Mammographic density – a useful biomarker for breast cancer risk in epidemiologic studies

Giske Ursin and Samera A. Qureshi

Department of Nutrition, University of Oslo

ABSTRACT

We describe mammographic density and how it is associated with breast cancer risk, what mammographic density represents biologically, as well as evidence that it is associated with breast cancer risk factors and is modifiable. Mammographic density has a large unused potential in mammographic screening programs. Currently mammographic density is being used as a biomarker or surrogate endpoint for breast cancer risk in a number of studies, and we discuss the rationale for doing this, as well as the challenges involved. A major challenge is the need for an automated method that can yield an even more precise estimate of the dense areas in the breast. Currently the most widely used methods are various computer-assisted methods. These are reader intensive, but so far the methods that yield the highest estimates for breast cancer risk. Once a robust automated method for assessing mammographic density or breast density is developed, this measure will probably become even more widely used, not just in epidemiology, but also in screening programs and in clinical practice.

INTRODUCTION – DEFINITION OF MAMMOGRAPHIC DENSITY

The relative amounts of fat, connective tissue, and epithelial tissue determines the radiographic appearance of the breast on a mammogram (Figure 1). Fat appears as dark or radiological lucent areas, whereas connective and epithelial tissue appear as areas of high radiologic density. Mammographic density represents the radiodense area. This is usually expressed as a percentage, where percent mammographic density is the percent of the breast area observed on a mammogram that is radiodense or white. Sometimes investigators will use the terms ‘breast density’ or ‘mammographic breast density’ to indicate mammographic density. However, the term ‘mammographic breast density’ seems to be redundant, and the term ‘breast density’ is not completely accurate, since it implies that this is a clinical, rather than a radiological measure. For this review we will use mammographic density to indicate what we can measure on a mammogram.

MAMMOGRAPHIC DENSITY, BREAST CANCER RISK AND MAMMOGRAM SENSITIVITY

Percent mammographic density has been found to be one of the strongest independent predictors of breast cancer risk (1-5), with risk increasing with increasing density. Women with the mammographically densest breasts have a 4-6 fold increased risk of breast cancer compared to women with the least dense breasts (5-12). Most women have some mammographic density, and the relative risk increases almost linearly with increasing density. It is estimated that 10% of postmenopausal women and 20% of premenopausal women

Figure 1. Mammographic density is the area or areas on a mammogram that are white (radiodense). This represents epithelial and connective tissue.
have mammographic density above 50% (5). This is therefore a common risk factor, and it has been estimated that about a third of all breast cancer cases can be explained by high mammographic density. Further, it has been argued that individual risk prediction models with mammographic density alone is as strong a predictor as the Gail model (13), which is sometimes used to identify women at high risk of breast cancer.

The sensitivity of a mammogram, or the ability of detecting an existing cancer is also reduced in women with high mammographic density (14,15). Data from various screening programs suggest that interval cancers are more prevalent in women with mammographically dense breasts (16,17). Despite this, mammographic density is in general not used to guide screening intervals, or even as a criterion for additional exams in large screening programs. Radiologists today use previous mammograms to compare with the current one for changes that could indicate the onset of a cancer. Although this improves detection rates for cancer, screening programs could probably improve their effectiveness even more by including mammographic density as a criterion for selecting women who need additional exams (17). Although additional exams such as ultrasound or magnetic resonance imaging are time consuming and costly, as they must be performed by trained radiologists, it is clear that they can improve detection rates of cancer substantially in women with dense breasts. However, today, women are often not told whether they have mammographically dense breasts, or how sensitive the mammogram is likely to be for them. To what extent this is acceptable from a clinical or even ethical point of view can be discussed. However, for epidemiologists there is a clear advantage when few women know how dense their mammograms are. This obliterates much concern about selection bias in designing studies of mammographic density as the outcome.

**WHAT DOES MAMMOGRAPHIC DENSITY REPRESENT BIOLOGICALLY?**

Although mammographic density is a clear risk factor for breast cancer, in order for it to represent a useful biomarker for breast cancer it also needs to have some biological correlates that can explain why this measure is important. There have been a number of studies correlating histopathological findings to mammograms, but it is not yet completely clear what mammographic density represents biologically. Nor is the biologic basis of the relationship between increased mammographic density and breast cancer risk completely understood. A number of early studies reported that mammographically dense breasts contained epithelial hyperplasia (18-21), but this was not consistently found (22-24). Further, there is no evidence that epithelial proliferation is higher in dense than non-dense areas (25,26). There is, however, some data that dense areas have an increased number of epithelial cells (25).

