Statistical analysis of mammographic breast composition measurements: towards a quantitative measure of relative breast cancer risk

C J KOTRE, PhD

Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne, UK

Objective: A number of studies have identified the relationship between the visual appearance of high breast density at mammography and an increased risk of breast cancer. With the advent of digital mammography and the promise of routine measurements of parameters associated with breast composition, the possibility arises of using breast composition in a quantitative manner to predict relative breast cancer risk. Previous measurements have shown that the average proportion of glandular and adipose tissue within the breast varies with both age and breast size. In order to be able to identify individual women with an unusually high volume of glandular tissue, it will therefore be necessary to make comparisons with a disease-free population matched for age and breast size.

Methods: A large number of breast glandular thickness measurements were analysed to investigate the statistics of breast composition across a disease-free population as a test of a suitable methodology for relative risk estimation. The large data set is also used to revisit the trends in breast composition used in the current UK method of breast radiation dosimetry.

Results: It is demonstrated that a non-linear transformation can be used to produce normal statistical distributions, suitable for producing a standardised “Z-score” for breast composition.

Conclusion: A standard “Z-score” approach to identify women with unusually glandular breasts is recommended and so provide a basis for cancer risk estimations.

Received 27 October 2009

Revised 6 January 2010

Accepted 4 February 2010

DOI: 10.1259/bjr/40806022

© 2011 The British Institute of
Radiology

It is generally accepted that a link exists between so-called “breast density”, *i.e.* the proportion and pattern of bright areas in a mammographic image denoting glandular tissue, and the relative risk of breast cancer. The field has been reviewed by Harvey and Bovbjerg [1], who identified 12 studies in which quantitative methods of measuring breast density showed a moderate to strong positive association with breast cancer risk. The risk of breast cancer for women with increased breast density in most of these studies is four to six times that for women with primarily adipose breasts; a relative risk greater than most traditional risk factors such as nulliparity and early menarche. With the advent of digital mammography and the promise of replacing visual assessment of breast density with quantitative measurements of parameters associated with breast composition, the possibility arises of using breast composition to predict relative breast cancer risk. In this article, a large number of breast glandular thickness measurements are analysed to investigate the statistics of breast composition across a disease-free population as a test of a suitable methodology for relative risk estimation.

Despite the structural and functional complexity of the female breast, in terms of its mammographic appearance, and for the purposes of constructing a simple cancer risk model, it is convenient to regard the breast as

being composed of two tissue components, adipose and glandular. These two components can be defined in terms of their X-ray attenuation properties [2], and phantom materials mimicking their physical characteristics at the X-ray energies commonly employed in mammography are commercially available (CIRS, Norfolk, VA). Any method of calibration of mammographic radiographic units which aims to measure breast composition that employs these phantom materials will therefore be contributing to a convenient standardisation of the definition of the adipose and glandular components. A number of such methods has been described, and the quantities proposed as the risk-related measure have covered a considerable range, with some form of percentage glandularity a frequent suggestion for the risk-related variable [3–6].

What is being described as glandular tissue here is in fact a combination of fibroglandular and connective tissue in which the proportion of luminal epithelium, from which most cancers arise, is known to be quite small [7]. However, if the basic assumption is made that an increased relative risk of breast cancer is associated with an unusually large number of glandular cells in the individual breast, that the number of cells at significant risk of carcinogenesis will be proportional to the volume of fibroglandular and connective tissue and that the relative risk associated with the volume of adipose tissue present is negligible, then it would seem reasonable to aim at quantifying the breast composition in terms of the volume of fibroglandular and connective tissue present [8]. With the advent of full-field digital mammography, such

Address correspondence to: Dr John Kotre, Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne, Tyne and Wear NE7 7DN, UK. E-mail: john.kotre@nuth.nhs.uk

volumetric measurements are possible, although there is likely to be a number of competing approaches to the problem in the medium term. Even with the promise of reliable delivery of an estimate of the volume of glandular tissue with each digital mammogram in the future, the question still arises as to the best way of employing such information to measure relative risk of breast cancer and make decisions on patient management.

