



**HOSPITAL CENTRAL
DR. IGNACIO
MORONES PRIETO**

UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE MEDICINA

Hospital Central “Dr. Ignacio Morones Prieto”

Trabajo de investigación para obtener el diploma en la especialidad de
Imagenología Diagnóstica y Terapéutica

**Concordancia inter e intraobservador para composición del tejido
mamario y categoría de valoración en la interpretación de
mastografías de tamizaje utilizando el sistema BI-RADS 5ª edición**

Dr. Daniel Sandoval Guerra

DIRECTOR CLÍNICO
Dr. Francisco Javier Marín Solares
Médico Radiólogo

DIRECTOR METODOLÓGICO
M. C. Mario Aurelio Martínez Jiménez
Cirujano General y Maestro en Ciencias en Investigación Clínica

Enero 2022



**HOSPITAL CENTRAL
DR. IGNACIO
MORONES PRIETO**

UNIVERSIDAD AUTÓNOMA DE SAN LUIS POTOSÍ

FACULTAD DE MEDICINA

Hospital Central “Dr. Ignacio Morones Prieto”

Trabajo de investigación para obtener el diploma en la especialidad de
Imagenología Diagnóstica y Terapéutica

**Concordancia inter e intraobservador para composición del tejido
mamario y categoría de valoración en la interpretación de
mastografías de tamizaje utilizando el sistema BI-RADS 5ª edición**

Dr. Daniel Sandoval Guerra

No. de CVU CONACYT: 1232624; ORCID: 0000-0001-5184-5255

DIRECTOR CLÍNICO

Dr. Francisco Javier Marín Solares

Médico Radiólogo

No. de CVU CONACYT: 1186447; ORCID: 0000-0002-8572-8230

DIRECTOR METODOLÓGICO

M. C. Mario Aurelio Martínez Jiménez

Cirujano General y Maestro en Ciencias en Investigación Clínica.

No. de CVU CONACYT: 302613; ORCID: 0000-0002-8905-849X

SINODALES

Dr. Raúl Martínez Martínez

Médico Radiólogo

Alta Especialidad en Radiología Intervencionista

Presidente

Dr. Alejandro Hernández Martínez

Médico Radiólogo

Alta Especialidad en Radiología Intervencionista

Sinodal

Dr. Carlos Lambert Cerda

Médico Radiólogo

Sinodal

Enero, 2022



This work © 2022 by Daniel Sandoval Guerra is licensed under Attribution-NonCommercial-NoDerivatives 4.0 International. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Resumen.

Objetivo: Determinar el grado de concordancia intra e inter observadores en la categorización de composición del tejido mamario y valoración definitiva en mastografías de tamizaje utilizando la quinta edición del sistema BI-RADS del Colegio Americano de Radiología.

Sujetos y Métodos: Utilizando tablas previamente desarrolladas para el cálculo del tamaño de muestra en estudios de concordancia y con base en los resultados recabados de estudios previos, se determinó un tamaño de muestra de 100 estudios, se seleccionaron aleatoriamente de una base de mastografías de tamizaje, aquellos que cumplieran con los estándares de calidad. Seis participantes (3 médicos radiólogos y 3 residentes de último año) evaluaron cada una de las mastografías y asignaron una categoría de composición del tejido y valoración final en dos ocasiones con un lapso de 45 días entre cada evaluación. Con los datos obtenidos se calculó la concordancia inter e intra observadores utilizando el índice de Kappa.

Resultados: En la primera ronda de evaluación, la concordancia interobservador entre los médicos radiólogos fue débil a buena (K 0.36 – 0.65) para la composición del tejido mamario y moderada a buena (K 0.50 – 0.61) para categoría de valoración final, la concordancia interobservador entre los médicos residentes fue moderada (K 0.50 – 0.52) para composición del tejido mamario e igualmente moderada (K 0.44 – 0.53) para categoría final. Durante la segunda ronda, la concordancia interobservador entre los radiólogos fue moderada a buena (K 0.43 – 0.71) para densidad y moderada a buena (K 0.45 – 0.65) para categoría final, la concordancia entre los médicos residentes fue moderada (K 0.41 – 0.63) para composición del tejido y débil a moderada (K 0.40 – 0.54) para categoría final. La concordancia intraobservador para densidad del tejido fue buena (K 0.69 – 0.76) en los radiólogos y moderada a buena (K 0.48 – 0.68) en los residentes, para categoría final la concordancia fue moderada a buena (K 0.60 – 0.69) en los radiólogos y moderada a buena (K 0.42 – 0.70) en los residentes.

Conclusiones: En este estudio, hemos encontrado una concordancia inter e intraobservador de moderada a buena entre los médicos radiólogos y residentes de último año, según los criterios establecidos en la 5ª edición del BI-RADS.

Palabras clave: Mamografía, tamizaje, variabilidad interobservador, variabilidad intraobservador, densidad mamaria, BI-RADS 5ª edición.

Índice.	
Resumen.....	1
Lista de Cuadros.	5
Lista de Figuras.....	6
Lista de Abreviaturas y Definiciones.....	7
Agradecimientos.	8
Dedicatorias.....	8
Antecedentes.....	9
Justificación.	15
Pregunta de investigación.....	16
Hipótesis.....	16
Objetivos.....	16
Objetivo general.....	16
Objetivos específicos.	16
Objetivos secundarios.....	17
Sujetos y Métodos.....	17
Diseño del estudio.	17
Lugar de realización.....	17
Universo de estudio.	17

Criterios de selección.....	17
Análisis Estadístico.....	18
Variables en el estudio.....	18
Tipo de muestreo.....	19
Cálculo del tamaño de la muestra.....	19
Aleatorización.....	23
Metodología.....	23
Plan de trabajo.....	26
Aspectos éticos y recursos.....	26
Recursos humanos y materiales.....	26
Capacitación de personal.....	27
Financiamiento.....	27
Resultados.....	27
Discusión.....	35
Limitaciones y/o nuevas perspectivas de investigación.....	37
Conclusiones.....	38
Referencias bibliográficas.....	39
Anexos.....	42
Anexo 1. Cronograma de actividades.....	42
Anexo 2. Formato de interpretación.....	43
Anexo 3. Dictamen comité de ética.....	¡Error! Marcador no definido.
Anexo 4. Dictamen comité de investigación.....	¡Error! Marcador no definido.

Lista de Cuadros.

Cuadro 1. Indicaciones de mastografía GPC S-001-08.	10
Cuadro 2. Cuadro de variables.....	19
Cuadro 3. Revisión bibliográfica de estudios de concordancia.	20
Cuadro 4. Cálculo del tamaño de la muestra para estudios de concordancia.....	22
Cuadro 5. Tabla de contingencia para composición del tejido mamario.....	25
Cuadro 6. Tabla de contingencia para categoría de valoración final.....	25
Cuadro 7. Valores de Kappa y su concordancia.	26
Cuadro 8. Concordancia interobservador en la primera ronda.....	28
Cuadro 9. Concordancia interobservador en la segunda ronda.	29
Cuadro 10. Concordancia intraobservador.....	30
Cuadro 11. Sumatoria de concordancia y discordancias para densidad mamaria.	31
Cuadro 12. Sumatoria de concordancias y discordancias para categoría final.	32
Cuadro 13. Media de resultados de todos los observadores.....	32

Lista de Figuras.

Figura 1. Resumen de las características generales del estudio.	24
Figura 2. Distribución de la densidad mamaria en la primera evaluación.....	33
Figura 3. Distribución de la densidad mamaria en la segunda evaluación.	33
Figura 4. Distribución de las categorías finales en la primera evaluación.	34
Figura 5. Distribución de las categorías finales en la segunda evaluación.....	34

Lista de Abreviaturas y Definiciones.

- **CM:** Cáncer mamario.
- **GPC:** Guía de Práctica Clínica.
- **ACR:** Colegio Americano de Radiología.
- **IC:** Intervalo de confianza.
- **BI-RADS:** Sistema de Informes y Registro de Datos de Estudios por Imágenes de la Mama. Método utilizado por los radiólogos para interpretar y comunicar de manera estandarizada los resultados de exámenes por imagen de la mama.
- **Concordancia:** Grado de correspondencia entre los valores de una variable proporcionados por dos métodos de medida u observadores diferentes, o bien entre medidas repetidas de la misma variable.
- **Mastografía/Mamografía:** Estudio radiológico de las mama, tomado con un aparato (mastógrafo) diseñado especialmente para este fin, con el que se podrán efectuar mamografías de pesquisa (tamizaje) y de diagnóstico.
- **Composición del Tejido Mamario:** Estimación visual del volumen de tejido fibroglandular denso que contiene la mama.
- **Categoría de Valoración:** Categoría definitiva que evalúa la impresión mamográfica general.
- **ACR a:** Mamas compuestas por tejido adiposo casi en su totalidad.
- **ACR b:** Sectores dispersos de densidad fibroglandular.
- **ACR c:** Mamas heterogéneamente densas, que pueden ocultar nódulos.
- **ACR d:** Mamas muy densas, lo que disminuye la sensibilidad mamográfica.
- **Categoría 0:** Estudio incompleto.
- **Categoría 1:** Resultado negativo, probabilidad de malignidad del 0%.
- **Categoría 2:** Hallazgos típicamente benigno, probabilidad de malignidad del 0%.
- **Categoría 3:** Hallazgo probablemente benigno, probabilidad de malignidad $\leq 2\%$.
- **Categoría 4:** Hallazgo sospechoso de malignidad, probabilidad de malignidad del 2 al 94%.
- **Categoría 5:** Hallazgo muy sugerente de malignidad, probabilidad de malignidad $\geq 95\%$.
- **Categoría 6:** Diagnóstico maligno confirmado mediante biopsia.

