



ORIGINAL

Competencia diagnóstica de los médicos de atención primaria con distintas técnicas de visualización de radiografías[☆]



Tamara Sigal* y Eduardo Durante

Servicio de Medicina Familiar y Comunitaria, Hospital Italiano de Buenos Aires, Buenos Aires, Argentina

Recibido el 23 de junio de 2016; aceptado el 3 de noviembre de 2016

Disponible en Internet el 13 de diciembre de 2016

PALABRAS CLAVE

Competencia diagnóstica;
Radiografías;
Atención primaria

Resumen

Objetivo: Explorar de forma pragmática si existen diferencias en la competencia diagnóstica de médicos de atención primaria utilizando diferentes técnicas de imagen.

Métodos: Estudio observacional de corte transversal incluyendo 10 médicos residentes (RMF) y 10 médicos posresidentes del servicio de medicina familiar (PRMF) y 16 médicos residentes del servicio de diagnóstico por imágenes (RDI). Los participantes evaluaron 7 casos clínicos de atención primaria con sus respectivas radiografías. Se utilizaron radiografías impresas y digitales. Cada radiografía presentaba un hallazgo principal y varios hallazgos sutiles, definidos por un especialista en diagnóstico por imágenes. Se calculó el acuerdo intraindividuo e intragrupo.

Resultados: Se decidió considerar a los participantes del grupo como un único observador para estimar un kappa global. Para RMF fue de 0,48 (IC 95%: 0,35-0,61), para PRMF de 0,48 (IC 95%: 0,35-0,61) sin encontrarse diferencias significativas entre ambos grupos. Para los RDI el kappa fue de 0,58 (IC 95%: 0,48-0,69). La puntuación de la descripción de cada pregunta podía variar 0 y 10. El promedio de puntuación obtenido para método impreso y digital respectivamente fue: 4,15 y 4,15 para RMF; 4,2 y 3,77 para PRMF y 7,03 y 7,04 para RDI.

Conclusión: No se encontraron diferencias significativas en el uso de ambas técnicas para los 3 subgrupos, aunque existe una tendencia a mejor desempeño con método el digital entre los médicos residentes. El desempeño de los RMF fue significativamente menor que los RDI.

© 2016 Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

[☆] Premio recibido por este trabajo: Mejor trabajo de investigación en Educación Médica-Congreso Iberoamericano de Medicina Familiar (CIMF-WONCA) 2015, Montevideo.

* Autor para correspondencia.

Correo electrónico: tamara.sigal@hospitalitaliano.org.ar (T. Sigal).

KEYWORDS

Radiological
diagnostic skills;
Radiography;
Primary care

Differences in radiological diagnostic skills in primary care physicians**Abstract**

Objective: A study was designed to assess whether there were differences in the diagnostic skills of Primary Care Physicians using different diagnostic techniques.

Methods: A cross-sectional observational study was conducted that included 10 Family Medicine Residents (FMR), 10 Specialists in Family Medicine (SFM), and 16 Radiology Residents (RR) as a control group. The participants evaluated 7 real Primary Care cases with their respective printed and digital X-rays. Each radiograph had a main finding and several subtle findings, defined by specialist in diagnostic imaging. The intra- and extra-individual agreement was calculated. It was decided to consider the participants of a group as a single observer to calculate an overall Kappa.

Results: The overall Kappa coefficients between film and screen X- rays for FMR was 0.48 (95% CI; 0.35-0.61), for SFM 0.48 (95% CI; 0.35-0.61), and for the RR Group, 0.58 (95% CI; 0.48-0.69). The mean score (range from 0 to 10) obtained for the printed and digital method was 4.15 and 4.15 for the FMR, 4.2 and 3.77 for SFM, and 7.03 and 7.04 for RR, respectively.

Conclusion: No significant differences were found in radiological diagnostic skills for the 3 groups in the use of both techniques, although there was a tendency of a better performance with the digital method among medical residents.