Methods and materials

Estimation and analysis of glandular thickness

A method for calibrating modern film-screen mammography units to give an estimate of glandular tissue thickness overlying the automatic exposure control (AEC) sensor has been described previously [9]. This measure of breast composition is clearly inferior to the full volumetric measurement of glandular tissue advocated above, but it does provide an intermediate step between the visual grading of breast density used in much of the existing literature [1] and true volume estimation. The method has the big advantage of being able to be applied retrospectively to the many thousands of existing mammograms on which the required exposure parameters have been routinely recorded. In this study over 11 000 mammographic exposures routinely undertaken in the NHS Breast Screening Programme (NHSBSP) on three calibrated radiographic units are analysed to determine trends with age and compressed breast thickness, and to examine the variation in statistical distributions across these variables.

Data collection

Three Lorad MIV mammography units (Hologic Inc., Bedford, NA) on mobile trailers were calibrated to allow estimation of glandular thickness by X-ray absorptiometry as previously described [9]. Data collection took place on women routinely attending for breast screening mammography in late 2008 and early 2009. The exposure parameters required for glandular thickness estimation were post-exposure mAs, indicated compressed breast thickness, compression force, AEC position, kVp, X-ray tube target and filter material. The gantry angle was also recorded so that lateral oblique (OBL) and craniocaudal (CC) view data could be analysed separately. Client identification numbers were removed from the data set before analysis, but age at the time of screening was retained. All of the required data can be recorded on disk by the Lorad MIV unit as well as being flashed on to the film, and collection on floppy disks allowed some 13 000 sets of data to be recorded relatively conveniently. No selection of clients was made to eliminate cases subsequently diagnosed to be cancer, but as this would be expected to be at a rate of around 0.8% [10] this would not be expected to significantly alter the statistics.

Before analysis proper, the data were examined to eliminate incomplete data sets (most frequently caused by the machine failing to register the compression force) and sets where no age was recorded. The glandular thickness estimation is known to be particularly sensitive

to errors in compressed breast thickness measurement, and a correction based on compression force and AEC position was applied to reduce this error as far as possible [9]. This source of error could not be eliminated, however, and where it resulted in a glandular thickness estimation being less than zero, or greater than the compressed breast thickness, these data sets were also rejected, leaving a total of 11 015 for further analysis.

Data grouping and transformation

Previous studies have shown that the average proportion of glandular and adipose tissue within the breast varies with both age and breast size, where breast size is defined in this case as the compressed breast thickness recorded during mammography. In order to be able to identify women with an unusually high volume of glandular tissue, it is therefore necessary to make comparisons with the statistics of a disease-free population with the same age and breast size. The data set was divided into small age and thickness ranges and analysed so that the trends in these statistics as functions of both age and breast thickness could be derived.

Age variations are due to the hormonal changes associated with the menopause. The median age of the menopause at 51 years in the UK [11] lies within the current NHSBSP invited screening age range of 50–70 (to be extended to 47–73), so the rate of change of breast composition from this source is relatively high in the screening population [12]. The average proportion of glandular tissue is also known to be smaller in larger breasts [12], and this finding is incorporated in the present definition of standard breast used for dosimetry in the NHSBSP [13]. This size relationship will also be examined in this study using the larger number of measurement points available.

The 11 015 valid glandular thickness data points were first separated into OBL and CC views. Although it might be expected that similar trends would be found for both views, it was felt important not to act on this assumption until it could be proven. The sets of OBL and CC data were then subdivided into the 5-year age bands 49–53, 54–58, 59–63, 64–68, 69–73 and 74+, and each of these was further divided into 1 cm compressed breast thickness bands extending from 2 and 9 cm. The recorded compressed breast thickness was corrected to thickness over the AEC sensor as previously described [9]. The resulting groups of data varied in size, and some groups at the extremes of age and compressed breast thickness contained too few elements to analyse further. This occurred throughout the age range in the 8–9 cm thickness group; therefore, this was merged with the 7–8 cm group and the results are presented with a combined 7–9 cm group at the top of the thickness range.