Further, it has become clear that stromal fibrosis is a prominent feature in mammographically dense breasts (2,22,24), and that dense breasts have higher level of collagen, and altered expression of stromal proteins (27,28). Although the role of collagen and stroma in causing cancer of epithelial cells have not yet been completely elucidated, stromal-epithelial interactions are known to be important in breast carcinogenesis (29). Thus, although, the exact mechanisms are not clear, mammographic density is associated with certain markers of epithelial growth, and most definitely with breast stroma. Finally, when we studied a set of in situ tumors, the in situ breast cancers were more likely to occur in the areas that were mammographically dense (30). Thus although the details are not completely clear, mammographic density has a biologic basis that explains its role in breast cancer development.

**EPIDEMIOLOGICAL DETERMINANTS OF MAMMOGRAPHIC DENSITY**

In order for mammographic density to be useful as a biomarker for breast cancer, we would also expect it to be modifiable. Mammographic density appears to have both a genetic component, but also a modifiable, non-genetic component.

**Genetic component**

Mammographic density has a strong genetic component. Studies of twins suggest that a large percent of the variance is due to genetic factors (31-33). A number of epidemiologic studies have tried to identify the important genes using a candidate gene approach. A recent review (34) suggests that this approach has only had limited success so far. So far, there is little evidence that genes known to be strong determinants of breast cancer risk predict mammographic density. Similarly, common genetic variants identified in genome wide association studies to play a modest role in breast cancer risk have not been strongly associated with mammographic density (35,36). However, there is some indication that some of the genes involved in hormone metabolism or that the insulin growth factor genes (37) play a role. A number of studies are under way to further explore the genetic basis of mammographic density, and more results on this topic should emerge over the next few years.

**The modifiable (non-genetic) component**

**Body mass index and reproductive factors**

The environmental or non-genetic risk factors for mammographic density have been much studied (1,2,4,38,39). Mammographic density has some similarities with serum estrogen levels in that it declines with age and with menopausal status. However, while estrogen postmenopausally is positively associated with body mass (BMI), the association between percent mammographic density and BMI is inverse. The reason for this
is obvious, women with large BMI tend to have large breasts with substantial amount of fatty (non-dense) tissue. Of other breast cancer risk factors, mammographic density is strongly inversely associated with parity, this effect is almost linear. Large studies have found that mammographic density, as breast cancer risk, increases with age at first birth (38,39). There is some indication that mammographic density is higher in women with early menarche (38), but these data are not completely consistent (39,40). It has been suggested that the genetic component that determine mammographic density may not be that different from the genetic components that explain breast cancer risk factors (41).

Postmenopausal hormone use
Mammographic density is clearly associated with use of postmenopausal hormone therapy regimens with combined estrogen and progesterin therapy (42). Numerous studies have reported mammographic density changes in women who start combined estrogen and progesterin therapy (EPT), most have been from the US (43-48). Two placebo-controlled randomized trials from the US, the Postmenopausal Estrogen and Progestin Interventions (PEPI) trial (46,48), and the Women’s Health Initiative (WHI) trial (49) found that women assigned to the EPT arm had on average a 5% and 6% increase in mammographic density respectively after 1 year, while there were only minor changes in the placebo group or the estrogen alone arm. In both studies women used conjugated equine estrogens combined with medroxyprogesterone acetate, and in PEPI there was also one arm combining these estrogens with micronized progesterone. There is a large individual variation in how these treatments affect mammographic density. Part of this variation is explained by changes in estrogen levels (50,51), suggesting that how women absorb or metabolize estrogen may determine this variation.

There are limited data on mammographic density changes associated with the EPT regimens commonly used in Scandinavia, which contain estradiol (E2) and norethisterone acetate (NETA) compounds. Two Swedish (52,53) and two Greek studies (54,55) correlated data on mammographic density changes with such hormone use. These studies used Wolfe parenchymal patterns categories to classify mammograms, and found that higher risk patterns were substantially more common in women starting EPT. Two Norwegian studies reported similarly higher mammographic density among women using the E2/NETA regimens using Madena (56,57). There was no indication that the E2/NETA regimens are better for the breasts than the US regimens, or that the mammographic density changes observed with E2/NETA regimens are smaller than those observed with US regimens.