© 2016 Elsevier España, S.L.U. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

La radiografía digital, debido a sus múltiples ventajas, se ha adoptado como herramienta diagnóstica de múltiples enfermedades. Los nuevos monitores, junto con los programas y *hardware* adecuados, permiten mostrar imágenes de alta calidad, similares a las proporcionadas por las copias en película radiográfica. El empleo de sistemas digitales para el diagnóstico clínico supone algunas ventajas: permite la modificación interactiva de los niveles de contraste de la imagen y mayor rapidez en la representación de las imágenes frente a los sistemas de impresión sobre película. Además, este sistema permite que la misma imagen sea vista simultáneamente en distintos lugares¹. Sin embargo, la validez para la aplicación clínica del uso de imágenes digitales aún no ha sido totalmente establecida.

En los últimos años se han llevado a cabo varios estudios que comparan la capacidad operativa de ambas técnicas con el fin de comprobar la eficacia de las técnicas digitales en la interpretación de diferentes enfermedades. Garmer et al.² compararon el desempeño diagnóstico de los médicos con las técnicas digitales vs. las radiográficas convencionales para diagnosticar imágenes patológicas de tórax. Se utilizó como estándar de oro el consenso de 3 radiólogos independientes, quienes se ayudaron de la imagen obtenida por tomografía computarizada. Las 160 imágenes (80 digitales y 80 convencionales) fueron aleatorizadas y evaluadas de forma separada por cada participante. La misma imagen fue nuevamente evaluada en un intervalo de 6 meses, utilizando el otro método de imágenes. Para comparar los desempeños de las 2 tecnologías se utilizó el método de análisis del área bajo la curva. El estudio concluyó que no existen diferencias significativas entre las áreas bajo la curva de cada método. Por otro lado, Thompson et al.³ compararon ambas tecnologías (digital y convencional) en la detección de dispositivos

cardiovasculares en las imágenes de tórax. Al igual que el trabajo previamente mencionado, se concluyó que no hubo diferencias significativas entre ambas técnicas. Elam et al.⁴ evaluaron la habilidad de radiólogos para detectar neumotórax en las radiografías digitales comparándolas con las convencionales. El análisis fue realizado utilizando el área bajo la curva. La media obtenida en las curvas no fue estadísticamente significativa entre ambos métodos. Por último, Thaete et al.⁵ compararon el desempeño de los observadores para interpretar imágenes patológicas de tórax, utilizando técnicas digitales, convencionales y en monitor. Los resultados sugieren que las observaciones de imágenes patológicas utilizando radiografías digitales son similares a las obtenidas con las imágenes convencionales. Sin embargo, estos fueron significativamente menores al evaluar imágenes de neumotórax y enfermedad intersticial con monitores. Los autores creen que esta diferencia se debería al bajo grado de resolución y de contraste de los monitores utilizados en el estudio.

Además de la evidencia citada, que muestra que no hay diferencias en la interpretación de enfermedades entre ambos métodos, la tecnología digital otorga beneficios significativos que fueron reconocidos en estudios previos: buena calidad, fácil distribución y acceso, lo que ha llevado a que en muchos centros las radiografías convencionales sean reemplazadas⁶⁻¹⁰.

Sin embargo, no existe evidencia sobre la competencia diagnóstica de los médicos de atención primaria (MAP) comparando ambos métodos, lo que ha motivado el presente estudio.

El objetivo de la investigación consistió en explorar de forma pragmática si existen diferencias en la competencia diagnóstica de médicos de Atención Primaria utilizando diferentes técnicas de imagen. El objetivo secundario consistió en explorar la validez y la fiabilidad de un nuevo instrumento

de evaluación de la competencia de diagnóstico utilizando radiografías digitales e impresas.

Material y método

Diseño

Se realizó un estudio observacional de corte transversal en el Hospital Italiano de Buenos Aires, entre julio de 2010 y diciembre de 2011.

