current Film-screen mammograms and differ mainly by being a positive image. Although the contrast of current Film-screen mammograms is higher, the original Xero-print mammograms also allow for a proper distinction between dense and non-dense tissue.

The print-mammograms were digitized on a flatbed scanner (Hewlett-Packard, HP5200) at 100 dpi and stored without any reference to outcome status. A trained observer (GH) evaluated the unclassified digital images with the program Image Xplorer (Imaging Sciences Institute, UMC Utrecht) by manually tracing the contours of the total breast and the dense tissue. The amount of non-dense tissue was calculated by subtracting the amount of dense tissue from the total breast size and the relative density was calculated as the amount of dense tissue divided by the total breast size times 100 for a percentage score. Intra-reader reliability for this method of measurement is >90%, which is similar to values in literature[16].

Characteristics at the time of examination were obtained for all participants as to a number of factors which are known to be related to breast density, breast cancer risk or both. These are age at examination, age at menarche, menopausal status, type of menopause and age at menopause, nulliparity, the number of full-term deliveries and the age at first childbirth, height, weight and Body Mass Index (BMI), current or past use of oral contraceptives and hormone replacement therapy (HRT), family history of breast cancer, smoking and socio-economic status based on the level of insurance.

The measurements of breast size, dense tissue, non-dense tissue and relative density were divided into tertiles for power considerations. The incidence rate was calculated for the total population and in the various strata. As some known cases could not be included due to unavailable mammograms, the incidence rates were adjusted for the fraction of missing cases (52/361).

The relation between the tertiles of these tissue measurements and breast cancer risk was analyzed with a weighted Cox regression analysis specifically designed for case-cohort analysis[17-19]. Follow-up time started at the examination where a questionnaire was completed and a mammogram taken (between 1983 and 1986) and ended at the date of diagnosis of primary invasive breast cancer (event) or censorship. Women that remained free of breast cancer during the period of observation were censored at date of movement, date of death or at 01-01-2000, whatever occurred first.

The weighted Cox regression was performed in SAS v8.2 (SAS Institute Inc., Cary, NC, USA) with the use of a macro (available at http://lib.stat.cmu.edu/general/robphreg) that

computes the weighted estimates together with a robust standard error, from which 95% confidence intervals (CI) were calculated.

In order to investigate the level of joint effects of the amounts of dense and non-dense tissue, these two variables were dichotomized as above or below median and subsequently cross tabulated. A weighted Cox analysis was performed as described above on the four strata with the stratum of below median amount of dense and non-dense tissue as the reference. The relative excess risk (RERI) of the stratum with above median amounts of dense and non-dense tissue was calculated to assess the level of additive interaction.

There has been debate on the possibility that dense tissue does not itself elicit an influence of breast cancer but that the dense tissue merely obscures existing tumours from detection. All analyses were therefore also performed on a study cohort that was restricted to include only cases that developed more than 2 years after start of follow-up, in order to eliminate any effects of this masking bias.

Results

Xero-print mammograms were retrieved for the cohort sample of 600 women. Of these women, one had to be excluded from the analysis because no anthropometric information was available, 484 (81%) were still alive on January 1st 2000, 41 (7%) were deceased, 48 (8%) had emigrated and 26 (4%) had been lost to follow-up. This cohort sample accrued a total of 8,168 person-years of follow-up, which relates to 134,395 person years for the total cohort when extrapolating on the basis of the sampling fraction.

Of the 361 women that were newly diagnosed with primary invasive breast cancer during the follow-up period, a Xero-print mammogram could be retrieved for 309 (86%). We excluded 48 women (13%) because only Film-screen films were available and 4 (1%) other women because no mammogram could be retrieved at all.

Baseline characteristics of the cases and the cohort sample-women are presented in table 1. On average, the cases were as old as the women in the cohort sample at the time of examination and had similar anthropometric features (height, weight and BMI). Whereas the women in both groups had started menstruating at about the same age, fewer cases had gone through menopause (73 vs. 84%). The mean age at menopause of the women that had passed through menopause was similar, as well as the percentage of women that had reported a natural menopause. Cases more often had used hormone replacement therapy or oral contraceptives in the year prior to examination, more often had a family history of breast cancer and more often smoked at the time of examination. Women in the cohort sample also appeared more likely to be parous and to have had more children on average, although the average age at first full-term childbirth was the same.

The measurements of the various breast tissues show that on average, cases have larger breasts (134 vs. 118 cm^2) that contain both more dense tissue (33.8 vs. 28.0 cm^2) as well as more non-dense tissue (91.6 vs. 81.8 cm^2) than the women in the cohort sample. Relative density is also higher in cases compared to controls (28.4% vs. 26.0%).

Table 1 Characteristics of cases and cohort sample for factors known to influence breast density or cancer risk.

	Cases (n=309)		Cohort Sample (n=599)	
	Mean (SD)	Range	Mean (SD)	Range
Age at mammogram(yr)	52.0 (5.7)	42.1-68.2	52.8 (5.8)	41.8-68.7
Weight (kg)	69.1 (10.7)	38.0-110	69.0 (11.2)	44-124
Height (m)	1.66 (0.58)	1.50-1.87	1.64 (0.55)	1.49-1.80
BMI (kg/m^2)	25.2 (3.7)	15.5-38.6	25.5 (4.0)	15.8-45.9
Age at Menarche (yr)	13.5 (1.6)	10.2-18.3	13.7 (1.6)	10.0-18.3
Age at Menopause[1] (yr)	48.9 (4.9)	26.0-57.0	48.7 (4.9)	28.0-59.0
Parity[2] (nr. of children)	2.8 (1.2)	1-9	3.2 (1.7)	1-9
Age at first childbirth[2] (yr)	26.7 (4.3)	15.3-40.3	26.8 (4.1)	18.1-41.7
	n (proportion)		n (proportion)	
Postmenopausal (yes)	225 (73%)		501 (84%)	
Natural menopause (yes)	162 (72%)		356 (71%)	
Nulliparous (yes)	48 (16%)		72 (12%)	
Used HRT in last yr (yes)	123 (40%)		172 (29%)	
Used OC in last yr (yes)	25 (8%)		21 (4%)	
Used OC ever (yes)	140 (45%)		251 (42%)	
Family History[3] (yes)	46 (15%)		53 (9%)	
Currently Smoking (yes)	65 (21%)		93 (16%)	
Social-economic status[4]				
Low	183 (59%)		360 (60%)	
Medium	30 (10%)		71 (12%)	
High	96 (31%)		168 (28%)	
	Median (interquartile range)	Range	Median (interquartile range)	Range
Breast Size (cm^2)	134 (94.3-197)	33-574	118 (86.1-167)	18-638
Dense Tissue (cm^2)	33.8 (20.2-54.4)	0.2-236	28.0 (12.9-47.0)	0.2-302
Non Dense Tissue (cm^2)	91.6 (53.1-147)	14.6-527	81.8 (50.8-127)	5.1-606
Relative Density (%)	28.4 (12.3-46.7)	0.1-82.4	26.0 (8.9-45.4)	0.1-90.8

[1] All postmenopausal women included; n=225 for cases and n=501 for random sample.

[2] Only parous women included; n=261 for cases and n=527 for random sample.

[3] Defined as at least one affected direct relative (mother and/or sister).

[4] Based on level of insurance (social service, civil servant or private).

Over a median follow-up period of 15.5 years, women had a significantly higher chance of developing breast cancer if their breasts showed more dense tissue or more non-dense tissue (table 2). For women in the upper tertile of dense tissue, the risk of breast cancer was 1.73 (95%CI 1.22-2.46), compared to women in the lowest tertile. Women in the upper tertile of non-dense tissue had a risk of 1.44 (95%CI 1.03-2.03) compared to women in the lowest tertile. Women in the upper tertile of relative density only had a non-significant relative risk of 1.17 (95%CI 0.83-1.64) as compared to those in the lowest tertile.

Because a case-cohort design was used, we were also able to compute absolute risks for our cohort of predominantly Caucasian women. The incidence rate for the study cohort is 2.69

cases per 1000 person-years, which is similar to a previous estimate in the larger, original cohort. Table 2 lists the absolute incidence rates for the various strata. The incidence rates range from 1.81 to 3.54 incidences per 1000 person-years of follow up.

