Acta Radiologica Short Reports
2014, Vol. 3(5) 1–5

DOI: 10.1177/2047981614536558
arr.sagepub.com

Comparison of Danish dichotomous and BI-RADS classifications of mammographic density

Rebecca Hodge[1,2], Sophie Sell Hellmann[1], My von Euler-Chelpin[1], Ilse Vejborg[3] and Zorana Jovanovic Andersen[1]

Abstract

Background: In the Copenhagen mammography screening program from 1991 to 2001, mammographic density was classified either as fatty or mixed/dense. This dichotomous mammographic density classification system is unique internationally, and has not been validated before.

Purpose: To compare the Danish dichotomous mammographic density classification system from 1991 to 2001 with the density BI-RADS classifications, in an attempt to validate the Danish classification system.

Material and Methods: The study sample consisted of 120 mammograms taken in Copenhagen in 1991–2001, which tested false positive, and which were in 2012 re-assessed and classified according to the BI-RADS classification system. We calculated inter-rater agreement between the Danish dichotomous mammographic classification as fatty or mixed/dense and the four-level BI-RADS classification by the linear weighted Kappa statistic.

Results: Of the 120 women, 32 (26.7%) were classified as having fatty and 88 (73.3%) as mixed/dense mammographic density, according to Danish dichotomous classification. According to BI-RADS density classification, 12 (10.0%) women were classified as having predominantly fatty (BI-RADS code 1), 46 (38.3%) as having scattered fibroglandular (BI-RADS code 2), 57 (47.5%) as having heterogeneously dense (BI-RADS 3), and five (4.2%) as having extremely dense (BI-RADS code 4) mammographic density. The inter-rater variability assessed by weighted kappa statistic showed a substantial agreement (0.75).

Conclusion: The dichotomous mammographic density classification system utilized in early years of Copenhagen's mammographic screening program (1991–2001) agreed well with the BI-RADS density classification system.

Keywords
Breast cancer, breast density, mammographic density, mammographic screening, Kappa

Date received: 19 January 2014; accepted: 30 April 2014

Introduction

European countries have the highest incidence of breast cancer in the world, with Denmark among those at the top, also regarding breast cancer mortality (1,2). The Copenhagen mammography screening program began in 1991 with the goal of detecting breast cancer at earlier stages and reducing mortality from the disease (3). During the early screening period from 1991–2001, one- or two-view mammograms were taken and mammographic density assessed by highly trained radiologists. Mammographic density refers to the regions of

[1]Center for Epidemiology and Screening, Department of Public Health, University of Copenhagen, Copenhagen, Denmark
[2]Danish Institute for Study Abroad, Copenhagen, Denmark
[3]Diagnostic Imaging Centre, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

Corresponding author:
Zorana Jovanovic Andersen, Center for Epidemiology and Screening, Department of Public Health, University of Copenhagen, Øster Farimagsgade 5, 1014 Copenhagen K, Denmark.
Email: Zorana.Andersen@sund.ku.dk

non-radiolucent, or bright, portions of the radiographic image of the breast, and represents the relative proportion of fibroglandular to fatty tissue (4–7).

High mammographic density is one of the strongest known risk factors for breast cancer (7,8), associated with a four- to six-fold increased risk of breast cancer when comparing women with high density in more than 75% of the breast to those with less than 5% mammographic density (4). Additional risk factors for breast cancer include the female gender, genetic inheritance, age, country of birth, parity, body mass index (BMI), age at onset of menarche and menopause, and hormone replacement therapy (HRT) (5,9,10).

There are different methods of measuring mammographic density, including quantitative, semi-quantitative, and qualitative scales (7). In the Copenhagen mammographic screening program from 1991 to 2001, mammographic density was classified either as fatty or mixed/dense by the reading radiologist in order to determine if the woman required a one or two-view mammography at her next screening. Another qualitative classification system that is widely used is the Breast Imaging Reporting and Data System (BI-RADS), which has four density categories: BI-RADS 1, which indicates predominantly fatty breasts; BI-RADS 2, indicating scattered fibroglandular breasts; BI-RADS 3, indicating heterogeneously dense breasts; and BI-RADS 4, denoting extremely dense breasts to the extent that masking of potential lesions is possible (6).

The dichotomous mammographic density classification used during the first 10 years of mammographic screening program in Copenhagen is unique internationally, and has not been validated before. In this study we compared the Danish dichotomous mammographic density classification system from 1991 to 2001 with the BI-RADS density classifications in an attempt to assess the agreement between the two and validate the Danish classification system.