Other medications – tibolone, tamoxifen and raloxifene
Although the effects of tibolone on the breast are not completely clear, evidence so far suggest that it does not increase mammographic density (58,59). Tamoxifen treatment reduces mammographic density, at least in premenopausal women (60,61), and treatment with a hormone regimen that reduces serum levels of estrogen and progesterone also reduces mammographic density (62,63). Raloxifene, which is used to prevent osteoporosis, but which has also been found to prevent breast cancer (64,65), does not have much effect on mammographic density (66,67).

Thus the evidence so far suggests that mammographic density is associated with a number of environmental factors, in particular certain hormone therapies, as well as reproductive factors believed to act through hormonal mechanisms. How large a percentage of the variance in mammographic density is explained by non-genetic factors is not completely clear. Some will argue that it is less than 30%, others that it may be close to 50%, the discrepancy is due to what extent one believes the variance in mammographic density between monozygotic twins is solely due to shared genetics or could partially be due to shared environment (31-33).

The case for using mammographic density as a surrogate marker for breast cancer risk
The associations with hormonal factors suggest that mammographic density is modifiable. This, combined with mammographic density being so closely associated with breast cancer risk, is why it has been suggested that mammographic density be used as an intermediate endpoint in breast cancer intervention studies. An advantage of using mammographic density rather than cancer is that mammographic density is a quantitative trait that all women have, while very few women develop breast cancer. Some investigators have, however, argued that until it is demonstrated that a mammographic density increase results in cancer occurrence, the use of this marker is not interesting. However, this question is currently being addressed in a study within the Women’s Health Initiative trial, and should become available over the next year. Data from studies of mammographic density changes over time, do however, suggest that density increases are in fact predictive of risk (68).

What magnitude of mammographic density change is important?
If an intervention or risk factor changes mammographic density with on average 5%, is this important? It could be. Estimates of density changes are averages, which means that a subset of the women may experience substantially larger changes. For estrogen and progesterin therapy, the average change is 5-6%, but a subset of women have much higher changes, 20-25% have increases of 10% or more, and some women have a substantially larger increase (50). Similar magnitude
changes are seen with tamoxifen. The important clinical question is whether the women with the largest changes in density with an intervention are the women with the largest changes in breast cancer risk.

**HOW TO MEASURE MAMMOGRAPHIC DENSITY**

**Qualitative methods**

There are numerous methods of measuring mammographic density. Early studies used predominantly parenchymal patterns. The most commonly used such classification was developed by John Wolfe (69,70), a well known mammogram expert in the US. Wolfe described four parenchymal patterns (N1, P1, P2, DY) of increasing densities. In the N1 pattern, the breast consists almost entirely of fat, the P1 and P2 patterns represent increasing ductal prominence, and in the DY pattern the breast parenchyma consists of diffuse or extensive nodular densities. In his two original studies, Wolfe reported that the risk of incident breast cancer was substantially higher in women with the DY pattern than in women with the N1 pattern (69,70). Although later studies confirmed a higher risk of breast cancer in women with the DY/P2 high-risk patterns (1), results were not as impressive as in Wolfe’s first study. Other classification methods have some similarities with Wolfe patterns. The qualitative Breast Imaging Reporting and Data System (BI-RADS) method for density assessment developed by the American College of Radiology is one commonly used approach (71). Note that this BI-RADS density method is not the same as the clinical assessment categories that were created to indicate whether a mammogram represents a negative, benign or suspected malignant finding. Rather the BI-RADS mammographic density categories are four, originally qualitative, categories of density (almost entirely fat, scattered fibroglandular densities, heterogeneously dense and extremely dense) (71). Another set of patterns are those developed by a Swedish mammographer, Lazlo Tabar (72). These qualitative methods have been associated with breast cancer risk and breast cancer risk factors, but the magnitude of these associations are not as strong as those obtained with more quantitative approaches (5,8,73). It has further been suggested that qualitative patterns are not predictive of breast cancer risk after percent density has been taken into account (8,74).