Compared to a reference group of women with both below median breast density and below median breast non-density, we assessed the risk of above median absolute density, the risk of above median absolute non-density and the joint effect of both above median density and above median non-density (table 3).

Table 2 Risk of breast cancer in tertiles of breast tissue measures.

Breast Size (cm^2)		Measurement Median (range)	Cases n	Follow-up yr	Incidence IR/1000yr	Hazard HR (95% CI)
	1	75 (18-99)	88	48,314	2.12	1
	2	123 (100-153)	98	45,465	2.52	1.21 (0.85-1.71)
	3	200 (153-638)	123	40,616	3.54	1.68 (1.19-2.37)
Dense Tissue (cm^2)						
	1	9.2 (0.2-20)	78	50,291	1.81	1
	2	30 (20-42)	116	41,501	3.27	1.81 (1.28-2.57)
	3	71 (42-302)	115	42,372	3.17	1.73 (1.22-2.46)
Non-Dense Tissue (cm^2)						
	1	39 (5.1-60)	94	47,429	2.32	1
	2	85 (61-115)	98	45,802	2.50	1.09 (0.77-1.54)
	3	61 (115-606)	117	41,164	3.32	1.44 (1.03-2.03)
Relative Density (%)						
	1	6.6 (0.1-15)	99	45,273	2.55	1
	2	27 (15-40)	97	45,211	2.50	0.99 (0.70-1.39)
	3	55 (40-91)	113	43,911	3.01	1.17 (0.83-1.64)

Table 3 Risk of breast cancer in combined groups of dense tissue and non-dense tissue.

Dense	Non Dense	Relative Density Median (range)	Cases n	Follow-up yr	Incidence IR/1000yr	Hazard HR (95% CI)
Low	Low	26% (0.5-71)	42	27,352	1.79	1
Low	High	7% (0.1-24)	92	43,896	2.45	1.37 (0.89-2.11)
High	Low	54% (26-91)	104	42,068	2.89	1.60 (1.04-2.45)
High	High	28% (5-68)	71	21,079	3.94	2.20 (1.37-3.52)

Women with only a high amount of non-dense tissue had a relative risk of 1.37 (95%CI 0.89-2.11) and women with only a high amount of dense tissue had a relative risk of 1.60 (95%CI 1.04–2.45). The women with both a high amount of dense tissue as well as a high amount of non-dense tissue had a relative risk of 2.20 (95%CI 1.37-3.52) of breast cancer compared to the women who had low amounts of both tissues.

The relative risk in the women with high amounts of both dense and non-dense tissue was higher than the sum of the additional risks of the two groups of women with a high amount of either one tissue. The relative excess risk of this group was however not significant (RERI 0.23, 95%CI -0.66; 1.12).

Analyses in a sub-cohort that excluded women diagnosed with breast cancer within 2 years of the mammogram, in order to rule out masking bias, only showed minor changes to the point estimates of breast cancer risk (data not shown).

Discussion

Our results suggest an influence of non-dense breast tissue on breast cancer risk besides the established influence of dense tissue. Women with either a high amount of dense tissue or a high amount of non-dense tissue are both at increased risk to develop breast cancer. The risk appears highest in women with both a high amount of dense tissue and a high amount of non-dense tissue. The risk of having both high amounts of dense tissue and non-dense tissue appeared to be even higher that than the sum of the independent risks of either tissue, but the calculated relative excess risk of interaction showed that there was no statistically significantly interaction on an additive scale.

The influence of density on breast cancer risk is well known and has been established in many studies[1, 2]. In the current study, the risks for absolute and especially relative density are lower than those reported previously. This is most likely due to limited variation in density in our population in combination with the stratification that we applied. In several previous studies, the reference category was defined as women with less than 5% relative density and the highest relative risk was then reported for women with more than 75% relative density[2]. In our population, these extreme categories had too few women to allow a reliable analysis and we therefore chose to apply tertiles.

Although most attention in literature is on dense tissue and predominantly through assessment of the relative density, the influence of non-dense tissue on breast cancer risk that we find is not unexpected. Boyd et al.[12] recently extended a study by Lam et al.[20] on the independent association of BMI and relative density with breast cancer risk, whereas an increased mortality from breast cancer is also known to be associated with postmenopausal obesity[21]. The result that women with above median amounts of both dense tissue as non-dense tissue were at the highest risk for developing breast cancer indicates that the presence of non-dense tissue increases the risk of breast cancer, whereas this risk of breast density patterns is usually ascribed to dense tissue. As the dense tissue represents the target tissue for breast cancer, the non-dense tissue most likely asserts some modifying influence on the target cells in the dense tissue.

Two possible mechanisms for this effect are local peroxidation of lipids[22, 23] and elevated exposure of the stromal and epithelial cells in the dense tissue to estrogen due to high local levels of estrogen by local production of estrogen by the fat-cells and extended release after the menopause[24-26]. Tamimi et al.[27] recently reported on independent effects of breast density and estrogen on breast cancer risk after a relationship between endogenous sex-hormones and breast density could not be established[28]. This estrogen-pathway may therefore possibly run through non-density.

 Although we do have sufficient information on a number of established determinants of breast cancer and breast density, we chose not to apply an adjustment on our tissue

measurements with these. The rationale for this is that the known and established determinants of density, non-density and relative density tissue are not the same and that uniform adjustment was therefore not warranted. As the aim of our study was to compare the three structures, rather than to determine the most accurate estimate of risk associated with each tissue type, we felt that not applying an adjustment was more suitable than applying different, measure-dependent adjustments. This lack of adjustment may also be a further explanation to our relatively low cancer risk estimates associated with relative density.

One of the strengths of the present study is the case-cohort design. Through the extrapolation of randomly sampled data to a large cohort, the results obtain a firm basis. Furthermore, the current cohort is recruited from the general population with only age and place of residence as selective variables. Also, due to the long follow-up, densities of well before (determinable) onset of the disease can be investigated. These aspects make the results reliable and able to be generalised within the studies' population, which primarily consists of Dutch, Caucasian women. The Dutch population has been known for decades to have one of the highest incidences of breast cancer in the world.

In summary, we show that non-dense tissue may also play a role in the risk of breast cancer, especially in combination with high amounts of dense tissue. This implies joint influences of non-dense and dense tissue that modulate the risk of breast cancer. Local lipid peroxidation and local estrogen levels are possible underlying mechanisms. These results need to be confirmed by other studies as they offer new insight into the origin and modification of breast density, the most prominent risk factor for breast cancer.

