

## ORIGINAL

# Valoración de las categorías BI-RADS® de la densidad mamaria entre radiólogos y respecto a un método cuantitativo de medición



Olga Martín Ruiz\*, Cristina Romero Castellano, Paul Martín Aguilar Angulo, Virginia Segur Nieto, Asunción Almenar Gil, Raquel Moreno de la Presa y José María Pinto Varela

Servicio de Radiodiagnóstico, Hospital Virgen de la Salud, Toledo, España

Recibido el 7 de marzo de 2014; aceptado el 3 de julio de 2014

Disponible en Internet el 26 de septiembre de 2014

### PALABRAS CLAVE

Densidad mamaria;  
Mamografía digital;  
Clasificación  
BI-RADS®

### Resumen

**Objetivo:** Valorar la concordancia al clasificar las densidades mamarias empleando el sistema BI-RADS® y la variabilidad entre este método empleado por un lector experto en comparación con los valores obtenidos mediante un sistema de medición automática (Quantra®).

**Pacientes y métodos:** Estudio descriptivo en 2 fases. Primero se evaluó la clasificación BI-RADS® de 5 lectores de una muestra seleccionada de forma aleatoria y estratificada por grupos (516 pacientes), para medir la variabilidad intraobservador e interobservador (índice kappa). Finalmente se analizó la concordancia entre el radiólogo con más experiencia y el software Quantra® que permite la valoración volumétrica de la densidad.

**Resultados:** La concordancia por parejas varió entre débil a moderada (kappa 0,35-0,67), el grupo presentó un valor moderado de reproducibilidad (0,58), siendo mayor en las densidades extremas y en los grupos de edad límite. La variabilidad intraobservador fue de moderada a buena (0,52-0,79). Respecto a la clasificación de la densidad por categorías entre el lector experto y la aplicación, no se evidenció concordancia (kappa = 0,0737).

**Conclusiones:** Los valores de concordancia obtenidos corroboran la evidencia descrita en la literatura, considerando que la lectura la realizaron radiólogos con experiencia y dedicación exclusiva en mama. No se apreció concordancia en la clasificación lector versus Quantra® (esta pondera las densidades en menor porcentaje respecto al análisis visual), y esto podría deberse al punto de corte por categoría.

© 2014 SESPM. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

\* Autor para correspondencia.

Correo electrónico: [olgamartinru@hotmail.com](mailto:olgamartinru@hotmail.com) (O. Martín Ruiz).

**KEYWORDS**

Breast density;  
Digital  
mammography;  
BI-RADS®  
classification

## Evaluation of the BI-RADS® categories between readers and with regard to a quantitative method

**Abstract**

**Objective:** To assess agreement in breast density classification using the BI-RADS® system and the concordance between this measurement method compared with a dedicated software program (Quantra®).

**Patients and methods:** A descriptive study was carried out in 2 phases. First, we analyzed the BI-RADS® breast density classification assessed by 5 readers in a sample of 516 patients randomly selected and stratified into groups. Intraobserver and interobserver reproducibility were calculated (kappa index). Second, we evaluated the agreement between the most experienced reader and a software package that calculates the volumetric fraction of fibroglandular tissue (Quantra®).

**Results:** The agreement between pairs ranged from fair to moderate (kappa 0.35-0.67). The reading group showed moderate reproducibility (0.58) with the highest values in high breast densities and in the youngest and oldest patients. Intraobserver agreement ranged from moderate to substantial (0.52-0.79). In density classification, there was no evidence of agreement between the expert reader and the software application (kappa = 0.0737).

**Conclusions:** The agreement values obtained corroborate the evidence described in the literature on reading performed by experienced radiologists working exclusively with mammograms. No agreement was found in reader classification versus Quantra® (which classifies densities with a lesser percentage than visual analysis), which could be due to the cut off by category.

© 2014 SESPM. Published by Elsevier España, S.L.U. All rights reserved.

## Introducción

Numerosos autores han expuesto que una alta densidad mamográfica (expresada como la razón entre el parénquima mamario y la grasa radiolúcida adyacente) incrementa el riesgo de cáncer de mama. Las causas potenciales del incremento del riesgo son, al parecer, multifactoriales<sup>1,2</sup>. Según el modelo de Pike et al.<sup>3</sup>, la exposición acumulada del tejido mamario a las hormonas, más la presencia de factores de crecimiento que estimulan la división celular, combinada con un daño genético, incrementarían la incidencia de cáncer de mama.