Agradecimientos.

Agradezco a todos los que contribuyeron en la realización de este proyecto, sobre todo a mi director de tesis, Dr. Francisco Javier Marín por estar siempre dispuesto y disponible.

A todos los que participaron como observadores, sé que el tiempo es algo muy valioso, agradezco profundamente el que invirtieron en apoyarme.

A mis maestros, que me inspiraron, enseñaron y respaldaron durante mi formación como radiólogo, especialmente al Dr. Marco Rivera y al Dr. Jorge Rodríguez por compartir su conocimiento conmigo pero sobre todo por brindarme su amistad, a ustedes les estoy eternamente agradecido.

Dedicatorias.

A mi familia, principalmente a mis padres, lo que he alcanzado como persona, como estudiante, como profesional, es gracias a ustedes.

A mis compañeros, Jackie y Gudi, a pesar de las adversidades, trabajando juntos, lo logramos con creces.

A todas las personas que, como pacientes, me ayudaron a aprender no sólo medicina sino también de la vida, a todos ellos, espero haberles podido ayudar a mantener o recuperar su salud.

Antecedentes.

El cáncer mamario (CM) constituye la principal causa de morbilidad y mortalidad por tumores malignos en las mujeres mayores de 30 años de edad en México y el resto del mundo.(1) Entre las mujeres representa 1 de cada 4 casos de cáncer y 1 de cada 6 muertes por cáncer.(1)

En México, durante 2020 se registró una tasa de incidencia de 23.05 casos de CM por cada 100 000 mujeres mayores de 10 años.(2) Particularmente a nivel local en el estado de San Luis Potosí la tasa fue de 16.19 casos por cada 100 000 mujeres mayores de 10 años durante ese mismo año. Existen incidencias más altas en las mujeres de los grupos de edad por encima de 45 años con incidencias reportadas de hasta 78.1 casos por cada 100 000 a nivel nacional y 41.91 a nivel local en el grupo de 60 a 64 años.(3)

Se ha demostrado, que la única herramienta de tamizaje efectiva para reducir la mortalidad por CM, es la mastografía. La mastografía de tamizaje es aquella que se realiza en toda mujer que se encuentra asintomática, consiste en adquirir dos proyecciones de cada mama, una oblicua mediolateral y una craneocaudal, con el objetivo de identificar al pequeño subgrupo de mujeres que requieren realizarse estudios de imagen complementarios o alguna otra intervención, como una biopsia, del amplio grupo de mujeres a quienes se recomienda solamente el tamizaje periódico. Lo anterior tiene como finalidad principal la detección temprana del CM cuando aún es asintomático y reducir la mortalidad asociada a este. (4)

Las indicaciones para la realización de una mastografía de tamizaje son variables alrededor del mundo, en México se encuentran establecidas en la Guía de Práctica Clínica (GPC) “Prevención, Tamizaje y Referencia Oportuna de Casos Sospechosos de Cáncer de Mama en el Primer Nivel de Atención”,(5) mostradas en el **Cuadro 1**.

Cuadro 1. Indicaciones de mastografía GPC S-001-08.

Mujeres asintomáticas de 40 a 49 años con riesgo medio: Mastografía anual.

Mujeres asintomáticas de 50 a 74 años: Mastografía cada 1 a 2 años.

Mujeres con riesgo alto de cáncer de mama: Mastografía anualmente a partir de los 30 años de edad pero no antes de los 25 años de edad.

Mujeres con hermanas o madres con cáncer de mama premenopáusico:
Mastografía anualmente a partir de los 30 años pero no antes de los 25 años de edad o 10 años antes de la edad de diagnóstico del familiar afectado más joven (lo que resulte más tardío)

Mujeres con antecedentes de radiación de tórax recibida entre los 10 y 30 años:
Mastografía inicial 8 años después de la radioterapia pero nunca antes de los 25 años de edad.

La precisión de la mastografía de tamizaje para la detección temprana de CM depende de varios factores como son los protocolos para la lectura del estudio, las características de la mama y la experiencia de los médicos radiólogos.

Existen estudios realizados en países desarrollados que han demostrado el impacto positivo de un programa de tamizaje organizado en la reducción de la mortalidad por CM (6–9), sin embargo, en países en vías de desarrollo, como México, el impacto de estos programas en la mortalidad por CM ha sido bajo debido a diversos factores, entre ellos, la carencia de sistemas de control de calidad y de capacitación del personal. (10–12)

En nuestro país, se ha visto que hay gran disparidad en la adherencia a las normas establecidas por instancias internacionales con el Colegio Americano de Radiología (ACR) y la Comunidad Europea. Dentro de las incongruencias más importantes destacan la obtención y procesamiento de la imagen, el acondicionamiento de la sala de interpretación y la concordancia entre los reportes radiológicos. (13,14)

En un esfuerzo para mejorar la precisión, disminuir la discordancia y estandarizar la evaluación e interpretación de estudios por imagen de las mamas, el Colegio Americano de Radiología, desarrolló el Sistema de informes y registro de datos de estudios por imágenes de la mama (BI-RADS) con su primera edición publicada en 1993 y su quinta y más reciente edición publicada en 2013.(15) Este sistema es un instrumento que propone un método normalizado para informar los estudios por imágenes de mama y que posee beneficios y limitaciones.

Por el grado de aceptación y validación que ha tenido a través del tiempo, se ha adoptado internacionalmente como una garantía de la calidad para normalizar los informes de estudios por imagen de la mama, disminuir la discordancia en la interpretación y recomendaciones y para facilitar la vigilancia de los resultados. En nuestro país actualmente es el sistema que se recomienda utilizar en la GPC mexicana. (5)

En su contenido, la 5ª edición del BI-RADS define cuatro categorías en las que se clasifica la composición del tejido mamario, definidas conforme a la estimación visual del tejido fibroglandular denso que contiene la mama, organizadas de la siguiente forma: ACR a (Mamas compuestas por tejido adiposo casi en su totalidad), ACR b (sectores dispersos de densidad fibroglandular), ACR c (mamas heterogéneamente densas, que pueden ocultar nódulos pequeños) y ACR d (mamas muy densas, lo que disminuye la sensibilidad mamográfica). Para categorías definitivas de valoración que evalúen la impresión mamográfica general, la 5ª edición establece siete categorías divididas en dos grupos, evaluación incompleta (categoría 0) y evaluación definitiva (categorías 1, 2, 3, 4, 5 y 6) que se ordenan jerárquicamente por el grado de sospecha de malignidad de la siguiente forma: categoría 1 (resultado negativo) con probabilidad de malignidad del 0%, categoría 2 (hallazgo típicamente benigno) con probabilidad de malignidad del 0%, categoría 3 (hallazgo probablemente benigno) con probabilidad de malignidad $\leq 2\%$, categoría 6 (diagnóstico maligno confirmado mediante biopsia), categoría 0 (estudio incompleto), categoría 4A (hallazgo sospechoso con escasa presunción de malignidad) con probabilidad de malignidad $>2\%$ pero $\leq 10\%$, categoría 4B (hallazgo sospechoso con presunción moderada de malignidad) con probabilidad

de malignidad $>10\%$ pero $\leq 50\%$, categoría 4C (hallazgo sospechoso con gran presunción de malignidad) con probabilidad de malignidad $>50\%$ pero $<95\%$) y categoría 5 (hallazgo muy sugerente de malignidad) con probabilidad de malignidad $\geq 95\%$.(4)

Los resultados otorgados por los médicos radiológicos en la interpretación de una mastografía de tamizaje desempeñan un papel clave, ya que estos dictaminan las recomendaciones a seguir y es el último filtro en el proceso de tamizaje que permitirá que se logre o no la detección temprana de CM, por lo tanto se debe de buscar mantener una certeza diagnóstica y fiabilidad lo más alta posibles.

La fiabilidad de un instrumento de medida no es una propiedad intrínseca de este, sino más bien de cuando se aplica en una muestra de individuos concreta en unas condiciones dadas, la medición de un fenómeno clínico está sujeta a la variación biológica, individual, del observador o del instrumento utilizado. La fiabilidad se evalúa repitiendo el proceso de medición, con el fin de analizar la concordancia entre las distintas medidas, cuanto menor variabilidad se obtenga, más fiable es el proceso. (16)

Debido a que el sistema BI-RADS es observador dependiente es extremadamente importante que exista una buena concordancia entre radiólogos que interpretan mastografías de tamizaje, dadas las implicaciones para la salud y médico-legales del reporte final.

Partiendo de la premisa que el sistema BI-RADS es una herramienta validada por sus elaboradores y a través del tiempo por evaluadores externos, en su utilización como instrumento para la evaluación de mastografías de tamizaje, conviene medir la variabilidad entre los observadores, en este caso los radiólogos, para valorar su desempeño utilizando el instrumento y como una medida indirecta de la precisión que nos permita identificar áreas de mejora. (17)

Idealmente las tasas de variabilidad inter e intraobservador en estas importantes evaluaciones deberían de ser bajas, sin embargo, en la práctica clínica diaria de nuestro hospital las mamografías son interpretadas por radiólogos y residentes con diferentes niveles de experiencia.