Reference List

1. Boyd NF, Guo H, Martin LJ et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3):227-236.
2. McCormack VA, dos SS, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006; 15(6):1159-1169.
3. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976; 37(5):2486-2492.
4. American College of Radiology. Breast Imaging reporting and data system, breast imaging atlas. 4th ed. Reston, VA: American College of Radiology, 2003. 2008.
5. Gram IT, Funkhouser E, Tabar L. The Tabar classification of mammographic parenchymal patterns. Eur J Radiol 1997; 24(2):131-136.
6. Byng JW, Yaffe MJ, Lockwood GA, Little LE, Tritchler DL, Boyd NF. Automated analysis of mammographic densities and breast carcinoma risk. Cancer 1997; 80(1):66-74.
7. Highnam R, Jeffreys M, McCormack V, Warren R, Davey SG, Brady M. Comparing measurements of breast density. Phys Med Biol 2007; 52(19):5881-5895.
8. Yaffe MJ, Boyd NF, Byng JW et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998; 7 Suppl 1:S47-S55.
9. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer 1998; 78(9):1233-1238.
10. Friedenreich CM. Review of anthropometric factors and breast cancer risk. Eur J Cancer Prev 2001; 10(1):15-32.
11. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control 2000; 11(7):653-662.
12. Boyd NF, Martin LJ, Sun L et al. Body size, mammographic density, and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2006; 15(11):2086-2092.
13. Haars G, van Noord PAH, van Gils CH, Grobbee DE, Peeters PHM. Measurements of breast density; no ratio for a ratio. Cancer Epidemiol Biomarkers Prev 2005; 14(11):2634-2640.
14. Collette HJ, Day NE, Rombach JJ, de Waard F. Evaluation of screening for breast cancer in a non-randomised study (the DOM project) by means of a case-control study. Lancet 1984; 1(8388):1224-1226.
15. de Waard F, Collette HJ, Rombach JJ, Baanders-van Halewijn EA, Honing C. The DOM project for the early detection of breast cancer, Utrecht, The Netherlands. J Chronic Dis 1984; 37(1):1-44.
16. Gao J, Warren R, Warren-Forward H, Forbes JF. Reproducibility of visual assessment on mammographic density. Breast Cancer Res Treat 2008; 108(1):121-127.
17. Barlow WE. Robust variance estimation for the case-cohort design. Biometrics 1994; 50(4):1064-1072.
18. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. J Clin Epidemiol 1999; 52(12):1165-1172.
19. Onland-Moret NC, van der A D, van der Schouw YT et al. Analysis of case-cohort data: a comparison of different methods. J Clin Epidemiol 2007; 60(4):350-355.
20. Lam PB, Vacek PM, Geller BM, Muss HB. The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in Vermont. Cancer 2000; 89(2):369-375.
21. Calle EE, Rodriguez C, Walker-Thurmond K, Thun MJ. Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults. N Engl J Med 2003; 348(17):1625-1638.
22. Boyd NF, McGuire V. Evidence of lipid peroxidation in premenopausal women with mammographic dysplasia. Cancer Lett 1990; 50(1):31-37.
23. Boyd NF, McGuire V. The possible role of lipid peroxidation in breast cancer risk. Free Radic Biol Med 1991; 10(3-4):185-190.
24. Jefcoate CR, Liehr JG, Santen RJ et al. Tissue-specific synthesis and oxidative metabolism of estrogens. J Natl Cancer Inst Monogr 2000;(27):95-112.
25. Kaaks R. Endogenous hormone metabolism as an exposure marker in breast cancer chemoprevention studies. IARC Sci Publ 2001; 154:149-62.:149-162.
26. Thijssen JH. Local biosynthesis and metabolism of oestrogens in the human breast. Maturitas 2004; 49(1):25-33.

27. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 2007; 99(15):1178-1187.

28. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. Breast Cancer Res 2008; 10(1):201.

6

The next phase in breast density studies

Background

Since the first classification of mammographic breast patterns by Wolfe, evaluation concentrated on relative density of the breast[1]. New, computer assisted techniques that were developed to measure breast density stimulated the transition from a crude, categorical evaluation to more detailed measurements on a continuous scale[2-4]. Over the three decades that have passed since Wolfe's position paper, it has become overwhelmingly clear that a high relative density increases the risk of breast cancer[5, 6]. The increase in risk varies depending on the method of measurement and the populations studied, but the conclusions were consistent. A recent meta-analysis including 42 studies estimated the breast cancer risk associated with 75% breast density or over at a four to five times higher compared to less than 5% breast density[6]. In this regard, relative breast density has established its value as a risk factor in breast cancer research and may even serve as a useful factor in prediction of breast cancer risk[7].

Remarkably, it has remained unclear for a long time what the dense tissue actually consists of. It was originally thought that the density represented glandular tissue, but more recently it has become clear that it is the combination of glandular and stromal tissue that cause a dense appearance of the breast on mammography and that interactions between the two tissue types may play a role in breast tumour development[8-10]. Definitive pathways are, however, not yet known and need much more investigation, but offer important potential for new relevant insight and modes of prevention.

Breast density and determinants of breast cancer

Childbirth, having gone through menopause and age in general, are well-established determinants of density level[11-18]. They account, however, only for a small proportion of the variability in relative density[19]. Numerous studies have been addressing a range of possible determinants and pathways but have not yet been able to establish the precise mechanisms that underlie the formation and modification of density.

Evidence had been accumulating for some time to suggest a relation between the Insulin-like Growth Factor I (IGF-I) system, breast cancer and density level, with several reports of a positive relation between serum levels of IGF-I and the ratio of IGF-I and IGF Binding Protein 3 (IGFBP-3) and mammographic density level. A summary of these reports is given by Martin and Boyd, indicating that this relation between the IGF-system and density has been observed more clearly in premenopausal than in postmenopausal women[9]. Some additional recent studies, including our own work, however, added to the literature not confirming a relation[20, 21]. At best, the association between the IGF-system and breast density as a possible pathway between the IGF-system and breast cancer is therefore still under debate.

The role of endogenous sex-hormones has also been studied extensively, as they are known to be strongly related to breast cancer risk[22-24]. Van Duijnhoven et al.[25] did find a relation between polymorphisms in the estrogen receptor and breast density, but there is little or no evidence to support a relation between systemic endogenous hormone levels and

mammographic density as can be seen from the review by Martin and Boyd[9]. In a recent study, Tamimi et al. actually showed that the effects of endogenous hormone levels and density on breast cancer risk were independent of one another[26].

There are more examples of breast cancer risk factors that, while increasing breast cancer risk, do not appear to modify breast density. A short but intense period of caloric restriction was shown to increase the risk of breast cancer later in life[27], without affecting breast density (chapter4). Similarly, the beneficial effect of physical activity on breast cancer risk could also not be ascribed to changes in mammographic density patterns[28-35]. An obvious explanation would be that there are other pathways leading to breast cancer independent from increased breast density.

Breast density measures

What also may have contributed to conflicting reports regarding mechanisms involving breast density is the emphasis on 'relative' density measures in earlier years. More recently, the number of studies including both absolute and percentage measurements of density has increased. The relative density is a combined score that may obscure the direct biological pathway of the mechanism under investigation. It combines the effects of two component tissues; dense and non-dense tissue. We showed that some of the established determinants of relative density have a less pronounced, or even no relationship with the actual dense tissue[36]. Age at time of examination and parity were found to have opposing effects on the amounts of dense and non-dense tissue, which enhances their effects on the relative density. BMI did not show any relationship with the absolute amount of dense tissue, the target tissue for breast cancer, whereas it is an established determinant of the relative density. These results show that to uncover etiologic aspects of breast density, absolute measures of breast tissues are more straightforward and may give a better insight in the underlying pathways.

It is not very likely that there is a physiological mechanism to support the use of a relative density measure. It is more likely that there are mechanisms that influence the amount of dense tissue and in addition there are (possibly other) mechanisms that influence the amount of non-dense tissue. Consequently, absolute measures of dense and non-dense tissue appear to be better equipped to make inferences about the developmental pathways of breast density.

The simplest explanation that has been proposed for the relationship between density and breast cancer risk is that the risk of breast cancer increases with the absolute amount of vulnerable glandular breast tissue present[37, 38]. This also supports the use of absolute measures.

Another limitation of the use of the relative density is that it ignores the effects of its denominator, total breast size. A publication by Nagata et al. as well as our data presented in chapter 5 show that this may not be justified[39]. In both studies, total breast size was found to be related to breast cancer risk. These effects were most likely due to effects of non-dense tissue size reflected in total breast size.

Our findings are the first to directly suggest that the amount of non-dense tissue in the breast, which includes fat tissue, may add to the total sum of breast cancer risk (chapter 5). Women that proved to be at highest risk were those in the group with high amounts of both dense and non-dense tissue. In particular the role of non-dense tissue in the association of levels and patterns of breast density and breast cancer risk has so far not received much attention in literature. Further investigation of this role may prove useful for better assessment of breast cancer risk as well as better understanding of etiologic pathways.

The fat that constitutes most of the non-dense tissue may confer a higher breast cancer risk through higher levels of mutagens from lipid peroxidation[40, 41]. Alternatively, it may influence breast cancer risk by increased and prolonged exposure of the dense tissue to estrogens. Although a relation between systemic estrogen levels and breast density is not supported by available data[9], local levels of estrogen are higher in the breast through local production by fat cells and continued release after menopause[42-44].

Room for improvement

Despite the biological arguments to the benefit of using 'absolute' measures of dense and non-dense tissue and the practical benefit of percent density as an indicator of risk in routine practice, there are limitations to all currently available measurement methods for breast density.