Figure 1 shows the frequency distribution of glandular thickness in the well-populated 59–63 years and 5.00–5.99 cm thickness group for OBL (Figure 1a) and CC (Figure 1b) views. Both show a strongly skewed distribution. This shape of distribution was evident in all other groups analysed. This result immediately invalidates the direct use of statistics such as mean, standard deviation and standard error of the mean (SEM), and somewhat weakens previous studies that

Figure 1. The frequency distribution of glandular thickness in the 59–63 years and 5.00–5.99-cm thickness group for (a) lateral oblique and (b) craniocaudal views.

have used these statistical measures on smaller data sets. In order to proceed with analysis using the preferred conventional statistics, a reversible non-linear transformation to a normal distribution is required. A small amount of experimentation demonstrated that a cube-root transformation of the glandular thickness values resulted in satisfactorily symmetrical and normal shaped distributions. Figure 2 shows the transformed versions of the distributions shown in Figure 1. The symmetrical fractional errors in the measurement process are driven by uncertainties in the measured compressed breast thickness and variations in recorded mAs due to the differences between individual film-screen cassettes [9]. At the most frequent compressed breast thickness of 6 cm, the fractional error on an individual measurement is approximately 40%. This uncertainty is dominated by the larger variation in the measured glandular thicknesses and does not

significantly affect the symmetry of the transformed distributions of Figure 2.

The analysis of trends with age and compressed breast thickness proceeded on the cube-root transformed data. The finding that the glandular thickness distribution can be simply transformed to a normal one also offers the possibility of using the normalised deviation from the mean (“Z-score”) as a working measure of comparison against a disease-free population. This is further explored below.

Results

Trends in glandular thickness and implications for dosimetry

Figure 3a–f shows the trend in average glandular thickness with compressed breast thickness for the OBL (circles) and CC (squares) views in each of the age bands.

Figure 2. The frequency distribution of transformed (cube root) glandular thickness in the 59–63 age and 5.00–5.99 cm thickness group for (a) lateral oblique and (b) craniocaudal views.

Figure 3. (a–f) The trend in average glandular thickness with compressed breast thickness for the lateral oblique (○) and craniocaudal (□) views in each of the age ranges: (a) 49–53 years, (b) 54–58 years, (c) 59–63 years, (d) 64–68 years, (e) 69–73 years and (f) 74+ years. The error bars show ± 1 standard error of the mean.

The average was calculated as the mean of the cube-root transformed data, which is then inverse-transformed to a value in centimetres. The error bars show ± 1 SEM, and these are slightly asymmetrical owing to the inverse transformation. Groups containing five or fewer measurements are not plotted. The compressed breast thickness plotted is in each case the mean thickness for the group of results within the age and thickness band analysed. The OBL and CC results are similar and mostly within 1 SEM of each other. The values for the thinnest breasts in the youngest two age bands are higher for the CC views than the OBL views. Radiographers reported that it is sometimes more difficult to apply adequate compression in the CC view for very small breasts, so this may be a real effect, but still equivocal because of the relatively small numbers of data points in these groups.

At this stage it was decided that the most reliable results would be produced by pooling the CC and OBL data sets and recalculating all of the statistics. Figure 4 shows the combined results. Error bars are omitted to avoid cluttering the figure. The fitted curves are derived from a two-dimensional (2D) surface fit to all of the groups analysed in line with the approach followed below. Data points for groups containing five or fewer measurements are not marked, although the surface fit still allows a trend line to be drawn. The curves show the expected trends with breast size and age, and the trends are smooth, an important consideration when the objective is to allow interpolation in both age and compressed breast thickness.