Material and Methods

Copenhagen Mammography Register

The data for this study came from an independent population-based mammographic screening program organized within the Copenhagen municipality since 1991. Personalized invitations were sent to all women aged 50–69 years residing in the Copenhagen municipality to participate in biennial mammographic screening, free of charge. Invitations were based on the updated central population register (11) which included information on personal identification number, historical addresses, emigration, immigration, and vital status for all people residing in Denmark at any point since 1968. Approximately 40,000 women aged 50–69 years were invited at the start of each biennial invitation round.

Dichotomous mammographic density classification

At a woman's first initial screening, a two-view mammography, craniocaudal and oblique, was performed. Analog mammography was used during the study period, and mammograms were evaluated independently by visual assessment by two highly trained radiologists. Women designated as having fatty mammographic density and a negative screening was scheduled for a one-view oblique mammography at their subsequent mammography screening. Women with mixed/dense mammographic density and a negative screening would receive a two-view mammography at the subsequent mammographic screening. Mammograms would be compared with those from earlier screenings. Radiologists have recorded their recommendation for one- (1V) or two-view (2V) mammography at the subsequent screening for each woman, available as a variable from the Copenhagen Mammography Register, which is a proxy of fatty (1V) or mixed/dense (2V) mammographic density. Fatty (1V) mammographic density corresponds to BI-RADS code 1 and part of code 2, while mixed/dense (2V) mammographic density corresponds to part of BI-RADS code 2, and BI-RADS codes 3 and 4, but these data were never validated against BI-RADS codes.

This procedure was applied continuously from 1991 to 2001 and the dichotomized outcome for mammographic density was utilized in several studies (4,12,13). Starting in 2001 until required in 2004, standard procedure became the two-view mammography, and thus fatty or mixed/dense classifications could not be extrapolated from the use of one- or two-view mammograms.

There were no other historical changes in the screening program throughout the study period. Copenhagen's mammography screening program maintained compliance with the quality performance indicators specified by the European guidelines for the duration of the study (14). The dichotomized outcome for mammographic density has in a previous study performed in accordance with existing evidence of a positive relation between mammographic density and breast cancer risk, finding a doubled risk of breast cancer in women of mixed/dense as compared with fatty mammographic density, and was applied in several other published studies exploring associations of mammographic density with breast cancer risk factors (4,12,13). However, the dichotomous mammographic density scale has never before been compared to the BI-RADS density scales.

The study was entirely based on data from the Copenhagen Mammography Register and approved by the Danish Data Inspection Agency by Danish law serving as ethical approval of register-based research, which does not require informed consent from study participants.

BI-RADS mammographic density classification

Mammograms for 295 women who participated in Copenhagen mammographic screening between 1991–2001 and tested false positive were re-assessed in 2012 and classified according to BI-RADS classification, as part of a larger study by von Euler-Chelpin et al. (15), which examined the risk of breast cancer after false-positive testing. The 295 mammograms were visually assessed and BI-RADS density codes assigned by a single experienced radiologist, or, when in doubt, in a consensus with a second radiologist. Of these 295 available mammograms, 177 were excluded from this analysis: 13 due to missing BI-RADS codes, and 162 excluded for which the dichotomous mammographic codes were assigned at or after the false-positive test, to avoid a possible misclassification bias from an increased tendency of a radiologist to request two-view mammography after a false-positive test.

Our study sample consisted of 120 mammograms taken between 1991 and 2001, for which dichotomous mammographic classification was assigned based on the latest negative screen prior to the false-positive screen, and which were in 2012 re-assessed and classified according to the BI-RADS classification system.

Statistical analysis

We calculated the linear weighted Kappa statistic between the Danish dichotomous mammographic classification as fatty (1V) or mixed/dense (2V) and the four-level BI-RADS density classification, by the *kappa* procedure in STATA version 11.2.

Results

The mean age at screening of the 120 women in the study was 58.3 years, with a standard deviation of 5.1 years. The mean age (standard deviation) in women with fatty breast was 61.2 (4.7) years and in women with mixed/dense breast 57.3 (4.8) years.

Out of the 120 women, 32 (26.7%) were classified as having fatty (1V), and 88 (73.3%) as having mixed/dense (2V) mammographic density (Table 1). According to BI-RADS density classification, 12 (10.0%) women were classified as having predominantly fatty breasts (BI-RADS code 1), 46 (38.3%) as having scattered fibroglandular breasts (BI-RADS code 2), 57 (47.5%) as having heterogeneously dense breasts (BI-RADS 3), and five (4.2%) as having extremely dense breasts (BI-RADS code 4).

Table 1. Comparison of the Danish dichotomous and BI-RADS classifications of mammographic density.