Un número reducido de estudios previos han reportado variabilidad sustancial entre radiólogos para composición del tejido mamario y un número aún más limitado de investigaciones han analizado la concordancia para categorías definitivas utilizando el sistema BI-RADS 5ª edición encontrando de igual forma alta variabilidad.

La última actualización del sistema BI-RADS se realizó en 2013 por lo que se decidió filtrar la búsqueda bibliográfica por tiempo incluyendo los estudios realizados de 2013 a la actualidad, utilizando los metabuscadores PubMed y Biblioteca Virtual en Salud (BVS), para la búsqueda en PubMed se usaron los Encabezados de Temas Médicos (MeSH) "Mammography" "Mass Screening" "Early Detection of Cancer" "Observer Variation" y "Breast Density" y para la búsqueda en la BVS los Descriptores en Ciencias de la Salud (DeCS) "Mamografía" "Tamizaje Masivo" "Detección Precoz del Cáncer" "Variaciones Dependientes del Observador" y "Densidad de la Mama".

A continuación se expone una breve revisión de los artículos de investigación más relevantes encontrados en esta búsqueda:

- Masroor et. al., en el año 2016 realizaron un estudio transversal para evaluar la variabilidad inter e intraobservador para densidad mamaria y categoría de valoración BI-RADS. El tamaño de la muestra fue de 254 mastografías de tamizaje de mujeres de 35 a 65 años, evaluadas por tres radiólogos cegados en dos ocasiones con 3 meses de diferencia. Para evaluar la concordancia se utilizó el índice de Kappa obteniendo concordancia interobservador e intraobservador moderada para densidad mamaria y para categoría de valoración BI-RADS concordancia interobservador débil e intraobservador moderada. (18)

- Ekpo et. al., realizaron en el año 2016 un estudio transversal para evaluar la concordancia interradiólogos para clasificación de densidad mamario utilizando el BI-RADS 5ª ed. El tamaño de la muestra fue de 1000 mastografías de mujeres de 26 a 89 años con y sin síntomas, evaluadas por 5 radiólogos cegados en una ocasión y posteriormente de esa misma muestra se seleccionaron 100 para una segunda evaluación en un mismo escenario. Para evaluar la concordancia se utilizó el índice de Kappa ponderado obteniendo buena concordancia interobservador y muy buena intraobservador. (19)
- Ishrad et. al. , realizaron en el año 2016 un estudio transversal para evaluar la diferencia en la concordancia intra e intraobservador para densidad mamaria utilizando el BI-RADS 4ª ed. y 5ª ed. El tamaño de la muestra fue de 104 mastografías de tamizaje de mujeres elegidas aleatoriamente sin límites de edad, la interpretación fue realizada por 5 radiólogos con subespecialidad en mama con experiencia mayor a 3 años cegados a la información del paciente y el resultado de los otros radiólogos, se realizaron dos evaluaciones de las mismas mastografías con 4 semanas de separación. El análisis estadístico se realizó utilizando el coeficiente de Kappa ponderado de Fleiss-Cohen obteniendo una concordancia intraobservador buena e interobservador moderada.(20)
- Alomaim et. al., realizaron en el año 2019 un estudio transversal interinstitucional para evaluar la variabilidad en la clasificación de la densidad mamaria entre radiólogos estadounidenses y británicos. El tamaño de la muestra fue de 250 mastografías obtenidas de 18 centros diferentes de las cuales 180 fueron escogidas aleatoriamente y 70 fueron escogidas de casos que se enviaron previamente para repetir. 49 radiólogos participaron de los cuales 25 era estadounidenses y 24 británicos todos con subespecialidad en mama. Para el análisis estadístico se utilizó Kappa ponderada encontrando concordancia interobservador buena. (21)

- Pesce et. al., realizaron en el año 2020 un estudio transversal descriptivo para evaluar la variabilidad inter e intraobservador en la determinación de la densidad mamaria según la 5ª edición del Atlas BI-RADS. El tamaño de la muestra fue de 451 mastografías de tamizaje de mujeres de 40 a 90 años seleccionadas aleatoriamente. Los estudios fueron evaluados por 6 radiólogos cegados a la información del paciente y el resultado de los otros radiólogos, se realizaron dos lecturas de las mismas mastografías con un mes de separación. El análisis estadístico se realizó utilizando el coeficiente ponderado de Kappa obteniendo concordancia interobservador e intraobservador buena. (22)

Este estudio se diseñó para determinar el grado de concordancia inter e intraobservador en la categorización de la composición del tejido mamario y la categoría definitiva utilizando la 5ª edición del sistema BI-RADS de la ACR entre un grupo de radiólogos y residentes del servicio de imagenología de nuestra institución.

Justificación.

El objetivo de un programa de detección temprana de cáncer de mama mediante mastografía es reducir la mortalidad por esta patología. En comparación con los países desarrollados, en países en vía de desarrollo, como el nuestro, el impacto de estos programas en la reducción de la mortalidad por CM ha sido bajo debido, entre otros factores, a:

- Baja cobertura.
- Deficiente calidad en el proceso de atención.
- Falta de homologación de sistemas de control de calidad y de capacitación continua del personal.

Si bien son muchas las características que deben de ser evaluadas con el propósito de mejorar la calidad del servicio y su eficiencia, en este estudio nos interesa evaluar específicamente la concordancia en la interpretación entre los médicos radiólogos y residentes de último año, para poder comparar los resultados obtenidos con los de otras publicaciones internacionales y así valorar el desempeño de nuestro departamento

en la calidad de la interpretación mamográfica y determinar la necesidad de implementar programas de capacitación y certificación.

Actualmente no existe un estudio que aborde esta cuestión a nivel local, por lo cual, contar con un protocolo que lo haga sería de gran valor y sentaría bases para futuras investigaciones e intervenciones.

Pregunta de investigación.

¿Qué concordancia inter e intraobservador existe entre los médicos radiólogos y residentes del departamento de imagenología del Hospital Central en la asignación de categoría de composición del tejido mamario y categoría definitivas al interpretar mastografías de tamizaje de acuerdo con el sistema BI-RADS 5ª edición?

Hipótesis.

Existe concordancia moderada entre los médicos radiólogos y residentes para composición del tejido mamario y débil para categoría definitiva.

Objetivos.

Objetivo general.

Evaluar el grado concordancia inter e intraobservador existente entre los médicos radiólogos adscritos y residentes del departamento de imagenología en la asignación de categoría de composición del tejido mamario y categoría definitiva al interpretar mastografías de tamizaje de acuerdo con el sistema BI-RADS 5ª edición.

Objetivos específicos.

- Determinar la concordancia entre médicos radiólogos adscritos.
- Establecer la concordancia entre residentes de último año.
- Calcular la concordancia interobservador entre médicos radiólogos y residentes.
- Evaluar las medidas anteriores 45 días después.
- Obtener la concordancia intraobservador para cada médico radiólogo y residente.

Objetivos secundarios.

- Registrar la prevalencia de mamas con alta densidad (categorías c y d) en la muestra.
- Determinar la tasa promedio de rellamado del departamento (estudios categoría 0, 3, 4 y 5).

Sujetos y Métodos.

Diseño del estudio.

Estudio descriptivo transversal

Lugar de realización.

Departamento de imagenología Hospital Central “Dr. Ignacio Morones Prieto”, San Luis Potosí, S. L. P.

Universo de estudio.

- Mastografías de tamizaje realizadas entre 2021 y 2022 a mujeres mayores de 40 años seleccionadas aleatoriamente.
- Cada estudio mamográfico consiste de cuatro placas, una medio-lateral oblicua y una cráneo-caudal para cada mama.
- Tres Médicos radiólogos adscritos al departamento de imagen con más de 5 años de experiencia y tres residentes de imagenología del Hospital Central “Dr. Ignacio Morones Prieto”.

Criterios de selección.

- Inclusión:
 - Mujeres mayores de 40 años.
 - Estudio mamográfico de tamizaje completo (proyecciones craneocaudal y mediolateral oblicua de cada mama)

- Médicos radiólogos adscritos al departamento de imagen con más de 5 años de experiencia.
- Residentes de tercer año del departamento de imagenología.
- Exclusión:
 - Mastografías que no cumplan con los criterios de calidad.
 - Mastografías de mujeres sintomáticas, con antecedente de cáncer de mama, cirugía o implantes

Análisis Estadístico.

Variables en el estudio.

El diseño principal del estudio corresponde con uno de concordancia que no es sinónimo de correlación. La correlación se refiere a la presencia de relación entre dos variables diferentes (independiente y dependiente), mientras que la concordancia busca el acuerdo entre dos o más medidas de una variable por lo que no existen variables dependientes e independientes. **El Cuadro 2.** Muestra las variables utilizadas y su descripción.

Cuadro 2. Cuadro de variables.