With the large number of data available, and the consequently improved statistics, there is an opportunity to re-evaluate the trends in glandular thickness within the context of the definition of the standard breast currently used for dosimetry in the NHSBSP [13]. This is defined with an adipose shield of 5 mm at the beam entry and exit surfaces of a semi-circular breast phantom, and the glandularity is defined as the percentage thickness of glandular tissue within the parenchyma, which therefore

Figure 4. Curve-fitted trends in average glandular thickness from the lateral oblique and craniocaudal results combined for age ranges: (○) 49–53 years, (□) 54–58 years, (△) 59–63 years, (▽) 64–68 years, (◇) 69–73 years and (★) 74+ years.

has a thickness of 1 cm less than the recorded compressed breast thickness. Two percentage glandularity curves are currently defined, for the 40–49 and 50–64 years age groups. The latter of these is reproduced in Figure 5 (dotted line) together with the corresponding curve calculated from the glandular thickness data for the combined 49–53, 54–58 and 59–63 years age groups. A third-order polynomial fit is used, and the fit is interpolated back to 2 cm breast thickness for comparison.

The curves in Figure 5 are clearly different, owing not only to the additional data and revised method of calculating the average used here, but also to the convention of forcing a fit to 100% glandularity for the dosimetric model at 2 cm thickness. The results suggest that, not unexpectedly, the dosimetric convention of a 0.5 cm thickness of adipose tissue at the entrance and exit surfaces of the compressed breast is not realistic for the smallest breasts. The dosimetric consequences of this are, however, relatively minor. The method of estimation of mean glandular dose used in the NHSBSP [13] employs a composition correction or “c-factor” to correct for any difference in breast composition from a glandularity of 50%. The maximum difference between the curves in Figure 5 towards the middle of the thickness and tube filtration range, at say 5 cm compressed breast thickness and 0.35 mm Al half value layer (HVL), results in an increase in the c-factor, and hence mean glandular dose estimate of 5% if the glandularity estimate given here is used. The present dosimetric model does not define glandularity values above 100% [13], so the deviation at small breast thicknesses cannot be evaluated within the constraints of this model. If, however, it is postulated that glandularity values above 100% imply a spread of glandular tissue into the 0.5 cm layers presently defined as adipose, then these elevated values might provide at least an approximation to estimate the resulting change in mean glandular dose. Following this approach, if the established c-values are simply interpolated up to the 2 cm glandularity of 144% suggested by Figure 5, the c-factor and hence the mean

Figure 5. The 50–64 years age group percentage glandularity curve used in the present dosimetric definition [13] (dotted line) compared with the equivalent curve calculated from the combined 49–53, 54–58 and 59–63 years age groups analysed in this article (solid line, ○).

glandular dose estimate decreases by only 10%. It should be noted that other extrapolations to the glandularity curve are possible, but the one used would suggest that no urgent change in the established standard breast is indicated as a result of the findings given here.

Trends in variation of glandular thickness

In order to apply a statistical test of significance to the glandular thickness value for an individual woman, it is necessary to know the variation of the value in the disease-free population against which it will be compared, and how this variation changes with age and compressed breast thickness. Figure 6 shows the standard deviation of the cube-root transformed glandular thickness measurements for each of the six age groups plotted against the average compressed breast thickness for each group. The standard error on the standard deviation is also plotted. The conclusions that can be drawn from Figure 6 are that there are no clear trends in standard deviation with increasing age (with the ordering of age groups inconsistent at each thickness), and that the values cluster together in the mid-thickness range where the largest number of data points contribute. It is therefore hypothesised that, provided the standard deviation is calculated with reference to the age and thickness-matched transformed average, the resulting value is relatively age insensitive. On this basis, it was decided to produce weighted average values of standard deviation at weighted average values of compressed breast thickness. These are plotted in Figure 7 together with a second-order curve fit.

Quantifying deviation from the disease-free population

A parallel problem of identifying high-risk individuals using a measurement variable that varies in the disease-free population is found in osteoporosis assessment

Figure 6. The standard deviation of the cube-root transformed glandular thickness measurements for each of the six age ranges plotted against the average compressed breast thickness for each group: (○) 49–53 years, (□) 54–58 years, (△) 59–63 years, (▽) 64–68 years, (◇) 69–73 years and (★) 74+ years.