The main limitation of the current measurement method is that the measured surface area on a mammogram is a 2-dimensional representation of a 3-dimensional structure. Variation can be caused by the amount of breast tissue that the radiologist manages to include in the mammogram and also in the amount of pressure that is applied to compress the breast at mammography. In addition, compression may be different for different breast structures. Information about the compression-thickness of the breast could be collected or 3-dimensional techniques such as MRI (magnetic resonance imaging) could be applied so that density can be assessed volumetrically[45]. The use of these methods is particularly important if one is interested in measuring changes in breast density over time. The (random) measurement error described above will occur in each of the density measurements, increasing the chance of observing artificial increases or decreases in density over time. Despite this 'noise', it has been possible to estimate biological changes in density over time[46-50], but the volumetric method will allow more precise estimation of effects.

Longitudinal effects

The importance of studying changes in density over time can be illustrated by the model of Pike et al. on 'breast tissue ageing' due to hormonal risk factors[51]. In this model, the breast tissue age is regarded as a reflection of the absolute risk of developing breast cancer. As the breast tissue age increases, so does the breast cancer risk. The breast tissue age is assessed as the cumulative exposure of the breast tissue to hormones such as oestrogen and prolactin. As of menarche, the cumulative exposure, and thus breast tissue age, starts increasing at a high rate due to high hormone levels. After a first full term pregnancy and

especially during the menopause, hormonal levels decrease. Subsequently, the cumulative exposure and thereby breast tissue age still increases, but at a lower rate.

This model was shown to follow the breast cancer incidence pattern in the USA. Before the average age of menopause, breast cancer incidence increases strongest, as does the breast tissue age. Parallel to the decrease in hormone levels after menopause and the subsequent slower increase in breast tissue age, the incidence of breast cancer also increases less. So as proposed by Pike et al., absolute breast cancer risk and the incidence of breast cancer increases throughout life, but the rate of increase in these slows down as the breast tissue starts to age more slowly, due to lower levels of hormones.

As Boyd et al. recently concluded, these developments are in line with the changes in (relative) breast density during a woman's life[52]. The most ideal predictor of breast cancer risk should thus not only incorporate measures such as the absolute amount of dense tissue, the absolute amount of non-dense tissue and the relative density, but possibly also longitudinal information on breast development.

Three studies have since then investigated longitudinal effects in large-scale studies. In the first, Maskarinec et al.[47] showed that the cumulative percent density and the breast cancer incidence rate indeed increase in parallel, as proposed by Boyd et al.[52] However, they did not find a direct relation between case-status and the rate of change in density in a multilevel growth model. In a matched case control study, Vachon et al.[48] also reported that the rate of change in percent density (assessed as the average individual difference over time) was not an indicator of subsequent breast cancer risk.

The most extensive of the three is a study in more than 300,000 women on changes in the BI-RADS (Breast Imaging Reporting and Data System) breast density score. This classification system of the American College of Radiology is not specifically aimed at categorising breast density, but is widely used in routine care and also allows for assessment of breast cancer[53,54]. In the study, Kerlikowske at al.[46] found that an increase in BI-RADS breast density score within 3 years of the assessment was associated with a corresponding increase in breast cancer risk, whereas a decrease in BI-RADS score indicated a lower risk.

Conclusion

Breast density is an important breast cancer risk factor but there remains ample room for further investigation into the origin and modification of breast density. To achieve this, first and foremost improvements are necessary in the assessment and interpretation of breast density. In any case, volumetric and longitudinal measures of either relative or absolute density are likely to be the better options for the future. The application of these in studies on known or disputed determinants and on the role of density in pathways leading to breast cancer may help to finish ongoing debates and provide new insights.

Reference List

1. Wolfe JN. Risk for breast cancer development determined by mammographic parenchymal pattern. Cancer 1976; 37(5):2486-2492.
2. Byng JW, Yaffe MJ, Lockwood GA, Little LE, Tritchler DL, Boyd NF. Automated analysis of mammographic densities and breast carcinoma risk. Cancer 1997; 80(1):66-74.
3. Highnam R, Jeffreys M, McCormack V, Warren R, Davey SG, Brady M. Comparing measurements of breast density. Phys Med Biol 2007; 52(19):5881-5895.
4. Yaffe MJ, Boyd NF, Byng JW et al. Breast cancer risk and measured mammographic density. Eur J Cancer Prev 1998; 7 Suppl 1:S47-S55.
5. Boyd NF, Guo H, Martin LJ et al. Mammographic density and the risk and detection of breast cancer. N Engl J Med 2007; 356(3):227-236.
6. McCormack VA, dos SS, I. Breast density and parenchymal patterns as markers of breast cancer risk: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2006; 15(6):1159-1169.
7. Vachon CM, van Gils CH, Sellers TA et al. Mammographic density, breast cancer risk and risk prediction. Breast Cancer Res 2007; 9(6):217.
8. Li T, Sun L, Miller N et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. Cancer Epidemiol Biomarkers Prev 2005; 14(2):343-349.
9. Martin LJ, Boyd NF. Mammographic density. Potential mechanisms of breast cancer risk associated with mammographic density: hypotheses based on epidemiological evidence. Breast Cancer Res 2008; 10(1):201.
10. Patocs A, Zhang L, Xu Y et al. Breast-cancer stromal cells with TP53 mutations and nodal metastases. N Engl J Med 2007; 357(25):2543-2551.
11. Boyd NF, Lockwood GA, Byng JW, Little LE, Yaffe MJ, Tritchler DL. The relationship of anthropometric measures to radiological features of the breast in premenopausal women. Br J Cancer 1998; 78(9):1233-1238.
12. Butler LM, Gold EB, Greendale GA et al. Menstrual and reproductive factors in relation to mammographic density: the Study of Women's Health Across the Nation (SWAN). Breast Cancer Res Treat 2007; .
13. Byrne C, Schairer C, Wolfe J et al. Mammographic features and breast cancer risk: effects with time, age, and menopause status. J Natl Cancer Inst 1995; 87(21):1622-1629.
14. de Waard F, Rombach JJ, Collette HJ, Slotboom B. Breast cancer risk associated with reproductive factors and breast parenchymal patterns. J Natl Cancer Inst 1984; 72(6):1277-1282.
15. El Bastawissi AY, White E, Mandelson MT, Taplin SH. Reproductive and hormonal factors associated with mammographic breast density by age (United States). Cancer Causes Control 2000; 11(10):955-963.
16. Gram IT, Funkhouser E, Tabar L. Reproductive and menstrual factors in relation to mammographic parenchymal patterns among perimenopausal women. Br J Cancer 1995; 71(3):647-650.
17. Titus-Ernstoff L, Tosteson AN, Kasales C et al. Breast cancer risk factors in relation to breast density (United States). Cancer Causes Control 2006; 17(10):1281-1290.
18. Vachon CM, Kuni CC, Anderson K, Anderson VE, Sellers TA. Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States). Cancer Causes Control 2000; 11(7):653-662.
19. Maskarinec G, Pagano I, Chen Z, Nagata C, Gram IT. Ethnic and geographic differences in mammographic density and their association with breast cancer incidence. Breast Cancer Res Treat 2007; 104(1):47-56.
20. Maskarinec G, Takata Y, Chen Z et al. IGF-I and mammographic density in four geographic locations: a pooled analysis. Int J Cancer 2007; 121(8):1786-1792.
21. Verheus M, Peeters PH, van Noord PA, van der Schouw YT, Grobbee DE, van Gils CH. No relationship between circulating levels of sex steroids and mammographic breast density: the Prospect-EPIC cohort. Breast Cancer Res 2007; 9(4):R53.
22. Eliassen AH, Missmer SA, Tworoger SS et al. Endogenous steroid hormone concentrations and risk of breast cancer among premenopausal women. J Natl Cancer Inst 2006; 98(19):1406-1415.
23. Key T, Appleby P, Barnes I, Reeves G. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. J Natl Cancer Inst 2002; 94(8):606-616.

24. Missmer SA, Eliassen AH, Barbieri RL, Hankinson SE. Endogenous estrogen, androgen, and progesterone concentrations and breast cancer risk among postmenopausal women. J Natl Cancer Inst 2004; 96(24):1856-1865.

25. Van Duijnhoven FJ, Bezemer ID, Peeters PH et al. Polymorphisms in the estrogen receptor alpha gene and mammographic density. Cancer Epidemiol Biomarkers Prev 2005; 14(11 Pt 1):2655-2660.