Los estudios epidemiológicos, aunque heterogéneos, evidencian que la presencia de una proporción de tejido glandular por encima del 75% se asocia a un mayor riesgo relativo de cáncer de mama<sup>4</sup>.

Además, hay que resaltar que el problema de la densidad mamográfica no es solo el aumento del riesgo de cáncer de mama, sino una marcada disminución de la sensibilidad de la mamografía<sup>5</sup>.

Dada la relevancia cada vez más importante de definir la densidad mamaria, parece conveniente hallar el método más adecuado que nos permita cuantificar y definir el tipo de tejido en la mamografía<sup>6,7</sup>.

El hecho de que la densidad mamaria sea un marcador de riesgo tal vez nos obligue a desarrollar modelos de medición de densidad mamaria lo más reproducibles posible.

Se han descrito diferentes métodos para la valoración de la densidad empleando escalas de valoración cualitativa (Wolfe<sup>8</sup>, descriptores BI-RADS®<sup>9</sup>; esta última es la más empleada en la práctica diaria y estratifica la densidad mamaria según porcentajes durante la lectura mamográfica en: mamas de predominio graso [ $<25\%$ ], patrón de

densidades glandulares parcheadas [25-50%], heterogéneas [51-75%] y extremadamente densas [ $\geq 75\%$ ]) y diferentes técnicas de valoración cuantitativa (como, por ejemplo, la valoración volumétrica de la densidad a través de un software como es el Quantra®), que tratan de aproximarse de manera más objetiva y precisa a la medición de la densidad mamaria. Otras técnicas de medición cuantitativa incluyen métodos bidimensionales (planimetría, análisis de textura)<sup>10</sup>, volumétricos (TC, tomosíntesis, absorciometría de rayos X, entre otros)<sup>11</sup> y aquellos que no emplean radiaciones ionizantes (como el ultrasonido y la resonancia magnética)<sup>12</sup>.

Los objetivos del presente estudio son valorar el grado de concordancia entre lectores expertos de nuestro centro al clasificar las densidades mamarias usando el sistema BI-RADS®, y además, evaluar la concordancia entre el método cualitativo descrito, aplicado por el lector de mayor experiencia, frente a los valores obtenidos mediante la aplicación Quantra®.

## Material y métodos

Se trata de un estudio descriptivo en 2 fases. El periodo de inclusión fue desde enero de 2011 a enero de 2012. Este estudio fue aprobado por el Comité Ético de nuestro hospital. En la primera fase se evalúa la categorización BI-RADS® entre lectores: se incluyeron un total de 6.521 mujeres. Los criterios de inclusión fueron: mujeres de más de 30 años, con mamografía completa (proyecciones en craneocaudal y oblicuamediolateral de cada mama) considerada normal (mediante la lectura única por un radiólogo, apoyado con un sistema de detección asistido por ordenador), tener al menos 2 mamografías previas normales realizadas en

**Tabla 1** Variabilidad intraobservador tras la segunda lectura de mamografías de la muestra

Radiólogo	Radiólogo				
	1	2	3	4	5
1	0,7900	0,4856	0,5831	0,5445	0,6250
2		0,5200	0,3514	0,3615	0,4477
3			0,6700	0,6656	0,5618
4				0,6200	0,5388
5					0,5900

nuestro centro, tener 2 mamografías en 2 años posteriores, que también fueran normales, realizadas en nuestro centro. Se descartaron las portadoras de prótesis (205 mamografías), y de las restantes, se seleccionaron 516 mamografías (7,9%) de forma aleatoria y estratificada por grupos de edad (30-44 años: 169 mujeres, 45-54 años: 181 mujeres, 55-64 años: 106 mujeres, 65-más años: 60 mujeres).

Todas las mamografías fueron realizadas en un mamógrafo digital MAMMOMAT® (Siemens Healthcare).