Variable	Definición operacional	Valores posibles	Unidades	Tipo de variable
Composición del tejido mamario	Estimación visual del volumen de tejido fibroglandular denso que contiene la mama	a. Mamas compuestas por tejido adiposo casi en su totalidad b. Sectores dispersos de tejido fibroglandular c. Mamas heterogéneamente densas d. Mamas muy densas	N/A	Cualitativa ordinal
Categoría de Valoración	Categoría definitiva que evalúa la impresión mamográfica general	0: Estudio incompleto 1: Resultado negativo 2: Hallazgo típicamente benigno 3: Hallazgo probablemente benigno *4: Hallazgo sospechoso de malignidad 5: Hallazgo muy sugerente de malignidad	N/A	Cualitativa ordinal

*Con fines de análisis estadístico se optó por no utilizar la subdivisión de la categoría 4.

Tipo de muestreo.

No probabilístico.

Cálculo del tamaño de la muestra.

Se realizó una revisión de la bibliografía existente relacionada con el tema de investigación, los grados de concordancia encontrados para las variables en estudio se resumen en el **Cuadro 3**.

Cuadro 3. Revisión bibliográfica de estudios de concordancia.

Título del Estudio	Variable	Concordancia Encontrada (κ)	
		Interobservador	Intraobservador
To asses inter- and intra-observer variability for breast density and BIRADS assessment categories in mammographic reporting(18)	Composición Categoría de valoración	Moderada $\kappa = 0.48$ Débil K = 0.26	Moderada $\kappa = 0.57$ Moderado K = 0.41
Assessment of Interradiologist Agreement Regarding Mammographic Breast Density Classification Using the Fifth Edition of the BI-RADS Atlas(19)	Composición	Buena K = 0.79	Muy Buena 0.87
Comparison of variability in breast density assessment by BI-RADS category according to the level of experience(23)	Composición	Buena K = 0.67	Muy Buena 0.85
Consistency of breast density categories in serial screening mammograms: A comparison between automated and human assessment(24)	Composición	Buena K = 0.76	
Effects of Changes in BI-RADS Density Assessment Guidelines (Fourth Versus Fifth Edition) on Breast Density Assessment: Intra- and Interreader Agreements and Density Distribution(20)	Composición	Moderada K = 0.57	Buena K = 0.77
Interobserver and intraobserver variability in determining breast density according to the fifth edition of the BI-RADS® Atlas(22)	Composición	Buena K = 0.74	Buena K = 0.8
Variability of Breast Density Classification Between US and UK Radiologists(21)	Composición Categoría de Valoración	Buena K = 0.76 Moderada K = 0.56	Muy Buena K = 0.95

Con base en estos resultados se determinó que para la variable composición del tejido mamario existe concordancia interobservador moderada (K = 0.41 – 0.60) a buena (K = 0.61 – 0.80) e intraobservador moderada a muy buena (K >0.80). Mientras que para la variable categoría de valoración existe concordancia interobservador débil (K = 0.21 – 0.40) a moderada e intraobservador moderada.

Para tamaño mínimo de la muestra, se utilizaron las tablas desarrolladas por Bujang y Baharum(25) utilizando y la fórmula de Flack y colaboradores, **Cuadro 4.** (26) Para la variable composición del tejido, se empleó la tabla con 4 categorías y la tabla con 6 categorías para la variable categoría de valoración final, con un nivel de significancia (α) de 0.05, potencia estadística de 90%, concordancia inicial igual a la hipótesis nula ($K1 = 0$) y concordancia esperada ($K2$) alrededor de los valores de la hipótesis alternativa para composición mamaria ($K2 = 0.5$) y para categoría de valoración ($K2 = 0.3$).

Utilizando estos valores y tablas se obtuvo un tamaño mínimo de muestra de 32 para la variable composición del tejido y de 64 para la variable categoría de valoración. Al representar estos números un mínimo, se decidió aumentar el tamaño de la muestra a 100 para mejorar la calidad del estudio.

Se seleccionaron aleatoriamente 100 estudios mamográficos de tamizaje provenientes de diferentes centros pertenecientes a los Servicios de Salud del Estado.

Cuadro 4. Cálculo del tamaño de la muestra para estudios de concordancia.

TABLE 1. Sample size calculation for kappa at category 2 until 7 when proportion each category is assumed at proportionate

Category	K1	K2	n ^a	n ^b	Category	K1	K2	n ^a	n ^b	Category	K1	K2	n ^a	n ^b			
2x2	0.0	0.2	194	259	4x4	0.0	0.2	71	97	6x6	0.0	0.2	49	68			
		0.3	85	113			0.3	32	44			0.3	23	32	X2 = 64		
		0.4	47	62			0.4	18	25			0.4	13	18			
		0.5	29	38			0.5	12	16			0.5	8	12	X2 = 32		
		0.6	20	25			0.6	8	11			0.6	6	8			
		0.7	14	17			0.7	6	7			0.7	4	6			
		0.8	10	12			0.8	4	5			0.8	3	4			
		0.9	7	8			0.9	3	4			0.9	2	3			
		0.3	0.4	698			927	0.3	0.4			348	465	0.3	0.4	283	380
		0.5	0.6	563			742	0.5	0.6			317	420	0.5	0.6	271	359
0.7	0.8	363	471	0.7	0.8	223	290	0.7	0.8	196	255						
0.9	0.9	79	96	0.9	0.9	49	60	0.9	0.9	43	53						
3x3	0.0	0.2	117	157	5x5	0.0	0.2	56	77	7x7	0.0	0.2	46	64			
		0.3	52	70			0.3	26	36			0.3	21	30			
		0.4	29	39			0.4	15	20			0.4	12	17			
		0.5	18	24			0.5	9	13			0.5	8	11			
		0.6	12	16			0.6	7	9			0.6	6	8			
		0.7	9	11			0.7	5	6			0.7	4	5			
		0.8	6	8			0.8	3	4			0.8	3	4			
		0.9	5	5			0.9	2	3			0.9	2	3			
		0.3	0.4	469			624	0.3	0.4			304	407	0.3	0.4	272	364
		0.5	0.6	397			525	0.5	0.6			286	380	0.5	0.6	261	347
0.7	0.8	266	345	0.7	0.8	206	267	0.7	0.8	190	247						
0.9	0.9	58	71	0.9	0.9	45	55	0.9	0.9	41	51						

n^a Result derived based on power at 80.0% and alpha of 0.05

n^b Result derived based on power at 90.0% and alpha of 0.05

Aleatorización.

La muestra de 100 mastografías se eligió de una base de mastografías de tamizaje realizadas entre 2021 y 2002 utilizando aleatorización simple con apoyo de un equipo de cómputo.

Metodología.

El estudio se conducirá de la siguientes forma: en la primera vuelta, los 6 participantes (3 médicos radiólogos adscritos y 3 residentes de último año) interpretarán 100 estudios mamográficos a una tasa de 20 estudios por día para evitar sesgo por sobrecarga y fatiga, se les otorgará a cada uno un formato de llenado donde anotarán una categoría para composición del tejido mamario y otra para categoría de valoración definitiva, utilizando el sistema BI-RADS 5ª edición para cada estudio. Después de 45 días volverán a evaluar los mismos estudios que serán aleatorizados en orden diferente pero con un escenario igual. Los seis observadores estarán cegados de la identidad de las pacientes y de los resultados de los demás observadores, solamente se les proporcionará la edad de las pacientes.

El diseño del estudio se ilustra resumido en la **Figura 1**.

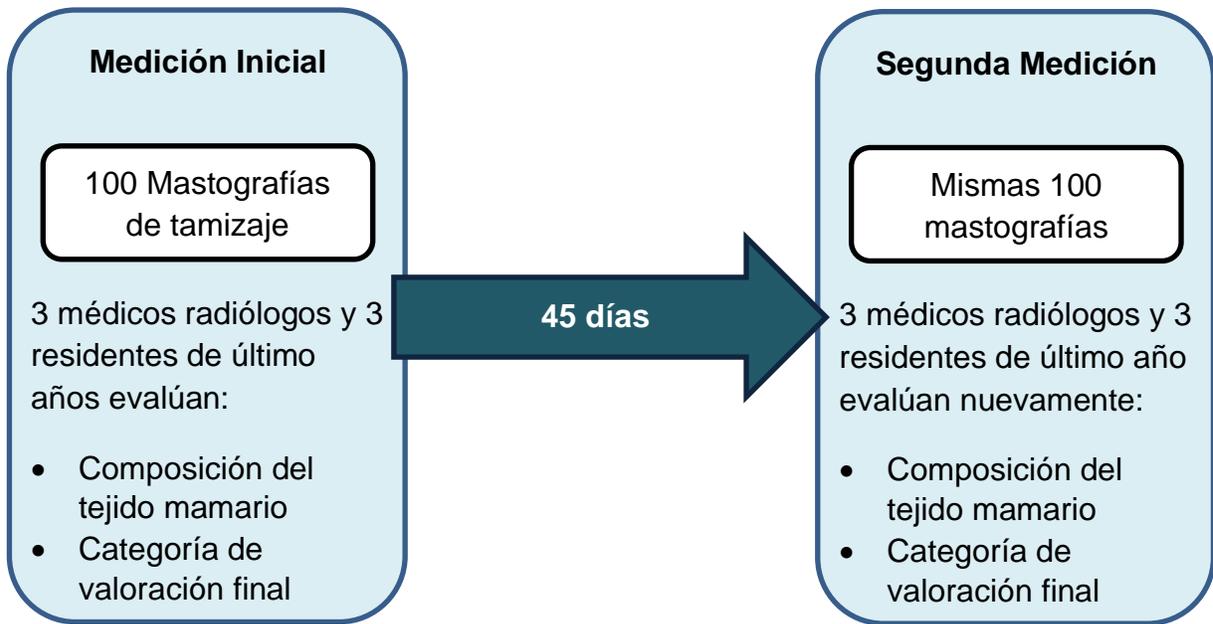


Figura 1. Resumen de las características generales del estudio.