	Danish dichotomous classifications		
BI-RADS classification	(1) Fatty	(2) Mixed/ Dense	Total
(1) Predominantly fatty	11 (91.7%)	1 (8.3%)	12 (10.0%)
(2) Scattered	19 (41.3%)	27 (58.7%)	46 (38.3%)
(3) Heterogeneous	2 (3.5%)	55 (96.5%)	57 (47.5%)
(4) Extremely dense	0 (0.0%)	5 (100.0%)	5 (4.2%)
Total	32 (26.7%)	88 (73.3%)	120

Of 12 women classified as having predominantly fatty breasts by BI-RADS classification (code 1), 11 (91.7%) were also classified as fatty (1V) by the Danish classification, and only one woman (8.3%) was misclassified as having mixed/dense breasts (2V) (Table 1). Of the five women classified as having extremely dense breasts by BI-RADS classification (code 4), all were classified as having mixed/dense breasts by dichotomous classification, achieving 100% agreement in this category. Of the 57 women classified as having heterogeneously dense breasts by the BI-RADS classification (code 3), 55 (96.5%) were classified as having mixed/dense breasts by the Danish classification system, while two (3.5%) were misclassified as having fatty breasts. Women with mammographic density classified in the BI-RADS code 2 of scattered fibroglandular density were, as expected, found in both categories of the Danish dichotomous classification system, with 27 (58.7%) in the mixed/dense category in the Danish classification system, and 19 (41.3%) in the fatty category.

According to interpretation by Landis JR and Koch GG (16), we found a substantial inter-rater agreement between the Danish classification system and BI-RADS mammographic density score assessed by kappa statistics (0.75).

Discussion

We found a substantial agreement between the Danish dichotomous and BI-RADS four-level classification system, based on a sample of 120 mammograms obtained from the Copenhagen Mammographic Register between 1991 and 2001. The agreement was excellent for BI-RADS categories 1, 3, and 4, and, as expected, poorer in category 2, which according to the Danish dichotomous classification, represented a mix

of women with fatty and mixed/dense mammographic densities.

The dichotomous mammographic density classification utilized in the Copenhagen, mammographic screening program from 1991 to 2001 agreed well with the four-level BI-RADS classification system in 75% of screened women. The BI-RADS category 2, denoting scattered density, which comprised of 38.3% of the study population, included parts of both the fatty and mixed/dense mammographic densities according to the Danish classification protocol. For all other BI-RADS density categories, the Danish dichotomous method mammographic density classification was directly comparable to the four-level BI-RADS density classification. Only three (4.2%) out of 74 mammograms classified in BI-RADS categories 1, 3, or 4, were misclassified by the Danish dichotomous system. There was 100% agreement with the BI-RADS category 4, that of women with extremely dense breasts, and 91.7% agreement (only 1 misclassified mammogram) in the BI-RADS category 1, of predominantly fatty breasts.

The Danish dichotomous classification was derived from the radiologist's assessment of a need for a one- or two-view mammogram in future mammographic screenings based on a given mammogram, because of potential masking of lesions associated with mixed/dense breast. The Danish classification was not developed as a systematic method of estimating a woman's mammographic density. In contrast, the BI-RADS density classification was developed to standardize reporting of mammographic findings and aid in the estimation of a woman's individual risk of breast cancer outcomes, as well as to facilitate research (17). However, our study provides evidence that the Danish classification of mammographic density is comparable to and in substantial agreement with the BI-RADS classification system widely used in US clinical radiology.

The dichotomous outcome has been utilized successfully in several earlier studies, and validated by showing the expected associations of mammographic density with breast cancer risk (12) and breast cancer mortality (12), childhood obesity (13), and birth cohort (4). Olsen et al., studying a total of 48,052 women participating in Copenhagen mammography screening program between 1991 and 2001, showed an expected doubling of the breast cancer risk in women with mixed/dense compared to women with fatty breasts, with a hazard ratio (HR) of 2.45 (95% confidence interval [CI], 2.14–2.81) (12), as did Andersen et al. (HR, 2.34; 95% CI, 1.97–2.78) in a subset of 13,572 screened women for whom childhood obesity data were available (13). Both of these studies show the expected doubling of breast cancer risk when comparing mixed/dense to fatty mammographic density, in agreement with Boyd et al. (7).