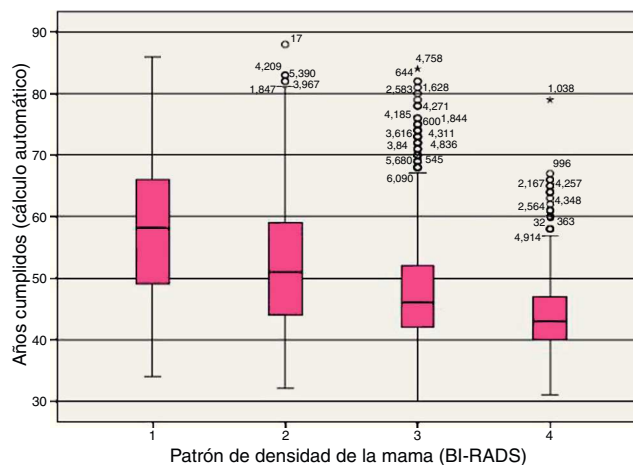
Las mamografías de la muestra seleccionada fueron leídas de manera independiente y alterando el orden, por 5 radiólogos expertos (con una experiencia profesional de entre 2 y 12 años, con lectura de al menos 5.000 mamografías/año) que clasificaron la densidad de las mamografías según las categorías BI-RADS®. Posteriormente, se realizó una segunda lectura por los mismos radiólogos de la misma muestra, alterando el orden de las mamografías para medir la variabilidad intraobservador. Con los resultados obtenidos de cada uno de los 5 lectores, se calculó el índice kappa con el fin de analizar la concordancia entre ellos.

En una segunda fase se correlacionó la concordancia entre el lector de mayor experiencia y los valores Quantra® (software de Hologic®). Se analizaron un total de 1.678 mamografías sin hallazgos patológicos (2 proyecciones por cada mama), consideradas como normales por un radiólogo con dedicación exclusiva con 12 años de experiencia. La lectura se realizó en lectura única, clasificando la densidad de las mamografías según la categorización BI-RADS®. Simultáneamente, para cada una de las mamografías se obtuvo el porcentaje de tejido fibroglandular mediante el método Quantra®. En primer lugar, se calculó el índice kappa entre el valor BI-RADS® de cada mamografía y su correspondiente valor de Quantra® reagrupado en 4 categorías (< 25%, 25-50%, 51-75% y ≥ 75%) para evaluar la concordancia de resultados entre ambos métodos. Luego se compararon los valores promedio de los porcentajes obtenidos con la herramienta Quantra® de cada una de las categorías de BI-RADS® con el propósito de objetivar la posible diferencia entre ellas, empleándose para ello el análisis de la variancia.

Los cálculos estadísticos se realizaron con el programa estadístico SPSS versión 18.0.0 y con el programa Stata/IC versión 11.

## Resultados

En el análisis de la categorización BI-RADS® por lectores se observó que la proporción de tejido glandular fue inversamente proporcional a la edad, siguiendo un modelo lineal aproximado (fig. 1).



**Figura 1** Distribución de las densidades mamarias según la edad.

La concordancia por parejas de radiólogos osciló en valores de kappa de 0,35 a 0,67. La variabilidad intraobservador osciló entre los valores de 0,52 y 0,79 (tabla 1).

La concordancia global entre radiólogos según grupo de edad presentó un índice kappa que varió entre 0,49 y 0,56 para los intervalos 45-54 y más de 65 años, respectivamente. En cuanto a la densidad mamográfica, se obtuvo un valor kappa global de la muestra de 0,58, observándose los máximos valores en las densidades BI-RADS® 1 y 4 (0,54 y 0,64, respectivamente) (tabla 2).

Por otro lado, el índice kappa entre la clasificación BI-RADS® de densidades realizada por el radiólogo de más experiencia del grupo y los porcentajes obtenidos con la herramienta Quantra® reagrupados (según intervalos de densidad similares a la clasificación ACR) fue de 0,07 (p = 1,0) (tabla 3).