Procedimiento

Para el objetivo primario se seleccionaron imágenes radiográficas con diferentes hallazgos patológicos (tórax, columna lumbosacra, cadera y ambas manos), de pacientes entre 18 y 80 años de edad de ambos sexos atendidos en consultorios de atención primaria del Hospital Italiano en el año 2009. Estas imágenes fueron seleccionadas por un médico experto del Servicio de Diagnóstico por Imágenes («estándar de oro»), quien evaluó las radiografías digitales impresas y en monitor, determinando todos los hallazgos presentes y la jerarquía de los mismos. Este experto no fue incluido en el análisis del estudio.

Los participantes fueron seleccionados por un muestreo estratificado entre los médicos del Servicio de Medicina Familiar y Comunitaria del Hospital Italiano de Buenos Aires ($n=20$), de acuerdo a la experiencia en la práctica asistencial: residentes y especialistas (con más de 5 años de experiencia desde haber obtenido la especialidad).

Todos los participantes evaluaron las imágenes en las mismas condiciones de su práctica habitual, tanto para la interpretación de las radiografías impresas (iluminación de negatoscopio y de la sala) como de las radiografías digitales. En este último caso, los consultorios contaron con monitores de 17 pulgadas widescreen, con resolución 1440×90 píxeles y una aplicación java (Raim Java 2009) que permite utilizar una serie de herramientas para la manipulación de las imágenes (ampliación, reducción, rotación, función de espejo, inversión de los niveles de gris, función de lupa).

Con las radiografías seleccionadas se organizó un conjunto de casos que fueron presentados a los participantes para su evaluación. Este conjunto de radiografías fue representativo de los tipos de radiografías más frecuentemente solicitadas en atención primaria y con los hallazgos más frecuentes y/o importantes para hacer el diagnóstico. Cada caso consistió en la presentación de una breve viñeta clínica y de la correspondiente imagen mostrada en uno u otro método. Se excluyeron de las radiografías los nombres de los pacientes para preservar la confidencialidad. Para cada caso, el experto en diagnóstico por imágenes definió un criterio de aprobación (hallazgo índice), de forma tal que si el participante no lo reconocía el caso se consideraba no aprobado. Además, se definieron otros hallazgos radiológicos que podían observarse, pero que no eran fundamentales para el diagnóstico radiológico del caso (hallazgos secundarios). Los participantes debían describir las imágenes radiológicas halladas y definir un posible diagnóstico clínico. Se definió si cada participante «aprobó» o no cada caso según si reconoció o no el hallazgo índice. Por otra parte, a cada hallazgo se le otorgó una puntuación y se estableció una ponderación de

los hallazgos de acuerdo a su importancia clínica, teniendo el hallazgo índice la mayor puntuación. Los resultados fueron definidos como correctos o incorrectos utilizando una clave de corrección construida a partir de las definiciones del experto en diagnóstico por imágenes. Para cada participante se estableció el número de casos aprobados en cada método de imagen, así como la puntuación por cada caso y la puntuación global para cada técnica de imagen (digital o impresa).

El orden de los casos en el conjunto, así como el método de imagen (digital o impreso), fue aleatorizado para cada participante. Cada participante fue evaluado en 2 oportunidades con los mismos casos pero, en cada ocasión, las imágenes eran mostradas con uno u otro método. Después de un intervalo de 8 semanas los participantes evaluaron el mismo conjunto de casos, pero las imágenes fueron mostradas con el método que no se había utilizado en la primera ocasión, bajo las mismas condiciones descritas previamente. Es decir: la imagen vista inicialmente impresa fue vista en monitor y viceversa. Se computaron los casos aprobados y las puntuaciones de la misma manera que en la primera ocasión.

Cálculo del tamaño muestral

Debido a que no existe en la bibliografía suficiente información sobre posibles diferencias con el uso de ambos métodos en esta población de médicos, se decidió realizar una prueba piloto para determinar el tamaño muestral, el tiempo necesario para la prueba, las puntuaciones y las posibles dificultades en la implementación de la misma. Esta prueba se realizó con 5 médicos, a los cuales se les presentaron 6 radiografías (3 digitales y 3 impresas) y después de 8 semanas se repitió la exposición con el método opuesto. Cada caso tenía una puntuación máxima de 10 puntos repartidos de forma ponderada entre el hallazgo índice y otros hallazgos de menor significación clínica. El promedio de puntuaciones obtenido para cada método fue de 26 y 29 puntos (sobre 60), con una desviación estándar de 8.

