

Document downloaded from:

<http://hdl.handle.net/10251/104248>

This paper must be cited as:

Martinez Gomez, I.; Casals El Busto, M.; Antón Guirao, J.; Ruiz Perales, F.; Llobet Azpitarte, R. (2014). Estimación semiautomática de la densidad mamaria con DM-Scan. Radiología. 56(5):429-434. doi:10.1016/j.rx.2012.11.007



The final publication is available at

<https://doi.org/10.1016/j.rx.2012.11.007>

Copyright Elsevier

Additional Information



La estimación semiautomática de la densidad mamaria con DM-Scan mejora la reproducibilidad y disminuye la variabilidad de la estimación visual

Resumen

Objetivo. Evaluar la reproducibilidad del cálculo de la densidad mamaria con la aplicación informática DM-Scan, basada en la segmentación semiautomática del tejido fibroglandular, y compararla con la de la inspección visual.

Material y métodos. El estudio incluyó 655 mamografías digitales directas en proyección craneo-caudal. Tres radiólogos expertos analizaron la densidad de las mamografías con DM-Scan, y se calcularon las concordancias inter e intraobservador entre pares de radiólogos para las escalas Boyd y BI-RADS®, utilizando el índice de correlación intraclase. Las concordancias se compararon con las obtenidas previamente para la inspección visual, en el mismo conjunto de imágenes, utilizando el índice Kappa.

Resultados. Con el análisis visual, la concordancia media interobservador fue de 0,876 (IC95% 0,873-0,879) para la escala de Boyd y 0,823 (IC95% 0,818-0,829) para BI-RADS®,. La concordancia intraobservador fue de 0,813 (IC95% 0,796-0,829) para la escala de Boyd, y 0,770 (IC95% 0,742-0,797) para la clasificación BI-RADS®. Con DM-scan, la concordancia media inter e intraobservador fue de 0,92, notablemente superior a las concordancias de la clasificación visual.

Conclusión. El cálculo de la densidad mamaria con la aplicación semiautomática DM Scan es más fiable y reproducible y disminuye la subjetividad y variabilidad de la estimación visual.

Palabras clave. Mamografía; Cáncer de mama; Densidad Mamaria; Procesamiento Digital de Mamografías; Diagnóstico Asistido por Ordenador; Variabilidad Intraobservador; Variabilidad Interobservador.

Introducción

Los programas de cribado con mamografía han permitido disminuir la mortalidad del cáncer de mama hasta un 22% en mujeres de más de 50 años, y un 15% en las mujeres de 40-49 años ([1,2](#)). Pero la sensibilidad de la mamografía disminuye con el aumento de la densidad mamaria que además se asocia a tasas más elevadas de cánceres de intervalo con peor pronóstico cuando se detectan clínicamente ([2-4](#)). El tejido mamario denso es por sí mismo un factor de riesgo de cáncer de mama, que es 4–6 veces mayor en las mamas muy densas (densidad mamaria >75%) frente a las mamas grasas (densidad mamaria <10%) ([4-6](#)).

La densidad mamaria expresa la cantidad de tejido fibroglandular respecto a la grasa de la mama. Desde que en 1976 Wolfe ([7](#)) asoció la densidad mamaria al cáncer de mama y definió una escala de 4 grados para graduarla, se han utilizado diferentes clasificaciones, todas ellas basadas en criterios cualitativos ([8](#)) o cuantitativos ([9-11](#)), a partir del análisis visual o semiautomático de las mamografías. La 4ª edición de la escala BI-RADS® ([10](#)) del Colegio Americano de Radiología (ACR) ha estandarizado la densidad mamaria en la inspección visual utilizando una escala cuantitativa discreta que clasifica la densidad en cuatro categorías. Esta clasificación, y la de Boyd, que establece seis categorías cuantitativas ([11](#)), son las más extendidas y aceptadas en la actualidad. Existe evidencia de que la medida cuantitativa del porcentaje de área densa en la

mamografía predice mejor el riesgo que la clasificación en categorías cualitativas (5,8).