**Tabla 2** Concordancia global y por categoría BI-RADS® de los lectores en su conjunto

BI-RADS®	Kappa	Z	Prob > Z
1	0,5451	39,08	0,0000
2	0,4102	29,41	0,0000
3	0,5161	37,00	0,0000
4	0,6383	45,76	0,0000
Combinado	0,5796	60,89	0,0000

**Tabla 3** Valores de concordancia de la clasificación BI-RADS® por un lector experto respecto a la aplicación Quantra®

Concordancia		Kappa	Error estándar	Z	Prob > Z
Observada	Esperada				
16,21%	21,96%	0,0737	0,0098	7,53	1,0000

**Tabla 4** Número de mujeres clasificadas según la densidad mamaria (BI-RADS®) por lectura y según los valores obtenidos por la aplicación Quantra®

Quantra®	BI-RADS®				Total
	1	2	3	4	
< 25	227	717	227	6	1.227
25-49	1	41	305	85	432
50-74	0	0	4	15	19
≥ 75	0	0	0	0	0
Total	228	758	586	106	1.678

En cuanto a la distribución según el número de mujeres, las categorías BI-RADS® y los valores obtenidos a través del software descrito, el hecho de no encontrar ningún valor recodificado en la categoría 4 y muy pocos en la 3 refuerza la hipótesis de no concordancia (tabla 4).

Con el análisis de la variancia se obtuvo un valor F de 738,97 ( $p < 0,001$ ), lo que indica que los valores obtenidos por cada categoría mediante el sistema Quantra® (< 25%, 25-49%, 50-74%, ≥ 75%) eran, estadísticamente, diferentes entre sí.

## Discusión

La densidad mamaria es un importante valor en el momento de predecir el riesgo de cáncer de mama y así seleccionar a potenciales pacientes en riesgo. Independientemente de los múltiples factores que condicionan una elevada densidad en cierto número de mujeres, lo cierto es que el patrón mamario denso reduce la sensibilidad de la mamografía como herramienta de cribado<sup>13</sup>.

La densidad mamaria ha ido adquiriendo cada vez más importancia, tanto en el momento de la emisión del informe radiológico estructurado, que debe incluir el tipo de densidad (nueva legislación americana sobre redacción del informe mamográfico, [http://leginfo.ca.gov/faces/billNavClient.xhtml?bill\\_id0201120120SB1538](http://leginfo.ca.gov/faces/billNavClient.xhtml?bill_id0201120120SB1538)), como en la valoración del riesgo y la propia fisiopatología del cáncer de mama.

La aplicación Quantra® es un software para los sistemas de mamografía digital que permite la cuantificación de la densidad mamaria. La imagen digital obtenida en un mamógrafo digital es procesada creando un modelo volumétrico que calcula, en las proyecciones craneocaudal y oblicuamediolateral, el volumen mamario, el volumen de tejido fibroglandular y el porcentaje de tejido fibroglandular, resultante del cociente de los 2 anteriores. El método Quantra® mejora las anteriores técnicas de medida volumétricas (Highnam et al. y Ding et al.)<sup>14,15</sup> al introducir un nuevo algoritmo de medida basado en parámetros de

medición de la resonancia magnética y al demostrar mayor grado de reproducibilidad de sus resultados (Hartman et al.)<sup>16</sup>.

En relación con nuestros resultados, al valorar la clasificación BI-RADS® en la lectura por 5 radiólogos con dedicación exclusiva a la enfermedad mamaria, la concordancia por parejas varió entre débil y moderada (0,35-0,67), objetivándose un valor moderado de concordancia del grupo en conjunto (kappa global de 0,58), con los mayores valores en las densidades extremas (tipo 1 y 4) y en los grupos de edad límite (30-44 y mayores de 65 años). La concordancia intraobservador varió de moderada a buena (0,52-0,79).

En estudios anteriores<sup>17,18</sup> encontramos resultados similares a los nuestros, con una variabilidad interobservador moderada (kappa de 0,59; 75% de acuerdo) y una mejor variabilidad intraobservador de la media (kappa de 0,72; 82% de acuerdo). Estudios posteriores<sup>19,20</sup> han demostrado una alta reproducibilidad interobservador (kappa 0,69-0,84) usando el sistema BIRADS® como medida de la densidad mamaria. Sin embargo, incluso estos estudios, con alta reproducibilidad interobservador, admiten que, aunque la clasificación visual parece reproducible en una base estadística, podría subestimar la densidad mamaria en la práctica clínica, ya que los sujetos clasificados como densidad tipo 3-4 variaron en un 25 a un 50%, según el lector clasificador.