**Quantitative methods**

There are a number of quantitative approaches. The simplest is the subjective evaluation approach, where radiologists categorize the mammograms into one of a number of preset categories, such as <25% density, 25-49% density etc. One such method is the six category subjective assessment method used by Boyd (75). Another method is the quantitative BI-RADS method: the qualitative BI-RADS categories described above have recently been linked to a quantitative description (<25%, 25-50%,51-75% and >75% density). This means that a vast number of mammograms read in the US are read with these categories (71). The Norwegian Breast Cancer Screening Program uses three categories of density, <30% glandular tissue, 30-70% glandular tissue and >70% glandular tissue (16). The choice of this few categories was unfortunate, and future screening programs would be better off using the BI-RADS 4-category approach, as this would at least enable comparisons with the vast amount of data collected in the US.

Computer-assisted methods have now become the most widespread method for assessing mammographic density in epidemiological studies. The method entails that the reader uses digitized versions of the analog image, and then using a specially developed software package, where the reader can outline the total area of the breast, as well as the area he/she considers to represent mammographic density. The dense area is identified using a threshold method, where the reader sets the threshold of ‘whiteness’ for what represents mammographic density after first excluding light artifacts. There are currently several such methods, including the Toronto method (Cumulus) (76,77), ours (Madena) (63), as well as others (11). The Madena method is displayed in Figure 2. Different amount of mammographic density is displayed in Figure 3. These computerized threshold methods have been well validated in the sense that they all have resulted in strong estimates of relative risk of breast cancer.

**Digital mammograms**

Digital images appear less dense, thus comparisons between analog and digital mammograms from the same woman over time can be problematic. Further, the methods described above for assessing mammographic density were developed for use of analog mammograms that are subsequently scanned into a computer. Few studies have examined to what extent these methods yield the same risk estimates when applied to digital mammograms. However, several of the automated methods and volumetric methods described below can use digital images, although as explained below, they have not yet become fully established, nor have they yielded as strong associations as the current methods.

**Automated methods**

A number of automated or semi-automated methods have been proposed to identify mammographic density using either a threshold based method, such as those described above, or fractal analysis or other texture-based techniques (78-86). However, so far none of these methods have become widely used.

**Volumetric methods**

Mammographic density as measured with the methods described above has been much used in epidemiologic studies. However, percent mammographic density is a
**Figure 2.** Example of a computerized mammographic density assessment program (Madena) (http://radon.usc.edu/uscradonc/madena/madena.html or http://www.eyephysics.com/Madena/TOC.html).

**Top**) On the digitized mammogram that has been imported into Madena, the reader outlines the total breast area by drawing a blue line around the breast. The size of this area is calculated by the software (here: 147.20 cm²).

**Middle**) The reader draws a region of interest in red around the areas in the breast considered to contain mammographic density.

**Bottom**) The reader decides on a threshold for what represents mammographically dense areas within the region of interest. Such dense areas are colored yellow. The size of the yellow area within the region of interest is estimated by the computer (here 58.68 cm²). Percent density can later then be calculated (100% x 58.68 cm² / 147.20 cm² = 39.9%).

simplified, two dimensional measure of a three dimensional structure, and introduces substantial measurement error of the actual biologic measure of interest, epithelial tissue (or epithelial-stromal tissue) in the breast. Volumetric measures of the dense tissue in the breast ought therefore to yield even higher estimates of breast cancer risk than mammographic density. Currently a number of research groups are working on developing automated volumetric methods to yield an estimate of breast density (either based on mammograms or other radiologic techniques). These include methods based on ultrasound tomography (87) and magnetic resonance imaging (88). In addition some investigators have developed methods that use digitized film mammograms (89-92). However, so far these automated methods have yielded weaker associations with breast cancer risk and with risk factors than the standard two dimensional mammographic density methods (93-95).

**CHALLENGES WITH USING MAMMOGRAPHIC DENSITY AS A SURROGATE MARKER FOR BREAST CANCER RISK IN EPIDEMIOLOGIC STUDIES**

Although it has yet to be proven that a change in density is associated with a change in breast cancer risk, mammographic density has already been used in a number of studies as a surrogate marker for breast cancer risk. In the following we discuss some of the challenges associated with such use.