En la actualidad hay técnicas de automatización que mejoran el cálculo de la densidad mamaria y la clasificación de los patrones parenquimatosos, como, por ejemplo, las presentadas por Cumulus, de la Universidad de Toronto (3), o por Madena, de la Universidad de Southern California (13). Estas técnicas utilizan mamografías analógicas, que necesitan ser escaneadas y digitalizadas. Sin embargo, la mamografía digital es hoy una técnica ampliamente extendida y aceptada. El propósito de este estudio es desarrollar una aplicación informática que permita analizar y calcular de una manera objetiva y reproducible la densidad mamaria en las mamografías digitales directas, sin procesamiento posterior (14-18).

Material y métodos

1. Pacientes

Se utilizaron 655 mamografías digitales directas que habían sido utilizadas en un proyecto anterior “Determinants of Density in Mammography in Spain” (DDM-Spain) (proyecto FIS PI060386) que investigó la densidad mamaria como uno de los factores de riesgo para cáncer de mama en 3.584 mujeres procedentes de siete centros de cribado españoles. Para nuestro estudio, las mamografías fueron seleccionadas aleatoriamente de tres de los centros de cribado que disponían de mamógrafos digitales: Barcelona, dotada con un Hologic Lorad M-IV (Bedford, MA EEUU); Palma de Mallorca, con un Siemens Novation (Muenchen, Germany); y Valencia, dotado con un Senographe 2000D GE Medical System S.A. (Buc Cedex,

France). La edad de las mujeres estaba comprendida entre 45-69 años. Se excluyeron las diagnosticadas de cáncer de ovario o mama, y las que habían sufrido una intervención quirúrgica mamaria o eran portadora de prótesis. El comité de bioética del instituto de salud Carlos III (Madrid) dio su aprobación y se obtuvo el consentimiento informado.

2. Técnica de estudio.

En una primera fase, tres radiólogos con amplia experiencia en mamografía de cribado clasificaron visualmente la densidad mamaria de las 655 mamografías, según las escalas de Boyd y BI-RADS® ([fig.1](#)). Previamente habían realizado un entrenamiento conjunto con 300 imágenes digitales directas procedentes de programas de cribado diferentes aunque técnicamente comparables. Para el cálculo de la densidad mamaria se utilizó la proyección cráneo-caudal izquierda, que presenta menos problemas técnicos para segmentar la imagen porque incluye menos músculo pectoral que otras proyecciones. Dado que para clasificar la densidad mamaria no se necesita una gran resolución de imagen, para el análisis se utilizó un monitor convencional de 17" y un software visualizador de imágenes radiológicas no diagnóstico (K-PACS V1 6.0; software libre, <http://www.k-pacs.de>).

En una segunda fase, tres meses después, utilizamos una aplicación informática (DM-Scan), desarrollada específicamente para este estudio en la Universidad Politécnica de Valencia (España). Se trata de una herramienta orientada al diagnóstico asistido por ordenador (CAD) y su finalidad es calcular la densidad mamaria de modo semiautomático para reducir la subjetividad. El programa se diseñó para trabajar con imágenes en formato PNG, que utiliza un algoritmo de compresión que reduce el peso de la imagen sin pérdida de calidad. Por este

motivo, las imágenes originales en formato digital DICOM fueron convertidas a PNG mediante el software dicom 2, software libre (19). Al cargar la imagen mamográfica, DM-Scan la procesa primero para identificar la mama y segmentarla automáticamente (define el contorno de la mama y la aísla del resto de la imagen). Cuando la segmentación no es precisa, por ejemplo cuando el músculo pectoral u otros elementos de la imagen se incluyen en el área segmentada, es posible excluirlas. Con este proceso se obtiene el tamaño total de la mama medido en número de píxeles. A continuación, el radiólogo define un umbral de brillo para establecer la separación entre el tejido denso y el graso, lo que permite conocer el área total de tejido denso, también medida en número de píxeles, y, por tanto, su porcentaje exacto respecto al tamaño total de la mama.