En nuestro estudio, los valores de concordancia obtenidos corroboran la evidencia descrita (a falta de una modalidad objetiva de evaluación consensuada y validada), teniendo en cuenta siempre que los intervalos presentados provienen de la lectura realizada por radiólogos con experiencia y dedicación exclusiva, y que evidentemente podrían tener una mayor variación si fuera otro el caso.

Al comparar la clasificación de los porcentajes de densidad tras la aplicación del software versus la lectura realizada por el radiólogo de mayor experiencia, no existía concordancia (kappa = 0,0737). Esto se produce debido a que la herramienta Quantra® brinda porcentajes clasificados mayoritariamente en las categorías 1 y 2 (< 25% de tejido fibroglandular), mientras que en el radiólogo predominan las

mamografías clasificadas como tipos 2 y 3 (50-75% de tejido fibroglandular). Esto concuerda con los resultados obtenidos en estudios posteriores<sup>19,20</sup>, que concluyen que el software proporciona sistemáticamente valores inferiores a los apreciados visualmente mediante el sistema BI-RADS®.

Dichos resultados son indicativos de la no equivalencia entre ambos métodos, que podría deberse a los puntos de corte que usa el sistema Quantra® para cada categoría (que tendrían que variar para obtener valores parecidos a la lectura visual mediante el sistema BI-RADS®).

Esta es una de las principales limitaciones que encontramos al realizar nuestro estudio, ya que estamos comparando medidas dadas por un sistema informático, completamente objetivo, que aplica complejos cálculos matemáticos basados en resonancia magnética para medir la densidad mamográfica, frente a medidas dadas por un radiólogo experto, especialista en mama, que aporta inconscientemente datos subjetivos (edad, necesidad de realizar ecografía, posibles hallazgos ocultos por la densidad mamográfica elevada, etc.) al aplicar el sistema BI-RADS® en la determinación de la densidad mamográfica.

En conclusión, según los resultados obtenidos en nuestro estudio, la densidad mamaria es inversamente proporcional a la edad, siendo las pacientes con mayor densidad las que presentan mayor riesgo de cáncer de mama, así como menor sensibilidad de la mamografía en el diagnóstico. La concordancia entre radiólogos expertos usando el sistema de clasificación BI-RADS® como método de medida de la densidad varía de débil a moderada, lo que implica una alta subjetividad. Así mismo, no encontramos concordancia entre las medidas obtenidas con la aplicación Quantra® y el radiólogo de mayor experiencia, ya que la aplicación Quantra® proporciona valores de la densidad mamaria sistemáticamente inferiores a los de la medición visual mediante el sistema BI-RADS®, por lo que habría que buscar puntos de corte diferentes, para que ambos métodos se aproximaran más.

Por tanto, se hacen necesarios mayores estudios para obtener una aplicación útil, válida, reproducible y aceptada, en general, que sea equivalente al léxico BI-RADS® visual, actualmente tan empleado por los radiólogos.

## Responsabilidades éticas

**Protección de personas y animales.** Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

**Confidencialidad de los datos.** Los autores declaran que en este artículo no aparecen datos de pacientes.

**Derecho a la privacidad y consentimiento informado.** Los autores declaran que en este artículo no aparecen datos de pacientes.

## Autoría

1. Obtención de datos, análisis e interpretación de los datos, redacción y revisiones del texto, aprobación de la versión final: Olga Martín Ruiz.

2. Concepción y diseño del artículo, obtención, análisis e interpretación de los datos, revisión del texto, aprobación de la versión final: Cristina Romero Castellano.
3. Obtención de datos, análisis e interpretación de los datos, redacción, revisiones del texto, aprobación de la versión final: Paul Martín Aguilar Angulo.
4. Obtención y análisis de datos, revisiones del texto, aprobación de la versión final: Virginia Segur Nieto.
5. Obtención y análisis de datos, revisiones del texto, aprobación de la versión final: Asunción Almenar Gil.
6. Obtención y análisis de datos, revisiones del texto, aprobación de la versión final: Raquel Moreno de la Presa.
7. Concepción y diseño del artículo, revisiones del texto, aprobación de la versión final: José María Pinto Varela.

## Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

## Bibliografía

1. Ginsburg OM, Martin LJ, Boyd NF. Mammographic density, lobular involution, and risk of breast cancer. *Br J Cancer*. 2008;99:1369-74.
2. Boyd NF, Martin LJ, Chavez S, Gunasekara A, Salleh A, Melnichouk O, et al. Breast-tissue composition and other risk factors for breast cancer in young woman: A cross-sectional study. *Lancet Oncol*. 2009;10:569-80.
3. 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;303:767-70.
4. McComarck VA, dos Santos Silva I. Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiol Biomarkers Prev*. 2006;15:1159-69.
5. Saftlas AF, Hoover RN, Brinton LA, Szklo M, Olson DR, Salame M, et al. Mammographic densities and risk of breast cancer. *Cancer*. 1991;67:2833-8.
6. Pollán M, Ascunce N, Ederra M, Murillo A, Erdozain N, Alés-Martínez JE, et al. Mammographic density and risk of breast cancer according to tumor characteristics and mode of detection: A Spanish population-based case-control study. *Breast Cancer Res*. 2013;15:1-11.
7. Goodwin PJ, Boyd NF. Mammographic parenchymal pattern and breast cancer risk: A critical appraisal of the evidence. *Am J Epidemiol*. 1988;127:1097-108.
8. Wolfe J. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer*. 1976;37:2486-92.
9. Breast Imaging Reporting and Data System: ACR BI-RADS®. Breast imaging atlas. 5th ed. Reston, VA: American College of Radiology; 2013.
10. Boyd NF, Byng JW, Jong RA, Fishell EK, Little LE, Miller AB, et al. Quantitative classification of mammographic densities and breast cancer risk: Results from the Canadian National Breast Screening Study. *J Natl Cancer Inst*. 1995;87:670-5.
11. Despina K, Bakic PR, Carton AC, Troxel A, Conant E, Andrew DA. Parenchymal texture analysis in digital breast tomosynthesis for breast cancer risk estimation: A preliminary study. *Acad Radiol*. 2009;16:283-98.
12. Nie K, Chen JH, Chan S, Chau MK, Yu JH, Bahri S, et al. Development of a quantitative method for analysis of breast density based on three-dimensional breast MRI. *Med Phys*. 2008;35:5253-62.

13. Highnam RP, Pan X, Warren R, Jeffreys M, Davey Smith G, Brady M. Breast composition measurements using retrospective standard mammogram form (SMF). *Phys Med Biol.* 2006;51:2695–713.
14. Highnam RP, Jeffreys M, McCormack V, Warren R, Davey Smith G, Brady M. Comparing measurements of breast density. *Phys Med Biol.* 2007;52:5881–95.
15. Ding J, Warren R, Warsi I, Day N, Thompson D, Brady M, et al. Evaluating the effectiveness of using standard mammogram form to predict breast cancer risk: Case-control study. *Cancer Epidemiol Biomarkers Prev.* 2008;17:1074–81.
16. Hartman K, Highnam R, Warren R, Jackson V. Volumetric assessment of breast tissue composition from FFDM images. *Digital mammography Lecture Notes in Computer Science.* 2008;5116:33–9.
17. Kerlikowske K, Grady D, Barclay J, Frankel SD, Ominsky SH, Sickles EA, et al. Variability and accuracy in mammographic interpretation using the American College of Radiology Breast Imaging Reporting and Data System. *J Natl Cancer Inst.* 1998;90:1801–9.
18. Kerlikowske K, Ichikawa L, Miglioretti DL, Buist DS, Vacek PM, Smith-Bindman R, et al., National Institutes of Health Breast Cancer Surveillance Consortium. Longitudinal measurement of clinical mammographic breast density to improve estimation of breast cancer risk. *J Natl Cancer Inst.* 2007;99:386–95.
19. Ciatto S, Bernardi D, Calabrese M, Durando M, Gentilini MA, Mariscotti G, et al. A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *Breast.* 2012;21:503–6.
20. Rafferty, E, Smith, A, Niklason, L, Comparison of three methods of estimating breast density: BI-RADS density scores using full field digital mammography, BI-RADS density scores using breast tomosynthesis, and volumetric breast density. *Radiological Society of North America 2009 Scientific Assembly and Annual Meeting, November 29-December 4, 2009, Chicago IL.*