Obtenidos los dos sets de resultados se realiza el análisis para calcular las concordancias inter e intraobservadores.

Para estimar la concordancia entre observadores para variables cualitativas, los datos se analizan mediante el cálculo del coeficientes de Kappa, que se reportan con un intervalo de confianza (IC) del 95%. Este coeficiente resume la concordancia entre dos o más medidas de una variable eliminando la fracción de la concordancia debida al azar, representado por la siguientes fórmula:

$$Kappa = \frac{Po - Pe}{1 - Pe}$$

Donde Po es la proporción total de concordancia observada y Pe la proporción de concordancia esperada por el azar.

Los datos obtenidos en el estudio para cada variable se organizarán en tablas de contingencia ejemplificadas en el **Cuadro 5.** y **Cuadro 6.**

Cuadro 5. Tabla de contingencia para composición del tejido mamario.

HALLAZGOS OBSERVADOR 1	HALLAZGOS OBERVADOR 2				
Medidas	A	B	C	D	TOTAL
A					
B					
C					
D					
TOTAL					

Cuadro 6. Tabla de contingencia para categoría de valoración final.

HALLAZGOS OBSERVADOR 2	HALLAZGOS OBERVADOR 3						
Medidas	0	1	2	3	4	5	TOTAL
0							
1							
2							
3							
4							
5							
TOTAL							

Una vez organizados los datos en las tablas de contingencia se puede obtener la proporción total observada de concordancia (P_o) y la proporción e concordancia esperada por el azar (P_e) para posteriormente calcular el índice de Kappa para cada variable utilizando la fórmula presentada anteriormente.

El valor máximo de un índice de Kappa es 1 cuando existe acuerdo total entre ambos observadores, el valor mínimo puede ser negativo si el porcentaje observado es menor que el esperado por el azar.

Los valores de Kappa se interpretarán utilizando el método estándar propuesto por Landis y Koch como se muestra en el **Cuadro 7**.

Cuadro 7. Valores de Kappa y su concordancia.

Valor de Kappa	Grado de Concordancia
<0.2	Muy débil
0.21 – 0.40	Débil
0.41 – 0.60	Moderada
0.61 – 0.80	Buena
>0.80	Muy buena

Plan de trabajo.

Consultar cronograma de actividades anexo.

Aspectos éticos y recursos.

De acuerdo con lo establecido en el artículo 17 del Reglamento de la Ley General de Salud en Materia de Investigación para la Salud, el estudio se considera una investigación sin riesgo que no requiere obtención de consentimiento informado. (27)

Se cuenta con dictamen aprobatorio por el Comité de Investigación y el Comité de Ética en Investigación del hospital sede del estudio **Anexo 3 y 4.**

Recursos humanos y materiales.

- Recursos humanos:
 - Investigador principal: Revisión y correcciones del proyecto, guía en el diseño del estudio, revisión y correcciones de los resultados y conclusiones.
 - Asesor metodológico: Revisión, guía y corrección en el diseño metodológico, tamaño de la muestra, análisis estadístico de los datos y resultados.
 - Colaborador: Apoyo en la obtención y selección de los estudios mamográficos de acuerdo con los criterios de calidad como observador externo.
 - Tesista: Diseño, desarrollo y correcciones del proyecto de investigación, revisión de la literatura disponible, obtención de la muestra y observadores,

ejecución del estudio, análisis de los datos obtenidos, obtención de resultados y conclusiones.

- Recursos materiales:
 - Estaciones de trabajo marca DELL disponibles en la sala de interpretación del departamento de radiología.
 - Formato de llenado para interpretación de los estudios desarrollado por el tesista.

Capacitación de personal.

No se requiere ningún tipo de capacitación o adiestramiento previo del personal involucrado.

Financiamiento.

Para este estudio no se contará con ningún tipo de financiamiento.

Resultados.

Sobrepasando el requisito de tamaño de muestra mínimo calculado de 64, el estudio comprendió 100 mastografías de tamizaje seleccionadas aleatoriamente. La edad media de las pacientes fue 51.4 (40 – 76 años).

Para objeto de análisis estadístico, nos referimos a los observadores con las siguientes variables: A y G: observador 1, B y H: observador 2, C e I: observador 3, D y J: observador 4, E y K: observador 5 y F y L: observador 6. A, B, C, D, E y F corresponden a la primera ronda y G, H, I, J, K y L a la segunda ronda de lectura para los observadores 1, 2, 3, 4, 5 y 6 respectivamente siendo los observadores 1, 2 y 3 los médicos radiólogos y 4, 5, y 6 los residentes de radiología de último año.

En la primera ronda de evaluación, la concordancia entre los médicos radiólogos (observadores A, B y C) fue de débil a buena (K 0.36 – 0.65) para composición del tejido y moderada a buena (K 0.50 – 0.61) para categoría de valoración final. La

concordancia entre los residentes (Observadores D, E y F) fue moderada (K 0.50 – 0.52) para composición del tejido mamario e igualmente moderada (K 0.44 – 0.53) para categoría de valoración, los grados de concordancia entre cada uno de los observadores se muestran en el **Cuadro 8**.

Cuadro 8. Concordancia interobservador en la primera ronda.

Composición del Tejido Mamario		Categoría de Valoración Final	
Observador	Valor de Kappa (IC 95%)	Observador	Valor de Kappa (IC 95%)
Radiólogos			
Entre A y B	0.36 (0.23 – 0.49)	Entre A y B	0.50 (0.36 – 0.64)
Entre A y C	0.65 (0.52 – 0.79)	Entre A y C	0.61 (0.47 – 0.75)
Entre B y C	0.52 (0.39 – 0.66)	Entre B y C	0.62 (0.47 – 0.76)
Residentes			
Entre D y E	0.50 (0.36 – 0.50)	Entre D y E	0.45 (0.30 – 0.59)
Entre D y F	0.53 (0.40 – 0.66)	Entre D y F	0.46 (0.31 – 0.60)
Entre E y F	0.52 (0.39 – 0.66)	Entre E y F	0.54 (0.39 – 0.68)

Después de 45 días, en la segunda ronda, la concordancia entre los médicos radiólogos para composición del tejido mamario incrementó a moderada a buena (K 0.43 – 0.71) y aunque hubo diferencias detectables en los resultados para categoría de valoración final, la concordancia general permaneció igual, moderada a buena (K 0.45 – 0.65), en comparación con los valores obtenidos en la primera ronda. Con respecto a los residentes de último año, la concordancia para composición del tejido mamario fue moderada a buena (K 0.41 – 0.63) representando un ligero incremento y para categoría de valoración una ligera disminución siendo débil a moderada (K 0.40 – 0.54), los grados de concordancia entre los pares de observadores se muestra en el **Cuadro 9**.

Cuadro 9. Concordancia interobservador en la segunda ronda.

Composición del Tejido Mamario		Categoría de Valoración Final	
Observador	Valor de Kappa (IC 95%)	Observador	Valor de Kappa (IC 95%)
Radiólogos			
Entre G y H	0.43 (0.30 – 0.56)	Entre G y H	0.56 (0.42 – 0.70)
Entre G e I	0.71 (0.60 – 0.84)	Entre G e I	0.65 (0.53 – 0.78)
Entre H e I	0.43 (0.30 – 0.57)	Entre H e I	0.45 (0.32 – 0.59)
Residentes			
Entre J y K	0.55 (0.42 – 0.69)	Entre J y K	0.54 (0.40 – 0.68)
Entre J y L	0.41 (0.28 – 0.55)	Entre J y L	0.40 (0.26 – 0.55)
Entre K y L	0.63 (0.50 – 0.76)	Entre K y L	0.47 (0.33 – 0.60)

Con respecto a la concordancia intraobservadores, para composición del tejido mamario la concordancia entre los médicos radiólogos fue buena (K 0.69 – 0.76) obteniendo el observador 1 el valor más alto, la concordancia entre los residentes fue moderada a buena (K 0.48 – 0.68) con el valor de kappa más alto obtenido por el observador 5. Para categoría de valoración final la concordancia intraobservador entre los médicos radiólogos fue moderada a buena (0.60 – 0.69) con el valor más alto para el observador 2 y entre los residentes de último año la concordancia fue moderada a buena (K 0.42 – 0.70) siendo la más alta obtenido por el observador 6. La concordancia personal obtenida por cada uno de los observadores se expone en el **Cuadro 10**.

Cuadro 10. Concordancia intraobservador.