Téngase en cuenta que la atenuación de los rayos X, y, por tanto, el nivel de brillo en la mamografía, depende no sólo de la densidad del tejido sino de su espesor. Esto significa que la parte de la mama más cercana al músculo pectoral en ocasiones es más brillante que la más periférica, por lo que la selección del tejido denso solo a partir de un umbral de brillo tiende a seleccionar regiones cercanas al músculo pectoral. Para corregir este efecto se aplica un filtro digital (“filtro de mama”) que estima el espesor de la mama en cada punto de la imagen y oscurece proporcionalmente el píxel correspondiente. Dado que el espesor real de cada mama no es conocido, la aplicación proporciona unos controles manuales que permiten modificar los parámetros que configuran este filtro. Concretamente, el brillo de cada píxel p_{ij} de la mama se multiplica por un coeficiente de corrección k_{ij} en función de un parámetro alfa que toma valores entre 0 y 1, definido por el usuario, tal que

$$k_{ij} = \alpha + (1 - \alpha) d_{ij}$$

donde d_{ij} representa la distancia relativa del píxel p_{ij} al borde de la mama. En la figura 2 se muestra la interfaz gráfica de la aplicación.

3. Análisis estadístico

Para estimar la concordancia intraobservador en las dos fases se eligieron al azar el 10% de las imágenes, que se analizaron dos veces con una separación de 2 meses entre ambas lecturas. En ambas fases también se calculó la concordancia interobservador entre pares de radiólogos, y posteriormente comparamos las medias de las concordancias de los métodos visual y semiautomático. Para los resultados dados como categorías (clasificación visual), la concordancia se analizó con el Índice de Kappa (k) con pesos cuadráticos e intervalos de confianza del 95% (IC95%). Para los resultados dados en escala continua (DM-Scan)) se utilizó el Índice de Correlación Intraclase (ICC). Para calcular el índice de kappa se utilizó una ecuación ponderada con pesos cuadráticos, de modo que los desacuerdos en categorías más distantes tienen mayor penalización. Dadas dos categorías i, j , el factor de ponderación $W_{i,j}$ utilizado fue:

$$W_{i,j} = 1 - \frac{|i - j|^2}{(N - 1)^2}$$

donde N representa el número de categorías. El índice kappa con pesos cuadráticos es estadísticamente comparable al ICC (20), lo que permite la comparación entre los dos métodos analizados.

El estudio estadístico se llevó a cabo con el software libre R (R Development Core Team 2011). R es un entorno de software gratuito para el análisis de resultados estadísticos. Para los cálculos de concordancia se hace uso del paquete irr (21).

Resultados

1. Clasificación mediante inspección visual

En el estudio de la concordancia interobservador la concordancia media fue de 0,876 (IC95% 0,873-0,879) para la escala de Boyd y 0,823 (IC95% 0,818-0,829) para BI-RADS®. La concordancia intraobservador media fue de 0,813 (IC95% 0,796-0,829) para la clasificación en la escala de Boyd, y 0,770 (IC95% 0,742-0,797) para la clasificación mediante BI-RADS®. Comparando ambas escalas se observa que la concordancia media es ligeramente superior con la escala de Boyd y llama la atención que la concordancia intraobservador sea inferior que la interobservador para ambas escalas.

2. Clasificación semiautomática

En el análisis realizado mediante DM-Scan la concordancia media, tanto interobservador como intraobservador, fue de ICC=0,92 (IC95% 0,916-0,928), notablemente mejor que la concordancia con respecto a la clasificación visual. En este caso, la concordancia intraobservador fue igual a la interobservador.

Discusión

En este estudio hemos mostrado que una aplicación informática que automatiza la determinación de la densidad mamaria consigue que concordancias intra e interobservador sean excelentes y mejoren las que se consiguen con la inspección visual.