26. Tamimi RM, Byrne C, Colditz GA, Hankinson SE. Endogenous hormone levels, mammographic density, and subsequent risk of breast cancer in postmenopausal women. J Natl Cancer Inst 2007; 99(15):1178-1187.

27. Elias SG, Peeters PH, Grobbee DE, van Noord PA. Breast cancer risk after caloric restriction during the 1944-1945 Dutch famine. J Natl Cancer Inst 2004; 96(7):539-546.

28. Irwin ML, Aiello EJ, McTiernan A et al. Pre-diagnosis physical activity and mammographic density in breast cancer survivors. Breast Cancer Res Treat 2006; 95(2):171-178.

29. Irwin ML, Aiello EJ, McTiernan A et al. Physical activity, body mass index, and mammographic density in postmenopausal breast cancer survivors. J Clin Oncol 2007; %20;25(9):1061-1066.

30. Peters TM, Ekelund U, Leitzmann M et al. Physical activity and mammographic breast density in the EPIC-Norfolk cohort study. Am J Epidemiol 2008; 167(5):579-585.

31. Reeves KW, Gierach GL, Modugno F. Recreational physical activity and mammographic breast density characteristics. Cancer Epidemiol Biomarkers Prev 2007; 16(5):934-942.

32. Samimi G, Colditz GA, Baer HJ, Tamimi RM. Measures of energy balance and mammographic density in the Nurses' Health Study. Breast Cancer Res Treat 2008; 109(1):113-122.

33. Sellers TA, Vachon CM, Pankratz VS et al. Association of childhood and adolescent anthropometric factors, physical activity, and diet with adult mammographic breast density. Am J Epidemiol 2007; 166(4):456-464.

34. Siozon CC, Ma H, Hilsen M, Bernstein L, Ursin G. The association between recreational physical activity and mammographic density. Int J Cancer 2006; 119(7):1695-1701.

35. Suijkerbuijk KP, Van Duijnhoven FJ, van Gils CH et al. Physical activity in relation to mammographic density in the dutch prospect-European prospective investigation into cancer and nutrition cohort. Cancer Epidemiol Biomarkers Prev 2006; 15(3):456-460.

36. Haars G, van Noord PA, van Gils CH, Grobbee DE, Peeters PH. Measurements of breast density: no ratio for a ratio. Cancer Epidemiol Biomarkers Prev 2005; 14(11 Pt 1):2634-2640.

37. Albanes D, Winick M. Are cell number and cell proliferation risk factors for cancer? J Natl Cancer Inst 1988; 80(10):772-774.

38. Trichopoulos D, Lipman RD. Mammary gland mass and breast cancer risk. Epidemiology 1992; 3(6):523-526.

39. Nagata C, Matsubara T, Fujita H et al. Mammographic density and the risk of breast cancer in Japanese women. Br J Cancer 2005; %20;92(12):2102-2106.

40. Boyd NF, McGuire V. Evidence of lipid peroxidation in premenopausal women with mammographic dysplasia. Cancer Lett 1990; 50(1):31-37.

41. Boyd NF, McGuire V. The possible role of lipid peroxidation in breast cancer risk. Free Radic Biol Med 1991; 10(3-4):185-190.

42. Dabrosin C. Increased extracellular local levels of estradiol in normal breast in vivo during the luteal phase of the menstrual cycle. J Endocrinol 2005; 187(1):103-108.

43. Kaaks R. Endogenous hormone metabolism as an exposure marker in breast cancer chemoprevention studies. IARC Sci Publ 2001; 154:149-62.:149-162.

44. Thijssen JH. Local biosynthesis and metabolism of oestrogens in the human breast. Maturitas 2004; 49(1):25-33.

45. van Engeland S, Snoeren PR, Huisman H, Boetes C, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms. IEEE Trans Med Imaging 2006; 25(3):273-282.

46. Kerlikowske K, Ichikawa L, Miglioretti DL et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst 2007; 99(5):386-395.

47. Maskarinec G, Pagano I, Lurie G, Kolonel LN. A longitudinal investigation of mammographic density: the multiethnic cohort. Cancer Epidemiol Biomarkers Prev 2006; 15(4):732-739.

48. Vachon CM, Pankratz VS, Scott CG et al. Longitudinal trends in mammographic percent density and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2007; 16(5):921-928.

49. Van Duijnhoven FJ, Peeters PH, Warren RM et al. Postmenopausal hormone therapy and changes in mammographic density. J Clin Oncol 2007; 25(11):1323-1328.

50. Verheus M, Peeters PH, Kaaks R, van Noord PA, Grobbee DE, van Gils CH. Premenopausal insulin-like growth factor-I serum levels and changes in breast density over menopause. Cancer Epidemiol Biomarkers Prev 2007; 16(3):451-457.

51. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. 'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer. Nature 1983; 303(5920):767-770.

52. Boyd NF, Rommens JM, Vogt K et al. Mammographic breast density as an intermediate phenotype for breast cancer. Lancet Oncol 2005; 6(10):798-808.

53. Kerlikowske K, Ichikawa L, Miglioretti DL et al. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. J Natl Cancer Inst 2007; 99(5):386-395.

54. Vacek PM, Geller BM. A prospective study of breast cancer risk using routine mammographic breast density measurements. Cancer Epidemiol Biomarkers Prev 2004; 13(5):715-722.

Summaries and acknowledgements

Summary
Samenvatting
Dankwoord
Curriculum Vitae

7
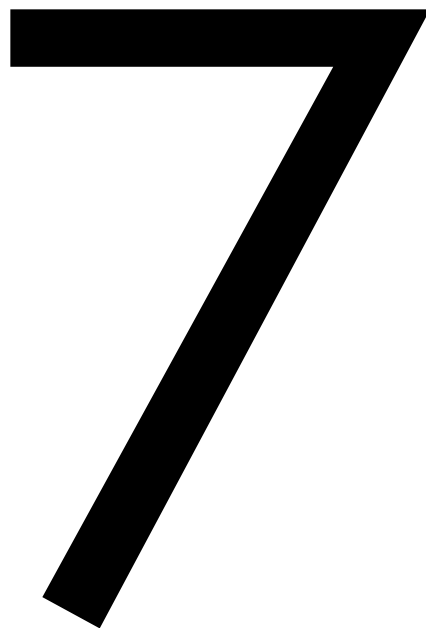
Summary – The new look in short

Breast parenchymal patterns, as visible on mammograms, are determined by the relative amount of radio dense tissue. This radio density comprises connective and epithelial tissue and can be seen to represent the glandular target tissue for breast cancer. First scored categorically through rough estimation of the relative density, current computerised techniques now enable more accurate, continuous measurement of density. High percentages of density are established to be one of the strongest risk factors of breast cancer risk. It has been hypothesised that the actual number of target cells is directly related to breast cancer. This has however not been established as such, which may be due to the use of the relative density in nearly all studies, rather than absolute amounts of dense tissue. There has been much research into the origin and modification of breast density and the way through which it influences breast cancer risk, but the biology is still largely unknown. Quite a number of determinants of relative density have been found of which some have been established and others have been or are being strongly debated.

In this thesis we aim to unravel more about the biology of breast density and the effect that using a relative measure has had on current knowledge. We do this by re-investigating known factors with absolute measures of breast density and non-density.

In Chapter 2 we investigated the heritability of breast density and we used this assessment as a reference to test a newly developed statistical model for sequential analysis. The heritability was studied to give a base estimate of the extent to which breast density is determined by 'innate' factors, and which room is left for other variables that are more susceptible to modification.

In chapter 2.1 we assessed heritability in dichotomously scored breast patters (high risk versus low risk) in 466 sister-, 25 dizygotic twin- and 26 monozygotic twin-pairs. This gave estimates of 38% for sisters, 34% for dizygotic twins and 88% for monozygotic twins. A model that combines monozygotic twins with dizygotic twins or sisters yielded an estimate of more than 100%, which indicated that dominant gene effects, genetic interactions or gene–environment effects could be involved. The last effect was seen for parity as heritability estimates ranging from 90% in sisters that were both nulliparous, to 2% in sister pairs discordant for nulliparity. In total the results indicated a substantial genetic influence, but a possible modifying ability of other factors, such as parity.