Composición del Tejido Mamario		Categoría de Valoración Final	
Observador	Valor de Kappa (IC 95%)	Observador	Valor de Kappa (IC 95%)
Radiólogos			
Entre A y G	0.76 (0.63 – 0.88)	Entre A y G	0.60 (0.45 – 0.74)
Entre B y H	0.73 (0.63 – 0.85)	Entre B y H	0.69 (0.55 – 0.82)
Entre C e I	0.69 (0.55 – 0.82)	Entre C e I	0.65 (0.52 – 0.78)
Residentes			
Entre D y J	0.56 (0.42 – 0.70)	Entre D y J	0.42 (0.28 – 0.57)
Entre E y K	0.68 (0.56 – 0.80)	Entre E y K	0.59 (0.45 – 0.73)
Entre F y L	0.48 (0.35 – 0.62)	Entre F y L	0.70 (0.57 – 0.83)

El **Cuadro 11.** muestra la distribución de la suma de todos los pares de resultados en ambas rondas de evaluación para composición del tejido mamario. La discordancia interobservador más frecuente fue entre categorías adyacente B y C (17%) seguido de las categorías C y D (6.8%). Para la combinación intraobservador la discordancia más frecuente fue igualmente entre las categorías B y C (13.3%) seguido de las categorías C y D (4.6%).

Cuadro 11. Sumatoria de concordancia y discordancias para densidad mamaria.

Categoría	A	B	C	D
Interobservadores				
A	29 (2.41%)	69 (5.75%)	1 (0.08%)	0
B		438 (36.5%)	205 (17.08%)	2 (0.16%)
C			304 (25.33%)	82 (6.83%)
D				70 (5.83%)
Intraobservador				
A	22 (3.66%)	20 (3.33%)	0	0
B		238 (39.66%)	80 (13.3%)	0
C			170 (28.33%)	28 (4.6%)
D				42 (7%)

Para categoría de valoración final la discordancia interobservador más alta fue entre las categorías 0 y 2 (17%) seguido de las categorías 1 y 2 (6.7%). Para la valoración intraobservadores la discordancia más frecuente fue entre las categorías 0 y 2 (14.2%) y en segundo lugar entre las categorías 1 y 2 (5.2%). La distribución de la suma de todos los resultados se muestra en el **Cuadro 12**.

Cuadro 12. Sumatoria de concordancias y discordancias para categoría final.

Categoría	0	1	2	3	4	5
Interobservadores						
0	484 (40.33%)	12 (1%)	205 (17.08%)	3 (0.25%)	6 (0.5%)	
1		12 (1%)	81 (6.75%)	0	1 (0.083%)	
2			384 (32%)	7 (0.58%)	5 (0.41%)	
3				0	0	
4					0	
5						
Intraobservadores						
0	252 (42%)	7 (1.16%)	85 (14.16%)	0	1 (0.16%)	
1		10 (1.66%)	31 (5.16%)	1 (0.16%)	0	
2			206 (17.16%)	4 (0.66%)	1 (0.16%)	
3				0	0	
4					2 (0.33%)	
5						

En el **Cuadro 13.** se presenta una media global del porcentaje de concordancia y de los valores del coeficiente de Kappa entre todos los observadores para ambas variables.

Cuadro 13. Media de resultados de todos los observadores.

Categoría	Interobservadores		Intraobservadores	
	% de concordancia	Kappa (IC 95%)	% de concordancia	Kappa (IC 95%)
Composición del tejido	70	0.52 (0.48 – 0.56)	78.65	0.66 (0.60 – 0.70)
Categoría final	73.33	0.53 (0.49 – 0.57)	61.15	0.62 (0.56 – 0.67)

La frecuencia de las categorías de densidad del tejido mamario asignadas por cada observador en la primera evaluación se muestran en las **Figura 2.** y las obtenidas en la segunda vuelta en la **Figura 3.** La distribución de la densidad mamaria en nuestra población en la primera ronda en orden descendente fue la siguiente: B (45%), C (39%), D (8%) y A (6%), para la segunda evaluación fue: B (50%), C (35%), D (10%) y A (4.7%).

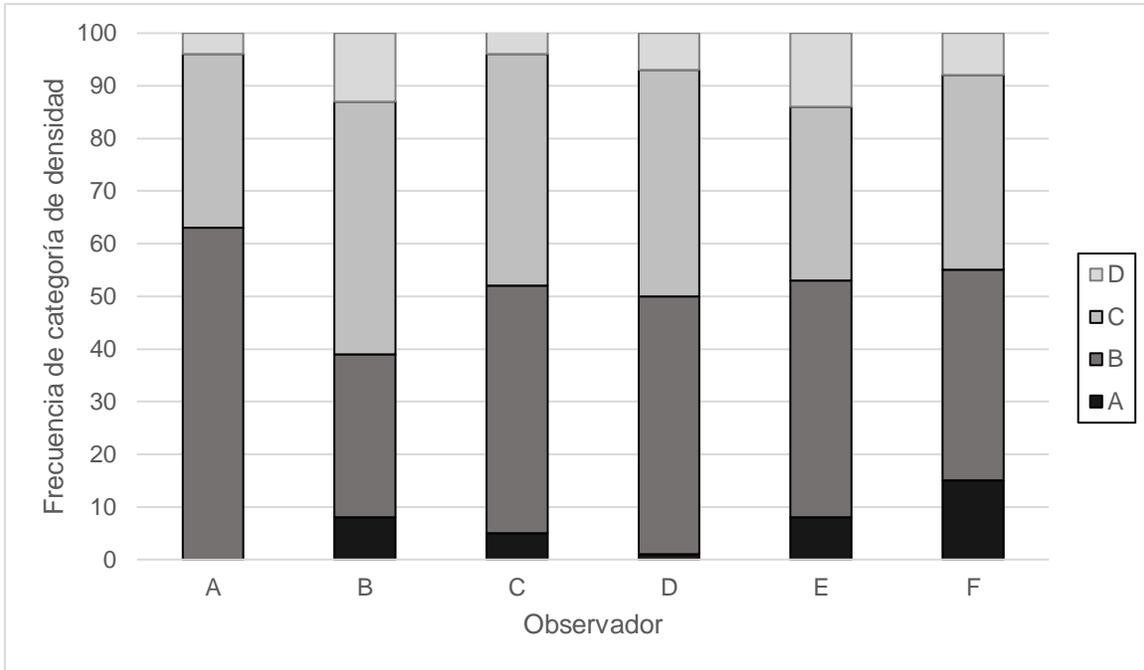


Figura 2. Distribución de la densidad mamaria en la primera evaluación.

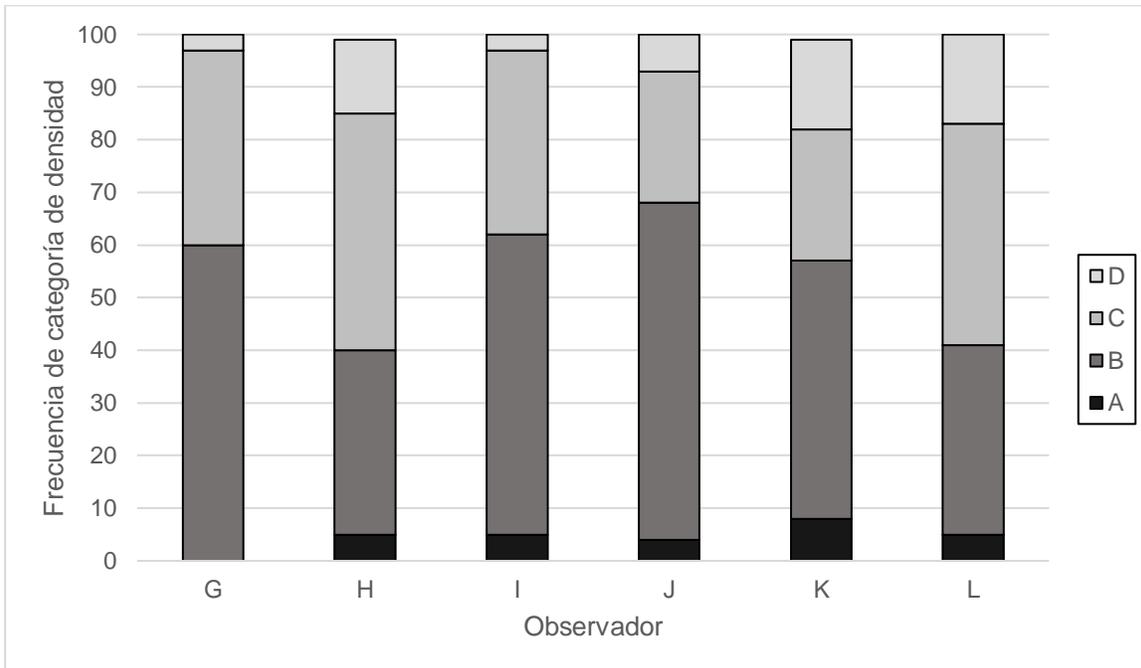


Figura 3. Distribución de la densidad mamaria en la segunda evaluación.

La distribución de las categorías categoría de valoración final asignadas por cada observador en la primera evaluación se muestran en las **Figura 4.** y las obtenidas en la segunda vuelta en la **Figura 5.** El orden descendente de las categorías asignadas en la primera ronda es: 0 (51.2%), 2 (43.8%), 1 (3.7%), 3 (0.7%), 4 (0.7%) y 5 (0%) para la segunda evaluación fue: 0 (50.8%), 2 (43.3%), 1 (5.3%), 4 (0.33%) y 5 (0%).