En la actualidad, la densidad mamaria se estima visualmente, pero hay variabilidad entre radiólogos o en un mismo radiólogo. Las concordancias descritas en estudios previos van desde moderadas hasta muy buenas, por ejemplo $k=0,54$ en el estudio

de Ciatto et al. (22), $k=0,77$ para Ooms et al. (16), $k=0,85$ para Perez Gómez et al. (23) y k por encima de 0,90 para Garrido-Estepa (9), en el que un solo radiólogo clasificó el mismo grupo de mamografías en cuatro escalas. Esta variabilidad no tiene fácil explicación pero puede ser debida a que los métodos empleados eran distintos, porque los estudios difieren en el número de mamografías analizadas, número de radiólogos que participan (entre uno y doce), experiencia de los radiólogos, entrenamiento previo al estudio, tipo de mamografía (analógicas o digitales) y ausencia de parámetros estándar difíciles de definir ya que resulta difícil asignar una categoría cuantitativa a aquellas imágenes que se encuentran en la frontera entre dos categorías. Además, muchos de los trabajos no especifican con precisión el método estadístico empleado para calcular el índice kappa lo que dificulta aún más la comparación de los resultados descritos. Con todo esto, se puede deducir que la estimación visual de la densidad mamaria, sujeta a un elevado grado de subjetividad, no es sencilla. A pesar de la dificultad intrínseca de la tarea, nuestras concordancias visuales se incluyeron entre valores de k de 0,61 y 0,80 y, por lo tanto, son buenas (24).

Llama la atención que nuestra concordancia media intraobservador en el análisis visual, a diferencia de otros estudios (9,23), fue menor que la interobservador. Cabe señalar que ambos estudios no se han realizado sobre la misma muestra (en el estudio intraobservador se utilizó únicamente el 10% de las imágenes). Además, el estudio interobservador se realizó inmediatamente después del entrenamiento previo, mientras que el intraobservador se llevó a cabo pasados unos meses, lo que podría justificar un ajuste menor a los criterios de clasificación acordados durante el entrenamiento previo. Sin embargo, la diferencia desapareció cuando el

análisis se hizo con DM-Scan, que además de obtener índices de concordancia excelentes, no mostró discordancias en la variabilidad inter e intraobservador, independientemente del paso del tiempo. Por lo tanto, la estimación de la densidad mamaria puede ser más precisa y objetiva con aplicaciones informáticas que la determinen automática o semiautomáticamente.

Uno de los primeros en sugerir el uso del ordenador como herramienta de ayuda para determinar la densidad mamaria fue Boyd ([11](#)), que presentó un método semiautomático (Cumulus) (3) que estimaba la densidad de la mama basándose en la selección manual de dos umbrales para segmentar la mama y el tejido denso respectivamente. La aplicación DM-scan, a diferencia de otras aplicaciones (3,25), hace un cálculo inicial del porcentaje de tejido mamario que se puede aceptar o modificar, lo que agiliza y facilita su manejo, y disminuye la subjetividad de la manipulación manual de la medida. Además corrige las diferencias de brillo debidas al grosor de la mama y no a la densidad.

Entre las limitaciones del presente estudio cabe destacar que no hemos calculado el tiempo extra necesario para añadir el cálculo de la densidad mamaria en la rutina diaria, que otros trabajos previos han estimado entre 18 y 40 segundos por proyección (14,26) y 5-8 minutos por estudio (25). Al igual que otras herramientas semiautomáticas orientadas al cálculo de la densidad mamaria, si bien pueden mejorar los resultados de la inspección visual, no están exentas de cierto grado de subjetividad, ya que requiere que un operador establezca el umbral que delimita el tejido denso del graso, lo que puede introducir cierta variabilidad en el cálculo. Sin embargo, esta no es una limitación real teniendo en cuenta que nuestra pretensión fue estimarla también para la medida semiautomatizada.