**Not all interventions work – the example of physical activity**

What we do know so far is that mammographic density does respond to hormone manipulations. However, this does not mean that it is useful for studies of every possible intervention for breast cancer. One example is physical activity. The association between physical activity and mammographic density is not straightforward. Although physical activity is a protective factor
for breast cancer, there is little evidence from epidemiological studies so far is that it is associated with reduced mammographic density (96-99). To the contrary, women with high levels of activity tend to have high percent mammographic density. The association with absolute area of the mammogram that is dense (absolute density) is also not that clear (99). Perhaps this suggests that it is difficult to distinguish the effects of physical activity completely from that of body mass (see above). What is clear is that mammographic density is not a useful marker for the beneficial effects of physical activity on the breast.

Current computer-assisted methods – reader intensive and time consuming

One main challenge with the current computer-assisted methods is that they require digitized mammograms, and this is time consuming and expensive. The other challenge is that the methods are not objective, but rather completely dependent on a subjective assessment by the reader. The measurements are also time consuming to obtain, each digitized image needs to be pulled up on the screen and read. An experienced reader can read anywhere between 30-100 mammograms per hour, fewer if this is part of a clinical trial and the images need to be compared.

Subjective measure, depends on reader

Another challenge with these current computer-assisted methods is that they are indeed subjective, i.e. reader dependent. Although reading densities is not that difficult to learn, negative findings in particular from small studies should be interpreted with caution. It ought to be a requirement that negative studies should provide evidence that the reader’s readings are valid. Usually readers will describe high correlation coefficients or high intra-class correlation coefficients. However, high correlations is expected on a variable with values from 0-100, where we use essentially the whole scale. Further, showing that a measure is reproducible does not necessarily indicate that it is valid. What investigators ought to do instead (or in addition to these measures) is to provide results on how their mammographic density estimates vary with age, or parity, or menopausal status or even BMI. If they cannot find associations with these variables in the expected direction, then there is little reason to believe that the measurements of mammographic density used in the paper are valid. Similar requirements ought to be placed on studies using new automated methods, both data on reliability and validity should be presented.

Measurement error – technical challenges, changes in projection of mammograms

It is difficult to assess changes in mammographic density if the films at two different time points have widely different exposures or, and this is more common, if the projection of the breast has changed. Sometimes one image will tend to display much more of the proximal area of the breast than the image obtained at the other time point, making any comparison impossible. If one image is analog and the others digital, the reader will guess that the analog image is older than the digital images, introducing possible systematic bias. At some large facilities, in particular in the United States, equipment, films and even technicians may change often. This is an additional challenge. However, all of these issues can be overcome in studies of mammographic density with adequate planning, size and making certain the mammograms are read in a random order, and that the reader is blinded to the treatment arm and timing of the images.

Automated volumetric methods of breast density – what to expect

Once a robust automatic volumetric method is developed, we should expect it to yield even stronger estimates of breast cancer risk than the current methods.
MAMMOGRAPHIC DENSITY AS A BIOMARKER

using mammographic density. However, any new such method needs to be able to show that it can find risk associations with breast cancer that are at least as strong as those with the conventional mammograms and computer-assisted methods. Thus, unless such methods can find at least relative risk increases of 4-6, they are not particularly useful. And, because we would expect volumetric methods to reduce the measurement error we are introducing by using a two-dimensional image when we use mammograms, we should expect solid volumetric methods to yield relative risk increases that are substantially larger than 6. Thus showing that a new volumetric method is highly reproducible, or correlated with, or as good as current computer-assisted methods of assessing percent mammographic density is not sufficient, the volumetric methods ought to be even better.

CONCLUSIONS AND FUTURE PERSPECTIVES
In conclusion, mammographic density is a strong breast cancer risk factor, one of the strongest risk factors known, apart from age and certain genetic mutations. It has been associated with other breast cancer risk factors, in particular those believed to act through hormonal mechanisms. Another advantage with this marker that can be measured on a continuous scale is that all women have measurable density, and most women have at least some density. Mammographic density may be a useful surrogate endpoint for breast cancer risk in clinical trials of agents that work through hormonal mechanisms. But, not all interventions may work on mammographic density, even if they ultimately turn out to reduce breast cancer risk. Therefore, studies selecting to use this measure must keep in mind how their intervention is likely to work. Probably the greatest challenge to mammographic density is that it is a two dimensional method, and there are still no automatic methods that have been found to work as well or better than the computer-assisted methods. Thus once a robust automatic volumetric method for mammographic density has been developed, and estimates are immediately provided to clinicians, then mammographic density may become much more widely used both in mammographic screening programs as well as in clinical practice. Until then, this is mostly a measure for epidemiologists.

REFERENCES