Figure 7. Weighted average values of standard deviation at weighted average values of compressed breast thickness, with a second-order curve fit.

using dual-energy X-ray absorptiometry. In this field, normalised deviation from the mean, or “Z-score”, is commonly used to measure how different an individual measurement is from the sex- and age-matched normal population. The Z-score is defined as:

$$Z = (x - \mu) / \sigma \quad (1)$$

where x is the measurement for the individual, μ is the mean and σ the standard deviation of the population. These latter statistics should be for the whole population rather than a sample, but in practice a large random sample is often used.

In order to use a similar approach to quantify breast composition, population data for the mean and standard deviation of a normally distributed breast composition variable are needed. As argued above, a simple second-order curve fit against compressed breast thickness seems appropriate to estimate standard deviation, but for (transformed) average glandular thickness a 2D surface fit against age and compressed breast thickness is required. A fourth-order surface was found to give an accurate overall result, with an root mean squared (RMS) error on the fit of $0.018 \text{ cm}^{1/3}$ and a maximum deviation of $0.04 \text{ cm}^{1/3}$.

The form of the 2D surface fit using fourth-order polynomials is:

$$w = a + bu + cv + du^2 + ev^2 + fu^3 + gu^2v + hu^2v^2 + iuv^2 + jv^3 + ku^4 + lu^3v + mu^2v^2 + nuv^3 + ov^4 \quad (2)$$

where u is the age, v is the compressed breast thickness in cm and w is the fitted variable, the cube root of glandular thickness in $\text{cm}^{1/3}$. The coefficients a–o for the surface fit are given in Table 1.

Table 1. Fit coefficients for the fourth-order surface fit to cube-root glandular thickness given in Equation 2

Fit coefficient	Value
a	9.51769
b	-5.46300×10^{-1}
c	1.62594
d	1.08825×10^{-2}
e	-1.32711×10^{-3}
f	-4.62296×10^{-1}
g	-8.76620×10^{-5}
h	-2.56483×10^{-4}
i	3.05540×10^{-3}
j	4.11779×10^{-2}
k	2.44955×10^{-7}
l	1.45987×10^{-6}
m	-2.53940×10^{-6}
n	-1.42924×10^{-4}
o	-1.26154×10^{-3}

The fitted surface plot is shown in Figure 8. Overall, the surface is smooth and suitable for the interpolation between measured thickness and age groups required to calculate Z-scores as suggested. The proposed general method for using the information from Figures 7 and 8 is shown diagrammatically in Figure 9.

Extrapolation to three dimensions

As indicated above, the quantity that might be expected to have the best correlation with relative cancer risk is comparison with the disease-free population of the total volume of glandular tissue within the breast (available in the future with digital mammography), rather than the average thickness of glandular tissue over the AEC sensor available as the source data here. A tentative extrapolation of the glandular thickness trends given above to three dimensions can be made, however, using a simple relationship between compressed breast thickness, average projected radius of the breast and thickness of the purely adipose layer previously established on more than 1000 OBL mammograms [14]. The data tabulated in this reference plot as a good fit to a straight line, and can be

Figure 8. The cube-root transformed average glandular thickness surface fitted to age and compressed breast thickness using Equation 2. **Figure 9.** Flow diagram of the general method proposed for identifying women with increased relative breast cancer risk using breast composition measurements. OBL, lateral oblique; CC, craniocaudal.

used to predict the average projected area of the breast parenchyma (*i.e.* excluding the adipose region) from the known compressed thickness. This area can then be multiplied by the average glandular thickness measured over the AEC sensor established above to give an extrapolation to glandular volume.

Figure 10 shows the estimated glandular volume as a function of compressed breast thickness for the six age ranges analysed above. The curves of Figure 10 show a

Figure 10. The estimated glandular volume as a function of compressed breast thickness for the six age ranges: (○) 49–53 years, (□) 54–58 years, (△) 59–63 years, (▽) 64–68 years, (◇) 69–73 years and (★) 74+ years.

similar shape, with a low average glandular volume for the smallest breasts as might be expected, and also a low average glandular volume for the largest breasts which have a greater adipose proportion. The apparent increase at the largest thickness is unlikely to be real, more likely being a combination of the simplistic area extrapolation and the relatively low number of original data points in these groups. The overall variation of the average glandular volume with breast size appears to be somewhat less than that of the variation with age. The regularity of the curves and smooth changes with age suggest that the approach to comparison with the statistics of a disease-free population explored here will be a suitable approach for full glandular volume data when that becomes available.