Nuestros resultados, con un número de mamografías significativamente mayor que el resto de estudios publicados, confirman los resultados obtenidos en otros estudios realizados mediante clasificación visual (16,22), semiautomática (14) o cuando se comparan ambas clasificaciones visual y semiautomática (14,15,25,26). Aunque ambos métodos (inspección visual y DM Scan) son válidos para el cálculo de la densidad mamaria, DM-scan proporciona es más precisa, reduce significativamente la subjetividad, aumenta la fiabilidad para establecer la densidad mamaria y, consecuentemente, permite homogeneizar criterios que puedan ayudar a elaborar protocolos de cribado más adecuados según la densidad de la mama. En conclusión, nuestro estudio muestra que el cálculo de la densidad mamaria con la aplicación semiautomática DM Scan es fiable y reproducible y disminuye la subjetividad y variabilidad de la estimación visual.

Agradecimientos

Agradecemos al Centro Nacional de Salud: Instituto de Salud Carlos III de Madrid, a la Fundación Gent x Gent (EDEMAC Project), al Fondo Español de Investigación Sanitaria y a la Federación Española de Pacientes con Cáncer de Mama (FECMA), por su colaboración en la financiación del proyecto “Determinants of Density in Mammography in Spain” (DDM-Spain) de la que proceden las mamografías del presente trabajo.

También agradecemos la colaboración de otros miembros del DDM-Spain: Pilar Moreo, Pilar Moreno y Soledad Abad (Aragón); Francisca Collado y Magdalena Moyá (Baleares); Isabel González, Carmen Pedraz y Francisco Casanova (Castilla-León); Mercé Peris (Cataluña); Carmen Santamaría, José Antonio Vázquez

Carrete, Montserrat Corujo y Ana Belén Fernández (Galicia); Nieves Ascunce, María Ederra, Milagros García y Ana Barcos (Navarra); Manuela Alcaraz, Jesús Vioque (C. Valenciana); Virginia Lope, Nuria Aragonés, Anna Cabanes (Madrid).

Al Instituto Tecnológico de Informática de la Universidad Politécnica de Valencia la realización del DM-Scan para la lectura semiautomática y automática.

A las Dras. Beatriz Pérez-Gómez y Marina Pollán del área de Epidemiología Ambiental y del Cáncer, Centro Nacional de Epidemiología, Instituto de Salud Carlos III, y las Dras. Dolores Salas y Josefa Miranda de la Dirección General Salud Pública, Programa de Cribado de Cáncer de Mama, Centro Superior de Investigación Salud Pública (CSISP), Valencia, así como a la Dra. Carmen Palop (Unidad de Prevención de Cáncer de Mama de Sagunto, Valencia) por su estrecha colaboración y ayuda en la elaboración de este trabajo.

Bibliografía

1. Bartella L, Smith CS, Dershaw DD, Liberman L. Imaging Breast cancer. *Radiol Clin North Am.* 2007 ;45:45-67
2. Tabár L, Vitak B, Chen TH, Yen AM, Cohen A, Tot T, et al. Swedish Two-County trial: Impact of mammographic screening on breast cancer mortality during 3 decades. *Radiology.* 2011;260: 658-63.
3. Byng JW, Boyd NF, Fishell E, Jong RA, Yaffe MJ. The quantitative analysis of mammographic densities. *Phys Med Biol.* 1994;39:1629–38.
4. Boyd NF, Guo H, Martin L J, Sun L, Stone J, Fishell E, et al. Mammographic density and the risk and detection of breast cancer. *N Engl J Med.* 2007; 356:227-36.