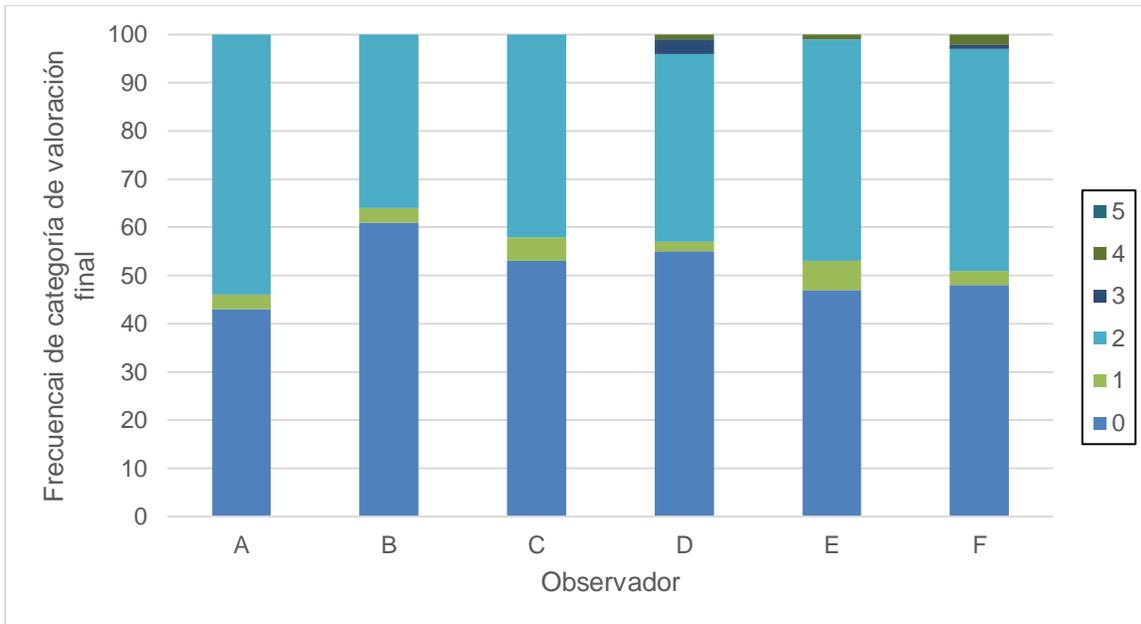


Figura 4. Distribución de las categorías finales en la primera evaluación.

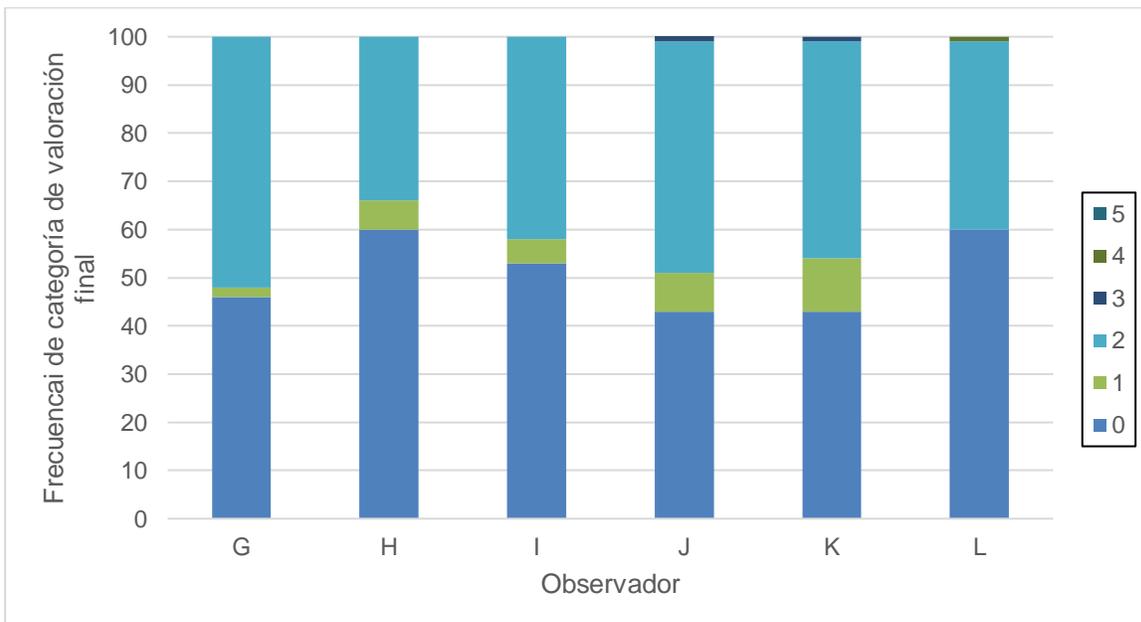


Figura 5. Distribución de las categorías finales en la segunda evaluación.

Discusión.

En el presente estudio, utilizamos el sistema BI-RADS 5ª edición para determinar la concordancia inter e intraobservador para las categorías densidad del tejido mamario y valoración final.

Interpretando con cuidado los resultados, el grado de concordancia puede ser utilizado como un sustituto de la precisión para este tipo de estudios en que no es factible la comparación con un estándar de oro(17). Por lo tanto minimizar la variabilidad en la interpretación de mastografías de tamizaje es vital para maximizar la precisión del estudio en la detección de lesiones sospechosas de cáncer de mama.

La concordancia interobservador para composición del tejido mamario fue moderada entre los médicos radiólogos e igualmente moderada entre los residentes de último año observando que existió un aumento en la concordancia en la segunda ronda de evaluación. En términos de categoría final la concordancia interobservador entre los radiólogos fue moderada y de igual forma moderada entre los residentes sin diferencias significativas entre la primer y segunda ronda.

Los resultados de concordancia intraobservador muestran para composición del tejido mamario una concordancia buena para los médicos radiólogos y moderada para los residentes de último año. Para categoría de valoración final la concordancia intraobservador fue buena para los radiólogos y moderada para los médicos residentes. Haciendo una revisión de los datos, se observó una variabilidad importante en el observador 4, el cual asignó más categorías definitivas en la segunda ronda, esto fue atribuido a que este observador acudió, durante el periodo de 45 días entre evaluaciones, a una rotación de 1 mes a un centro especializado en imagen mamaria, reflejando cómo la capacitación puede tener un efecto en el nivel de confianza del radiólogo.

En comparación con estudios previos que alcanzaron concordancia interobservador buena para densidad del tejido mamario, nuestros observadores solamente alcanzaron

un grado de concordancia moderado, sin embargo, para categoría final la concordancia interobservador alcanzada en nuestro estudio fue igual a la reportada en otros estudio siendo de grado moderado.

Para concordancia intraobservador, en la categoría de densidad del tejido, estudios previos reportan una grado máximo alcanzado de muy buena siendo en nuestro estudio el grado máximo alcanzado de buena y para categoría de valoración final se reporta una concordancia moderada, alcanzando en nuestro estudio una concordancia buena.

La densidad mamaria es considerada un factor de riesgo independiente para el desarrollo de cáncer de mama(4). La categorización de la densidad es fundamental para realizar una evaluación personalizada del riesgo y eficientizar el tamizaje poblacional con estudios más adecuados y con mayor sensibilidad en pacientes con patrones densos (C y D), como la tomosíntesis, ultrasonido y resonancia magnética. Esto evita que se realicen demasiadas o insuficientes pruebas complementarias y se presenten retrasos en el acceso a dichas pruebas.

El BI-RADS quinta edición, reporta una distribución histórica de la densidad mamaria de la siguiente forma: densidad adiposa (A), 10%; densidad dispersa (B), 40%; densidad heterogénea (C), 40% e hiperdensidad (D), 10% (4).

El promedio de la distribución de la densidad mamaria en este estudio y que representa una muestra de la distribución en nuestra población fue: densidad A, 5.3%; densidad B, 47.5%; densidad C, 37%; densidad D, 9%, similar a lo informado en el BI-RADS.

Las guías del colegio americano de radiología definen la tasa de rellamado como la proporción de estudios de tamizaje con una evaluación inicial positiva (es decir, categoría 0, 3, 4 o 5), una tasa de rellamado muy baja puede asociarse con una menor sensibilidad y aumento en los resultados falsos negativos, mientras que una tasa demasiad alta resulta en un aumento de los falsos positivos y una menor especificidad, mayores costos, ansiedad de las pacientes y sobretratamiento, por lo que el ACR

recomienda como objetivo una de tasa de rellamado del 5 al 12% para la mastografía de tamizaje(4). En este estudio encontramos una tasa de rellamado promedio de 12.79% para los residentes, 13.16% para los médicos adscritos y 13% para todos los observadores.

Limitaciones y/o nuevas perspectivas de investigación.

Existen algunas limitaciones encontradas en el estudio que pudieron haber influido en los resultados. Primero aunque todas las mastografías seleccionadas para el estudio cumplieron con los criterios de calidad en la adquisición, estas fueron obtenidas con diferentes mastógrafos, lo que pudiera incidir en distintos parámetros de la imagen como son contraste, resolución, artefactos, etc.

Segundo, si bien los médicos radiólogos que participaron en el estudio cuentan con al menos 5 años de experiencia en la práctica de la radiología general, no se tomó en cuenta el número de mastografías evaluadas por año o el número de horas laborales dedicadas a la interpretación de mastografías, por lo que el grado de experiencia en esta área específica no fue considerado.

Por último, el periodo de tiempo entre la primera y la segunda evaluación fue de 45 días, lapso que pudiera ser considerado insuficiente para evitar un sesgo por memoria de los observadores.

Conclusiones.