Our study was based on objectively collected data on mammograms from a comprehensive registry of the Copenhagen mammography screening program, which performs according to international quality indicators for mammographic screening limiting the possibility for bias (1). Some limitations of this study include the aforementioned use of women who experienced false-positive tests and who also later developed breast cancer, which were utilized in a study by von Euler-Chelpin et al. (14), and not selected for the purpose of this paper. These women may have different compositions of breast tissues than women who never received a false-positive test or a breast cancer diagnosis, since high mammographic density is associated with both higher risk of false-positive tests and risk of breast cancer (7,12,13). Indeed, the 120 women in our study population were older at screening (mean age, 58 years), and were more likely to have mixed/dense breasts (73.3%) than the background screening population in Copenhagen in 1991–2001, with a mean age of 55 years and between 55% and 60% with mixed/dense breasts (4,12,13). However, this would not invalidate this study's results on the agreement between two independent mammographic density classification scales. Another limitation of this study was that the BI-RADS categories were assigned by one radiologist, or when in doubt, in consensus with another radiologist. As there is a considerable inter-observer and reader variability in reading mammograms in general, and especially in assessing BI-RADS codes (18,19), it would have been more optimal to have two or more radiologist evaluate each mammogram. Another weakness of the study was the small size of the study population. A repeat analysis using a larger and more randomly chosen percentage of the screening population could allow for more definitive results in future studies and could potentially allow for better interpretation and understanding of the BI-RADS category 2 with regards to the Danish classification system's two categories. Still, this is the only dataset available with assigned BI-RADS codes on the historical data from mammographic screening in Copenhagen in 1991–2001.

In conclusion, the dichotomous mammographic density classification utilized in the Copenhagen's mammographic screening program between 1991 and 2001 agrees well with the BI-RADS classification system.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

1. Vejborg I, Olsen AH, Jensen MB, et al. Early outcome of mammography screening in Copenhagen 1991–99. J Med Screen 2002;9:115–119.
2. Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10, International Agency for Research on Cancer, http://globocon.iarc.fr (2010, cited 6 November 2013).
3. Utzon-Frank N, Vejborg I, von Euler-Chelpin M, et al. Balancing sensitivity and specificity: sixteen year's of experience from the mammography screening programme in Copenhagen, Denmark. Cancer Epidemiol 2011;35: 393–398.
4. Hellmann SS, Lynge E, Schwartz W, et al. Mammographic density in birth cohorts of Danish women: a longitudinal study. BMC Cancer 2013;13:409.
5. Hulka BS, Moorman PG. Breast cancer: hormones and other risk factors. Maturitas 2001;38:103–113.
6. Yaffe MJ. Mammographic density. Measurement of mammographic density. Breast Cancer Res 2008;10:209.
7. Boyd NF, Martin LJ, Yaffe MJ, et al. Mammographic density: a hormonally responsive risk factor for breast cancer. J Br Menopause Soc 2006;12:186–193.
8. Tice JA, Cummings SR, Ziv E, et al. Mammographic breast density and the Gail model for breast cancer risk prediction in a screening population. Breast Cancer Res Treat 2005;94:115–122.
9. Heine JJ, Malhotra P. Mammographic tissue, breast cancer risk, serial image analysis, and digital mammography. Part 1. Tissue and related risk factors. Acad Radiol 2002;9:298–316.
10. Wolfe JN, Saftlas AF, Salane M. Mammographic parenchymal patterns and quantitative evaluation of mammographic densities: a case-control study. Am J Roentgenol 1987;148:1087–1092.
11. Pedersen CB. The Danish Civil Registration System. Scand J Public Health 2011;39:22–25.
12. Olsen AH, Bihrmann K, Jensen MB, et al. Breast density and outcome of mammography screening: a cohort study. Br J Cancer 2009;100:1205–1208.
13. Andersen ZJ, Baker JL, Bihrmann K, et al. Birth weight, childhood BMI and height in relation to mammographic density and breast cancer: register based cohort study. Breast Cancer Res 2014;16:R4.
14. Vejborg I, Olsen AH, Jensen MB, et al. Early outcome of mammography screening in Copenhagen 1991–99. J Med Screen 2002;9:115–119.
15. von Euler-Chelpin M, Risor LM, Thorsted BL, et al. Risk of breast cancer after false-positive test results in screening mammography. J Natl Cancer Inst 2012;104: 682–689.
16. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33: 159–174.
17. Lehman C, Holt S, Peacock S, et al. Use of the American College of Radiology BI-RADS Guidelines by Community Radiologists: Concordance of Assessments and Recommendations Assigned to Screening Mammograms. Am J Roentgenol 2002;179:15–20.
18. Ciatto S, Houssami N, Apruzzese A, et al. Reader variability in reporting breast imaging according to BI-RADS assessment categories (the Florence experience). Breast 2006;15:44–51.
19. Redondo A, Comas M, Macià F, et al. Inter- and intraradiologist variability in the BI-RADS assessment and breast density categories for screening mammograms. Br J Radiol 2012;85:1465–1470.