Conclusions

As the NHSBSP changes from screen-film to full-field digital mammography, there will be a greater opportunity to employ quantitative breast composition information to assist in the identification of high-risk individuals as well as the image information for direct diagnosis. Since the imaging exposures are already solidly justified in the breast-screening context, this potential additional source of information comes at no detriment to the screening client. In order to be able to identify individual women with an unusually high volume of glandular tissue, it will be necessary to make comparisons with a disease-free population. A large number of breast glandular thickness measurements were analysed to investigate the statistics of breast composition across a disease-free population as a test of a suitable methodology for identifying individuals at increased risk. It is demonstrated that a non-linear transformation can be used to produce normal statistical distributions, suggesting the

use of a standard “Z-score” approach to identify women with unusually glandular breasts. A further clinical study is under way to establish the relative risk to such women.

Acknowledgments

The author would like to express his sincere thanks to the radiographers of the Newcastle Breast Screening Programme for the extensive data collection required for this work, and to his colleague Maria Robinson for its organisation and collation.

References

- Harvey JA, Bovbjerg VE. Quantitative assessment of mammographic breast density: relationship to cancer risk. *Radiology* 2004;230:29–41.
- Hammerstein GR, Miller DW, White DR, Masterson ME, Woodward HQ, Laughlin JS. Absorbed radiation dose in mammography. *Radiology* 1979;130:485–91.
- Kaufhold J, Thomas JA, Eberhard JW, Galbo CE, Gonzalez Trotter DE. A calibration approach to glandular tissue composition estimation in digital mammography. *Med Phys* 2002;29:1867–80.
- Heine J, Behera M. Effective x-ray attenuation measurements with full field digital mammography. *Med Phys* 2006;33:4350–66.
- Jamal N, Ng K-H, Looi L-M, McLean D, Zulfiqar A, Tan S-P et al. Quantitative assessment of breast density from digitized mammograms into Tabar’s patterns. *Phys Med Biol* 2006;51:5843–57.
- Shepherd JA, Kerlikowske KM, Smith-Bindman R, Genant HK, Cummings SR. Measurement of breast density with dual x-ray absorptiometry: feasibility. *Radiology* 2002;223:554–7.
- Bryant RJ, Underwood AC, Robinson A, Stephenson TJ, Underwood JC. Determination of breast tissue composition for improved accuracy in estimating radiation doses and risks in mammographic screening. *Breast* 1998;7:95–8.
- Yaffe MJ. Measurement of mammographic density. *Breast Cancer Research* 2008;10:209. [cited 2010 Sep 13] Available from: <http://breast-cancer-research.com/content/10/3/209/>.
- Kotre CJ. X-ray absorptiometry of the breast using mammographic exposure factors: application to units featuring automatic beam quality selection. *Br J Radiol* 2010;83:515–23.
- NHS Breast Screening Programme annual review 2008 Saving lives through screening. NHS Cancer Screening Programme, Sheffield.
- Whitehead MI, Whitcroft SIJ, Hillard TC. An atlas of the menopause. New York, NY: Parthenon, 1993.
- Beckett JR, Kotre CJ. Dosimetric implications of age related glandular changes in screening mammography. *Phys Med Biol* 2000;45:801–13.
- Dance DR, Skinner CL, Young KC, Beckett JR, Kotre CJ. Additional factors for the estimation of mean glandular dose using the UK mammography dosimetry protocol. *Phys Med Biol* 2000;45:3225–40.
- Dance DR, Thilander Klang A, Sandborg M, Skinner CL, Castellano IA, Alm Carlson G. Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: a Monte Carlo study. *Br J Radiol* 2000;73:1056–67.