In chapter 2.2 we used continuous measurements of absolute and relative density from 19 monozygotic twin pairs and 79 sister pairs to estimate heritability with a sequential analysis approach. An interim analysis was performed for each set of 1 monozygotic twin pair vs. 3 sister pairs. As the 19 monozygotic pairs were not adequate to come to a result, further sets were randomly sampled from the available measurements. The obtained results were therefore only a modelled outcome and not a true estimate. After 35 sets, the heritability of the amount of dense tissue was estimated at 74% and after 27 sets, the estimate for percent

density was 102%. Besides adding these rough estimates of the heritability of continuously measured amount of dense tissue and percent density to our knowledge about breast density, this study showed that the methodology worked and that sequential analyses can be a valuable tool is the sample size of a study needs to be reduced. The 35 sets needed for the analysis of the amount of dense tissue is only 41% of the amount of sets that would have been needed for a fixed sample size analysis.

In chapter 3 we aimed to assess whether there were differences between the effects of known determinants on relative density and on the absolute amounts of dense and non-dense tissues. We measured the absolute and relative densities on digitized mammograms in a population of 418 postmenopausal women. Information on determinants was available through physical examination and questionnaires. We found that age and parity decrease the amount of dense tissue and that the ages at menarche and menopause increase it (R^2 = 13%). The amount of non-dense tissue was increased by higher body mass index (BMI), age, and parity (R^2 = 43%). Relative density was affected by a combination of these factors (R^2 = 29%) in the same direction as dense tissue. However, the magnitudes of these effects were the resultant of the effects on dense and non-dense tissues. The influence of BMI on relative density was completely due to an effect on non-dense tissue. These results showed that, although relative density is a relevant prognostic factor, inferences about the aetiology of breast density should better be made on the basis of absolute measures.

In chapter 4, a previous study by our group on the relation between the Dutch famine and breast cancer risk is extended to breast density. Breast size, dense and non-dense tissue and the relative density were measured on a continuous scale from the mammograms of 1035 women. Mean densities were compared between three groups of ascending levels of famine-exposure; unexposed, moderately exposed and severely exposed. These levels were based on self reported information on exposure to cold and hunger and suffering from weight loss. The results of the measurements were adjusted for known determinants of breast density and stratified to age at exposure. Overall, there were no significant differences by exposure. The mean breast size varied from 124 cm^2 in unexposed to 121 cm^2 (p_{trend} =0.50) in severely exposed women. Mean amount of dense tissue varied from 23.4 to 21.8 cm^2 (p_{trend} =0.48), and mean amount of non-dense tissue from 87.7 to 85.4 cm^2 (p_{trend} =0.55). The relative density varied from 22.8 to 22.3% (p_{trend} =0.78). When analysed according to groups of 'age at exposure', the amount of non-dense tissue was significantly lower with higher exposure before age 10, with 53.1 cm^2 for severely exposed as to 77.8 cm^2 (p_{trend}=0.03) for unexposed. This group also appeared to have smaller breasts with more absolute and relative density, but these results were not statistically significant.

In chapter 5 we report on the risk of breast cancer associated with absolute measures of breast density compared to the established risk associated with percent density. In a case-cohort design within DOM, all newly diagnosed invasive breast cancers in 15.5 years of

follow-up were determined (361) and follow up experience was obtained for a random cohort sample of 600 women. The mammograms could be obtained, digitized and evaluated for 309 cases and the cohort sample. The risk of breast cancer was then compared between tertiles of either dense tissue, non-dense tissue or relative density and in combined categories of dense tissue and non-dense tissue (low/low vs. low/high, high/low and high/high). We found that breast cancer risk tended to increase independently with increasing amounts of density (HR 1.68; 95%CI 1.22-2.46 for highest vs. lowest tertile) and non-density (HR 1.44; 95%CI 1.03-2.03 for highest vs. lowest tertile). The highest risk for breast cancer was found in the women that had above median amounts of both tissues (HR 2.20; 95%CI 1.37-3.52) compared to women with below median amounts of both tissues. This indicated that non-dense tissue plays a role in the development of breast cancer that had not yet received much attention. Especially the result that having high amounts of both tissues gave the highest risk estimate is interesting as thus far the highest risk had always been associated with the highest percent density.

In chapter 6 we evaluate our results from the comparisons of absolute and relative measures of (non-)density. These results indicate that absolute density measures are absolutely a valuable resource in etiologic research, as they yield different information than the more established relative density measurements. Also the role of non-density should receive more attention. However, the currently available outcome measures also have shortcomings and are not the final, ideal measure of breast density. Volumetric and longitudinal information would provide additional relevant data and options for new insights into breast density.

Samenvatting – De nieuwe kijk in het kort

Parenchymale borstpatronen, die zichtbaar zijn op mammogrammen, worden bepaald door de relatieve hoeveelheid radiodicht weefsel. Dit radiodichte weefsel (densiteit) bevat bindweefsel en epitheliaal weefsel en kan worden gezien als een graadmeter van het weefsel dat gevoelig is voor borstkanker. Waar de densiteit eerst werd beoordeeld in categorieën op basis van ruwe schatting van de relatieve dichtheid, maken huidige computergestuurde technieken het mogelijk om de densiteit nauwkeuriger en op een continue schaal te meten. Hoge percentages densiteit zijn bekend als een van de sterkste risico factoren voor borstkanker. Het is voorgesteld dat het daadwerkelijke aantal borstweefselcellen dat tot een tumor uit zou kunnen groeien een direct verband heeft met de kans op borstkanker. Deze relatie is echter nog nooit direct vastgesteld, wat te wijten zou kunnen zijn aan het gebruik van de relatieve densiteit in de meeste studies, in plaats van de absolute hoeveelheid radiodicht weefsel.

Er is veel onderzoek gedaan naar de oorsprong en de beïnvloeding van borstdensiteit en de manier waarop dit met het risico op borstkanker is gerelateerd, maar de precieze biologie is nog grotendeels onbekend. Een aanzienlijk aantal factoren is al genoemd als determinant van de relatieve densiteit . Enkelen daarvan zijn goed geaccepteerd, terwijl de rol van anderen nog sterk ter discussie staat.

In dit proefschrift willen we meer van de biologie van borstdensiteit ontrafelen en bepalen welk effect het gebruik van een relatieve maat voor de dichtheid op de huidige kennis heeft gehad. We doen dit door bekende factoren opnieuw te onderzoeken met gebruik van absolute maten van borstdensiteit en niet-densiteit.

In hoofdstuk 2 onderzoeken we de overerfelijkheid (heritabiliteit) van borstdensiteit en gebruiken we deze bepaling als referentie om een nieuw ontwikkelde statistische methode voor sequentiële analyse te testen. De heritabiliteit werd onderzocht om een basis schatting te geven van de mate waarin borstdensiteit wordt bepaald door aangeboren factoren en welke ruimte er nog is voor variabelen die gevoeliger zijn voor beïnvloeding.

In hoofdstuk 2.1 bepaalden we de heritabiliteit in dichotoom gescoorde borstpatronen (hoog vs. laag risico) in 466 zusterparen, 25 dizygote tweelingparen en 26 monozygote tweelingparen. Dit gaf een schatting van 38% voor zussen, 34% voor dizygoten en 88% voor monozygote tweelingen. Een model dat monozygote tweelingen combineert met dizygote tweelingen of zussen gaf een schatting van de heritabiliteit van meer dan 100%. Dit wijst er op dat er dominante geneffecten, genetische interactie of gen-omgevings effecten zouden kunnen zijn. Een gen-omgevings effect werd gezien voor pariteit, waar de schatting van de heritabiliteit varieerde van 90% voor zussen die beide nullipaar waren, tot 2% voor zussen waarvan er één nullipaar was en de andere niet. In zijn totaliteit gaven de resultaten aan dat er een aanzienlijke genetische invloed is, maar met een mogelijkheid voor beïnvloeding door andere factoren, zoals pariteit.