La concordancia interobservador para composición del tejido mamario alcanzada en este estudio fue moderada, inferior a lo reportado en otros estudios y moderada para la categoría de valoración final, igual a lo encontrado en estudios previos.

La concordancia intraobservador para densidad mamaria y categoría de valoración final fue buena, hallazgo inferior y superior a lo encontrado en otros estudios respectivamente.

El nivel de experiencia del observador no se reflejó en una diferencia significativa en la concordancia interobservador, sin embargo, sí en una concordancia intraobservador más alta para los radiólogos más experimentados.

A pesar de la existencia de guías estandarizadas para reportar la densidad y categorías de evaluación, la variabilidad del observador continúa existiendo, la capacitación continua en el léxico y categorías BI-RADS es de vital importancia, con el objetivo aumentar el nivel de confianza de los radiólogos y aumentar la concordancia en la interpretación de imágenes.

Alrededor del 46% de las pacientes de nuestra población tendrán mamas con patrones densos que requerirán un estudio de imagen complementario, esta cifra debe de ser contemplada por las autoridades sanitarias al momento de diseñar programas que busquen reducir la mortalidad por cáncer de mama.

La tasa de rellamado de nuestro departamento de imagenología se encuentra por encima de lo recomendado, por lo que es importante implementar estrategias que busquen reducir este número.

Referencias bibliográficas.

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;71(3):209–49.
2. Incidencia de casos nuevos de enfermedad por grupos de edad Estados Unidos Mexicanos 2020 Población Femenina [Internet]. [citado el 18 de abril de 2022]. Disponible en: https://epidemiologia.salud.gob.mx/anuario/20200/incidencia/incidencia_casos_nuevos_enfermedad_grupo_edad.pdf
3. Incidencia de Tumor maligno de la mama (C50) por grupos de edad Estados Unidos Mexicanos 2020 Población Femenina [Internet]. [citado el 18 de abril de 2022]. Disponible en: https://epidemiologia.salud.gob.mx/anuario/20200/incidencia/enfermedad_grupo_edad_entidad_federativa/119.pdf
4. Edward A. Sickles, Carl J. D'Orsi. Atlas BI-RADS de ACR. Sistema normalizado para la redacción de informes de estudios por imágenes de la mama. 5a ed. Reston, VA; 2013.
5. GUÍA DE PRÁCTICA CLÍNICA: PREVENCIÓN, TAMIZAJE Y REFERENCIA OPORTUNA DE CASOS SOSPECHOSOS DE CÁNCER DE MAMA EN EL PRIMER NIVEL DE ATENCIÓN [Internet]. [citado el 18 de abril de 2022]. Disponible en: <http://www.cenetec-difusion.com/CMGPC/S-001-08/ER.pdf>
6. Van Ourti T, O'Donnell O, Koç H, Fracheboud J, de Koning HJ. Effect of screening mammography on breast cancer mortality: Quasi-experimental evidence from rollout of the Dutch population-based program with 17-year follow-up of a cohort. *Int J Cancer.* el 15 de abril de 2020;146(8):2201–8.
7. Kalager M, Zelen M, Langmark F, Adami HO. Effect of Screening Mammography on Breast-Cancer Mortality in Norway. *N Engl J Med.* el 23 de septiembre de 2010;363(13):1203–10.
8. Zielonke N, Gini A, Jansen EEL, Anttila A, Segnan N, Ponti A, et al. Evidence for reducing cancer-specific mortality due to screening for breast cancer in Europe: A systematic review. *Eur J Cancer.* el 1 de marzo de 2020;127:191–206.
9. Herrmann C, Vounatsou P, Thürlimann B, Probst-Hensch N, Rothermundt C, Ess S. Impact of mammography screening programmes on breast cancer mortality in Switzerland, a country with different regional screening policies. *BMJ Open.* marzo de 2018;8(3):e017806.

10. Anderson BO, Shyyan R, Eniu A, Smith RA, Yip CH, Bese NS, et al. Breast Cancer in Limited-Resource Countries: An Overview of the Breast Health Global Initiative 2005 Guidelines. *Breast J.* 2006;12(s1):S3–15.
11. da Costa Vieira RA, Biller G, Uemura G, Ruiz CA, Curado MP. Breast cancer screening in developing countries. *Clinics.* abril de 2017;72(4):244–53.
12. Iwamoto Y, Kaucher S, Lorenz E, Bärnighausen T, Winkler V. Development of breast cancer mortality considering the implementation of mammography screening programs – a comparison of western European countries. *BMC Public Health.* el 26 de junio de 2019;19(1):823.
13. Brandan ME, Ruiz-Trejo C, Verdejo-Silva M, Guevara M, Lozano-Zalce H, Madero-Preciado L, et al. Evaluation of equipment performance, patient dose, imaging quality, and diagnostic coincidence in five Mexico City mammography services. *Arch Med Res.* febrero de 2004;35(1):24–30.
14. Torres-Mejía G, Villaseñor-Navarro Y, Yunes-Díaz E, Ángeles-Llerenas A, Martínez-Montañez OG, Lazcano-Ponce E. Validez y reproducibilidad de la interpretación de la mamografía por radiólogos mexicanos, mediante el sistema BI-RADS. :11.
15. Burnside ES, Sickles EA, Bassett LW, Rubin DL, Lee CH, Ikeda DM, et al. The ACR BI-RADS® Experience: Learning From History. *J Am Coll Radiol JACR.* diciembre de 2009;6(12):851–60.
16. Josep M. Argimon Pallás, Josep Jiménez Villa. *Métodos de investigación clínica y epidemiológica.* 4a ed. Elsevier; 2013.
17. Harlold L. Kundel, Marcia Polansky. *Measurement of Observer Agreement.* *Radiology.* agosto de 2013;6.
18. Masroor I, Rasool M, Saeed SA, Sohail S. To asses inter- and intra-observer variability for breast density and BIRADS assessment categories in mammographic reporting. *JPMA J Pak Med Assoc.* febrero de 2016;66(2):194–7.
19. Ekpo EU, Ujong UP, Mello-Thoms C, McEntee MF. Assessment of Interradiologist Agreement Regarding Mammographic Breast Density Classification Using the Fifth Edition of the BI-RADS Atlas. *AJR Am J Roentgenol.* mayo de 2016;206(5):1119–23.
20. Irshad A, Leddy R, Ackerman S, Cluver A, Pavic D, Abid A, et al. Effects of Changes in BI-RADS Density Assessment Guidelines (Fourth Versus Fifth Edition) on Breast Density Assessment: Intra- and Interreader Agreements and Density Distribution. *AJR Am J Roentgenol.* diciembre de 2016;207(6):1366–71.

21. Alomaim W, O'Leary D, Ryan J, Rainford L, Evanoff M, Foley S. Variability of Breast Density Classification Between US and UK Radiologists. *J Med Imaging Radiat Sci.* marzo de 2019;50(1):53–61.
22. Pesce K, Tajerian M, Chico MJ, Swiecicki MP, Boietti B, Frangella MJ, et al. Interobserver and intraobserver variability in determining breast density according to the fifth edition of the BI-RADS® Atlas. *Radiologia.* diciembre de 2020;62(6):481–6.
23. Eom HJ, Cha JH, Kang JW, Choi WJ, Kim HJ, Go E. Comparison of variability in breast density assessment by BI-RADS category according to the level of experience. *Acta Radiol Stockh Swed* 1987. mayo de 2018;59(5):527–32.
24. Holland K, van Zelst J, den Heeten GJ, Imhof-Tas M, Mann RM, van Gils CH, et al. Consistency of breast density categories in serial screening mammograms: A comparison between automated and human assessment. *Breast Edinb Scotl.* octubre de 2016;29:49–54.
25. Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Cohen's Kappa. *Epidemiol Biostat Public Health.* el 1 de enero de 2017;14:e12267-1.
26. Flack VF, Afifi AA, Lachenbruch PA, Schouten HJA. Sample size determinations for the two rater kappa statistic. *Psychometrika.* el 1 de septiembre de 1988;53(3):321–5.
27. Reglamento de la Ley General de Salud en Materia de Investigación para la Salud. 2014.

Anexos.

Anexo 1. Cronograma de actividades.

CRONOGRAMA DE ACTIVIDADES									
Actividad	Mes								
	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre
Tema	✓								
Asesor	✓								
Delimitación del Tema	✓								
Planteamiento del Problema		✓							
Factibilidad		✓							
Objetivos		✓							
Antecedentes		✓							
Pregunta de Investigación		✓							
Justificación		✓							
Hipótesis		✓							
Metodología		✓							
Aspectos éticos		✓							
Referencias		✓							
Obtención de Muestra			✓						
Experimentación 1					✓				
Análisis estadístico 1					✓				
Experimentación 2						✓			
Análisis estadístico 2						✓	✓		
Resultados							✓	✓	
Conclusiones								✓	✓

Anexo 2. Formato de interpretación.

FORMATO PARA INTERPRETACIÓN DE MASTOGRAFÍAS.									
Número de Estudio	Composición del Tejido Mamario (ACR)				Categoría de Valoración Definitiva (BI-RADS)				
1	A	B	C	D	1	2	3	4	5
2	A	B	C	D	1	2	3	4	5
3	A	B	C	D	1	2	3	4	5
4	A	B	C	D	1	2	3	4	5
5..	A	B	C	D	1	2	3	4	5