5. McCormack 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.
6. Ursin G, Samera A, Qureshi S. Mammographic density – a useful biomarker for breast cancer risk in epidemiologic studies. *Nor Epidemiol.* 2009;19:59-68.
7. Wolfe JN. Breast patterns as an index of risk for developing breast cancer. *AJR Am J Roentgenol.* 1976;126:1130-7.
8. Gram IT, Bremnes Y, Ursin G, Maskarinec G, Bjurstam N, Lund E. Percentage density, Wolfe's and Tabar's mammographic patterns: agreement and association with risk factors for breast cancer. *Breast Cancer Res.* 2005;7:854-61.
9. Garrido-Esteba M, Ruiz-Perales F, Miranda J, Ascunce J, Gonzalez-Román I, Sanchez-Contador C, et al. DDM-Spain. Evaluation of mammographic density patterns: reproducibility and concordance among scales. *BMC Cancer.* 2010;10:485.
10. American College of Radiology Breast Imaging Reporting and Data System (BI-RADS®) 4th ed. Reston , VA: (American College of Radiology); 2003.
11. 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.
12. [Heine JJ](#), [Carston MJ](#), [Scott CG](#), [Brandt K](#), Wu F-F, Pankratz V, et al. An automated approach for estimation of breast density. [Cancer Epidemiol Biomarkers Prev.](#) 2008;17:3090-7.

13. Ursin G, Astrahan MA, Salane M, Parisky YR, Pearce JG, Daniels JR, et al. The detection of changes in mammographic densities. *Cancer Epidemiol Biomarkers Prev.* 1998;7:43-7
14. Tagliafico A, Tagliafico G, Tosto S, Chiesa F, Martinoli C, Derchi LE, et al. Mammographic density estimation: comparison among BI-RADS categories, a semi-automated software and a fully automated one. *Breast.* 2009;18:35–40.
15. Oliver A, Freixenet J, Marti R, Pont J, Pérez E, Denton ER, et al. A novel breast tissue density classification methodology. *IEEE Trans Inf Technol Biomed.* 2008;12:55-65.
16. Ooms EA, Zonderland H, Eijkemans M, Mahdavian Delavary B, Burger C, Ansink A, et al. Mammography: interobserver variability in breast density assessment. *Breast.* 2007;16:568-76.
17. Zhou C, Chan HP, Petrick N, Helvie MA, Goodsitt MM, Sahiner B, et al. Computerized image analysis: estimation of breast density in mammograms. *Med Phys.* 2001;28:1056–69.
18. Llobet R, Ruiz Perales, Solves JA, Perez-Cortés JC, Salas D, Pollán M. Estimación de la densidad mamaria asistida por ordenador. *Informática y Salud.* 2009; 84-9.
19. Barre S. Dicom2, a free medical image and dicom converter. URL <http://www.barre.nom.fr/medical/dicom2/>. 2007.
20. Fleiss JL, Cohen J: Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educ Psychol Meas.* 1973;33:613-9.

21. Gamer M, Lemon J, Ian Fellows Puspendra Singh. R: Various coefficients of inter-rater reliability and agreement. URL <http://rss.acs.unt.edu/Rdoc/library/irr/html/00Index.html>. 2010.
22. Ciatto S, Houssami N, Apruzzese A, Basetti E, Brancato B, Carozzi F, et al. Categorizing breast mammographic density: Intra and interobserver reproducibility of BI-RADS density categories. *Breast* 2005;14:269-75.
23. Pérez-Gómez B, Ruiz F, Martínez I, Casals M, Miranda J, Sánchez-Contador C, et al. Women's features and inter/intra rater agreement on mammographic density assessment in full-field digital mammograms (DDM-SPAIN). *Breast Cancer Res Treat.* 2012;132:287-95.
24. Altman DG. Practical statistics for medical research. Monographs on Statistics and Applied Probability (first ed.). New York: Chapman and Hall; 1991.
25. Lobbes MB, Cleutjens JP, Lima Passos V, Frotscher C, Lahaye MJ, Keymeulen KB, et al. Density is in the eye of the beholder: visual versus semi-automated assessment of breast density on standard mammograms. *Insights Imaging.* 2012;3:91-9
- 26.** Martin KE, Helvie MA, Zhou C, Roubidoux MA, Bailey JE, Paramagul C, et al. Mammographic density measured with quantitative computer-aided method: comparison with radiologists' estimates and BI-RADS® categories. *Radiology.* 2006;240:656-65