In hoofdstuk 2.2 gebruikten we continue metingen van absolute en relatieve densiteit van 17 monozygote tweelingparen en 79 zusterparen om de heritabiliteit te schatten met een sequentiële analyse methode. Na elke set van 1 monozygoot paar vs. 3 zusterparen werd een interim analyse uitgevoerd. Omdat er na 19 tweelingparen nog geen conclusie kon worden getrokken werden voor het testen van het model willekeurig verdere sets getrokken uit de beschikbare metingen. De verkregen uitkomsten zijn daardoor alleen een gemodelleerde uitkomst en geen ware schatting. Na 35 sets kon de heritabiliteit van de absolute densiteit worden geschat op 74% en na 27 sets was de schatting voor percentuele densiteit 102%. Naast het toevoegen van deze ruwe schattingen van de heritabiliteit van continue gemeten absolute en relatieve densiteit aan de bestaande kennis over borstdensiteit, gaf deze studie aan dat de methodologie werkte en dat sequentiële analyse een waardevol middel kan zijn als de groepsgrootte van een studie moet worden beperkt. De 35 sets die nodig waren voor de analyse van de hoeveelheid densiteit is maar 41% van het aantal sets dat nodig zou zijn geweest voor een analyse met een vaste groepsgrootte.

In hoofdstuk 3 beoogden we vast te stellen of er verschillen zijn tussen de effecten van bekende determinanten op relatieve densiteit en op de absolute hoeveelheden dense en niet-dense weefsel. De absolute en relatieve densiteit werd gemeten op de gedigitaliseerde mammogrammen van 418 postmenopauzale vrouwen. Informatie over determinanten was beschikbaar uit lichamelijk onderzoek en vragenlijsten. We vonden dat hogere leeftijd en pariteit de hoeveelheid dense weefsel verminderd en dat de hogere leeftijden bij menarche en menopauze een toename veroorzaken (R^2 = 13%). De hoeveelheid niet-densiteit nam toe met hogere 'body mass index' (BMI), leeftijd en pariteit (R^2 = 43%). De relatieve densiteit werd beïnvloed door een combinatie van deze factoren (R^2 = 29%) in dezelfde richting als voor de hoeveelheid dense weefsel. De grootte van de effecten op relatieve densiteit waren echter een resultante van de effecten op de hoeveelheden dense en niet-dense weefsel.
De invloed van BMI op de relatieve densiteit was volledig te wijten aan een verband met het niet-dense weefsel. Deze resultaten gaven aan dat, alhoewel de relatieve densiteit een belangrijke prognostische factor is, uitspraken over de etiologie van borstdensiteit beter zouden moeten worden gemaakt aan de hand van absolute metingen.

In hoofdstuk 4 verlengen we een eerder onderzoek door onze groep naar de relatie tussen de Nederlandse Hongerwinter en het risico op borstkanker naar borstdensiteit. Borstgrootte, de hoeveelheden dense en niet-dense weefsel en de relatieve densiteit werden op een continue schaal gemeten op de mammogrammen van 1035 vrouwen. Gemiddelde densiteiten werden vergeleken tussen 3 niveaus van blootstelling aan de Hongerwinter; geen, matige of zware blootstelling. Deze niveaus waren gebaseerd op zelf gerapporteerde blootstelling aan koude, honger en gewichtsafname. De resultaten van de metingen werden gecorrigeerd voor bekende determinanten van borstdensiteit en gestratificeerd naar de leeftijd bij blootstelling.

Over het geheel waren er geen statistisch significante verschillen door de blootstelling. De gemiddelde borstgrootte varieerde van 124 cm^2 in niet blootgestelden tot 121 cm^2 (p_{trend} =0.50) in zwaar blootgestelde vrouwen. De gemiddelde hoeveelheid dense weefsel varieerde van 23.4 tot 21.8 cm^2 (p_{trend} =0.48) en de gemiddelde hoeveelheid niet-dense weefsel van 87.7 tot 85.4 cm^2 (p_{trend} =0.55). De relatieve densiteit varieerde van 22.8 tot 22.3% (p_{trend} =0.78). In de analyses naar leeftijdsgroep bij blootstelling was alleen de hoeveelheid niet-dense weefsel statistisch significant lager bij blootstelling voor 10 jaar met 53.1 cm^2 voor zwaar blootgestelden tot 77.8 cm^2 (p_{trend}=0.03) voor niet blootgestelden. Deze groep leek ook kleinere borsten te hebben met een hogere absolute en relatieve densiteit, maar die resultaten waren niet statistisch significant.

In hoofdstuk 5 rapporteren we over het risico op borstkanker dat is verbonden met absolute maten van borstdensiteit, in vergelijking tot het bekende risico dat is verbonden met percentuele densiteit. In een case-cohort design binnen het DOM onderzoek (Diagnostisch Onderzoek Mammacarcinoom, Utrecht) werden alle invasieve borstkanker gevallen in 15,5 jaar follow-up vastgesteld (361) en werd follow-up verzameld voor een willekeurige steekproef van 600 vrouwen uit het cohort. De mammogrammen van 309 cases en alle controles konden worden gevonden, gedigitaliseerd en beoordeeld. Het risico op borstkanker werd daarna vergeleken tussen tertielen van dense weefsel, niet-dense weefsel of relatieve densiteit en in gecombineerde categorieën van dense en niet-dense weefsel (laag/laag vs. laag/hoog, hoog/laag en hoog/hoog). We vonden dat het risico op borstkanker onafhankelijk leek toe te nemen met toenemende hoeveelheden dense weefsel (HR 1.68; 95%CI 1.22-2.46 voor hoogste vs. laagste tertiel) en niet-dense weefsel (HR 1.44; 95%CI 1.03-2.03 voor hoogste vs. laagste tertiel). Het hoogste risico werd gevonden voor vrouwen die bovengemiddelde hoeveelheden van beide weefsels hadden (HR 2.20; 95%CI 1.37-3.52) in vergelijking tot vrouwen met benedengemiddelde hoeveelheden van beide weefsels. Dit gaf aan dat niet-dense weefsel een rol speelt in de ontwikkeling van borstkanker die tot nu toe weinig aandacht heeft gekregen. Met name het resultaat dat het hebben van grote hoeveelheden van beide weefsels het hoogste geschatte risico gaf is interessant aangezien tot dusverre het grootste risico altijd werd toegeschreven aan de hoogste percentuele densiteit.

In hoofdstuk 6 evalueren we onze resultaten van de vergelijkingen van absolute en relatieve maten van (niet-)densiteit. Deze resultaten geven aan dat absolute maten van densiteit zonder twijfel een waardevolle bron zijn in etiologisch onderzoek, aangezien ze andere informatie opleveren dan de meer geaccepteerde relatieve maat van densiteit. Ook de rol van niet-densiteit zou meer aandacht moeten krijgen. Toch hebben de huidig beschikbare uitkomstmaten ook hun beperkingen en zijn ze niet de definitieve, ideale maat voor borstdensiteit. Volumetrische en longitudinale informatie zou additionele, relevante gegevens en verdere mogelijkheden voor nieuwe inzichten in borstdensiteit geven.

Dankwoord

Voor ik een aantal mensen persoonlijk bedank, moet ik iedereen die ik ken bedanken. Bedanken voor hun geduld. Iedereen die me ook maar een beetje kent, weet hoe dit proefschrift tot stand is gekomen en hoe lang dat heeft geduurd. Toch is er nooit ook maar één iemand geweest die me heeft voorgesteld om er maar mee te stoppen. Stoppen zit ook niet in mijn aard, maar toch is het impliciete vertrouwen op een goed einde dat me altijd is gegund mij het meeste waard.

Verder wil ik een algemeen bedankje uitdoen naar alle vrouwen die in de loop der jaren mee hebben gedaan aan het DOM onderzoek. Zonder jullie zou ik geen data hebben gehad om mee te werken. Speciale dank ook voor de Vriendinnen van het DOM project die dit onderzoek hebben gesteund.

Bedankt Paul. Jij was het die mijn wetenschapshart kloppend hield en me bij het Julius Centrum binnenhaalde. Ik had je zelf gevonden in mijn zoektocht naar een 'droge' stage omdat het lab niet mijn plek was, maar waar de wereld van klinische studies aan het eind van mijn stage aan me trok, deden jouw enthousiasme en enorme creativiteit mijn onderzoekershart weer sneller kloppen. Ik ben je dan ook enorm dankbaar dat je me aanbood mijn stagestudie door te zetten als promotieonderzoek en dat we tot een oplossing konden komen die mijn interesses in onderzoek en klinische studies kon combineren. Het was een genot elke keer weer met je te discussiëren en speculeren over nieuwe hersenspinsels. Het is jammer dat je door je gezondheid de afrondingsfase van dit proefschrift aan je voorbij hebt moeten laten gaan, maar jouw belangrijkste bijdrage heb je geleverd door ervoor te zorgen dat het er überhaupt is gekomen.

Bedankt Petra. In de woelige wateren van Paul's creativiteit, was jij het baken die me de koers deed houden. Je bijdragen waren onmisbaar om ervoor te zorgen dat uiteindelijk één idee werd gekozen en tot een goed einde werd gebracht. Alhoewel je een paar keer stevig op de rem hebt moeten trappen en ook geregeld een duwtje hebt moeten geven, heb ik onze samenwerking altijd heel erg fijn gevonden. Ik waardeer je bijdrage enorm en denk dat je recente aanstelling als professor bij het Imperial College en daarmee als 2[e] promotor dan ook niet minder dan vanzelfsprekend is.

Bedankt Carla. Ik had geen beter iemand dan jij kunnen wensen om de oncologie-groep te komen versterken toen ik net begonnen was als promovendus. Alhoewel je in eerste instantie mijn co-promotorschap erbij nam naast je eigen onderzoekslijn, heb je me altijd je volle aandacht gegeven. Je hebt me talloze keren enorm geholpen met je uitvoerige en altijd up-to-date kennis. Je commentaren waren scherp en soms confronterend, maar altijd terecht en een enorme bijdrage. Helemaal bedankt dat ik, bij fysieke afwezigheid van Paul en Petra, tijdens de laatste maanden van het werk aan dit proefschrift altijd en onvoorwaardelijk bij je binnen kon vallen.

Bedankt Rick. Velen voor mij hebben zich verbaasd over de mate van detail waarin jij je weet te mengen in ieder van de vele onderzoeken onder je houde. Ook ik heb dat mogen ervaren en heb er alle respect voor. Meestal was je de manager van het Julius Centrum die na korte discussie de knopen doorhakte, maar tot zeker twee keer toe kwam je met dusdanig scherp inhoudelijk commentaar dat je mijn stuk, en daarmee voor even mijn wereld, volledig op zijn kop zette. Zowel mijn stukken als ik zijn er alleen maar beter van geworden. Waar ik je echter het meest voor wil bedanken is het vertrouwen dat je altijd in me hebt getoond toen de voortgang stagneerde en steeds meer tijd verstreek. Ik hoop dat het eindresultaat je vertrouwen waarmaakt.

Bedankt Ingeborg. Enkele malen hebben wij samengezeten als ik struikelde over een stukje statistiek en telkens heb je me op het juiste spoor geholpen. Ook op het laatst heb je als lid van de leescommissie nog zeer nuttige opmerkingen geleverd die het eindresultaat alleen maar beter maken.

Bedankt Sjoerd, Martijn, Franzel en Irene. Jullie kwamen na mij, maar hebben me allemaal ruimschoots ingehaald. Ik weet niet in hoeverre ik precies heb geholpen met jullie onderzoek, maar weet wel dat alle discussies met jullie en het werk dat jullie hebben geleverd mij enorm hebben geholpen.

Bedankt overige promovendi van het Julius Centrum. Jullie zijn met teveel om op te noemen, maar hebben allemaal via de JOB, de wandelgang of waar dan ook wel eens geholpen door op te treden als sparringpartner of kenniscentrum. Esther, jij in het bijzonder bedankt voor alle ruggespraak in de laatste maanden. Succes over een paar weken.

Bedankt Annette. Als mijn bazin bij de CTSU heb je maar wat met me te stellen gehad als ik weer eens kwam vragen om verlof om aan mijn proefschrift te werken. Uiteindelijk kwam het zelfs in mijn functioneringsverslag, terwijl het niet eens mijn werk (meer) was en al helemaal niet binnen de CTSU. Bedankt voor de steun en de herhaaldelijke motivatie. Nu dit boekje klaar is heb je eindelijk mijn volle aandacht.

Bedankt Mariska. Door de jaren heb ik de meeste tijd op het Julius met jou op een kamer doorgebracht. Soms had ik het idee dat je me dan expres precies op de drukste momenten naar mijn promotie vroeg om me op de kast te jagen, maar over het algemeen was je volgens mij echt betrokken. Wat het ook was, je was altijd mijn praatpaal en daarmee heb je een onschatbare bijdrage geleverd.

Bedankt CTSU collega's. Geen week ging er voorbij zonder dat één van jullie me naar mijn proefschrift vroeg. Soms werden er hele lunchpauzes aan besteed. Al die gesprekken hielden me scherp en gemotiveerd om verder te gaan.

Bedankt Heren van VV. Onvoorwaardelijke vriendschap. Welgemeende interesse. Sociale dwang. Ontspanning. Peer review. Afleiding. Dat en veel meer vond ik bij jullie. Als mede medisch biologen/biomedisch wetenschappers kon ik altijd wel bij één van jullie terecht voor inhoudelijke of principiële discussie. Als vrienden kon ik altijd bij jullie terecht voor alles.

Bedankt (schoon)familie. De interesse die jullie altijd weer toonden in mijn werk en de uitleg die ik dan geregeld moest geven hebben niet alleen enorm geholpen om me door te laten gaan, ze hielpen me ook om mijn werk te focussen en hebben me zelfs een paar keer op een idee gebracht. Opa Piet, de trots in je ogen toen je op je sterfbed nog de kracht vond om je schoon-kleinzoon te feliciteren met zijn aanstaande promotie zal ik nooit vergeten.

Bedankt pa en ma. Jullie hebben dit mogelijk gemaakt. Al zo lang als ik me kan herinneren hebben jullie alles gedaan om mij de ruimte en mogelijkheden te geven om mezelf te ontwikkelen. Ik hoop dat het me lukt hetzelfde te doen voor mijn kinderen. Ik ben wie ik ben door jullie en ik heb kunnen doen wat ik heb gedaan door jullie.

Bedankt Karlijn. Jij bent in mijn leven gekomen tijdens het laatste wat ik heb gedaan vóór ik begon aan dit proefschrift, maar blijft er hopelijk in tot lang na de afronding. Door alles heen ben je er geweest. Twee keer als 'de-vrouw-met-het-slechte-gevoel-van-de-wereld-voor-timing-om-te-bevallen', maar daarnaast altijd als mijn rots in de branding. Bedankt dat je er was voor me. Ik weet wat jij hebt moeten verduren met me en zonder jou zou dit proefschrift er niet zijn geweest.

Bedankt Jeroen en Nienke. Bedankt dat jullie er zijn.

Curriculum Vitae

Gerco Haars was born in Lopik, The Netherlands on January 13th, 1976. Following primary school in Lopik, he attended VWO ("Preparatory Scientific Education") at the Willem de Zwijger college in Schoonhoven from 1988 to 1994. In the summer of 1994 he graduated VWO and started his studies of Medical Biology at Utrecht University. As part of his studies he did internships at the department of Infectious Diseases and at the Julius Center for Health Sciences and Primary Care, both of the University Medical Center Utrecht. During these years he was also an active member of the student faculty on both social and educational level, culminating in chairmanship of the Medical Biologists Society 'Mebiose' in 1996-1997. After graduating Medical Biology in the summer of 2000, he extended his work at the Julius Center with a position that combined research as a PhD student at the oncology group and work in clinical trials as a Clinical Research Associate at the Julius Clinical Trials Services Unit (CTSU). He held this combined position until October 2004 at which time he moved into a full-time position as Quality Assurance Manager of the CTSU. Since then he has continued his work on this thesis outside of working hours and has expanded his position of QA Manager to currently also include a role as QA Manager for the Bureau for Quality Assurance in Research of the UMC Utrecht and a role as Lead QA manager for the Academic Alliance for Clinical Trials (AACT), an international collaboration of the Julius Center.