Tras el análisis de esta prueba piloto se hicieron ajustes sobre el formulario de recolección de la información y se calculó el tamaño muestral, considerando ambos métodos como diferentes cuando existiera entre ellos una diferencia igual o mayor a 5 puntos en la puntuación promedio de competencia.

En función de la hipótesis previa, de que el nivel de entrenamiento y la práctica estuvieron relacionados con la competencia en la interpretación de las imágenes, se procedió a la estratificación de los médicos de familia en los 2 grupos de residentes y especialistas.

Dada la esperable correlación de las respuestas de cada médico para cada caso, y utilizando ambos métodos, el diseño del estudio consideró a cada uno como un clúster con un coeficiente de correlación intraclúster de 0,05. Asumiendo una diferencia de puntuación entre grupos de 5

Tabla 1 Coeficiente kappa global para cada subgrupo

	Kappa	DS	IC 95%
Residentes medicina familiar	0,48	0,06	0,35-0,61
Especialistas medicina familiar	0,48	0,06	0,35-0,61

Tabla 2 Promedio de puntuaciones de descripciones con ambos métodos

	Residentes medicina familiar		Especialistas medicina familiar		Residentes diagnóstico por imágenes	
Promedio de puntuación (DS)	Impreso	Digital	Impreso	Digital	Impreso	Digital
	4,20 (2,9)	4,10 (3,0)	4,20 (2,87)	3,77 (2,89)	7,03 (2,28)	7,04 (2,19)

puntos (30 a 25), una desviación estándar de 8 y un poder del 80%, se debía incluir al menos a 10 médicos (clústers) por subgrupo (residentes y posresidentes) y 14 observaciones por clúster; es decir, que de esta forma se incluirían al menos 140 observaciones por estrato.

Protocolo de puntuación

La competencia de los médicos en diagnosticar imágenes radiológicas se estableció determinando el número de casos aprobados y la puntuación global obtenida con cada método. El grado de acuerdo de las medidas con ambas técnicas se estableció calculando el coeficiente kappa.

Análisis estadístico

Para el objetivo primario se calculó el coeficiente kappa intraindividuo. Asumiendo que los observadores de un mismo grupo presentarían similar competencia diagnóstica, se decidió considerar a todos los participantes del grupo como un único observador para estimar un kappa global.

Para el objetivo secundario se diseñó un estudio de validez con grupos extremos. Se comparó a los médicos de familia con un grupo de 16 médicos residentes de Diagnóstico por Imágenes del Hospital Italiano de Buenos Aires. Para calcular la fiabilidad se utilizó el coeficiente alfa de Cronbach.

Resultados

Se incluyeron 20 médicos de familia (10 residentes y 10 especialistas) y 16 médicos residentes de diagnóstico por imágenes. Cada uno de ellos observó 7 radiografías con ambos métodos (digital e impresa).

El acuerdo intraindividuo calculado mediante el coeficiente kappa arrojó resultados que variaron entre 0,13 y 0,73 para los médicos de familia. Esto implica un bajo nivel de acuerdo entre las observaciones con ambos métodos.

En la [tabla 1](#) se pueden observar el kappa global para cada grupo. Para residentes de medicina familiar fue de 0,48 (DS: 0,06) y para posresidentes de 0,48 (DS: 0,06), sin encontrarse diferencias significativas entre ambos grupos. Para los residentes de imágenes el kappa fue de 0,58 (DS: 0,05).

En la [tabla 2](#) se observan los promedios de las puntuaciones de descripción obtenidos para cada pregunta con cada uno de los métodos (puntuación de cada pregunta varía entre 0 y 10). La puntuación obtenida para ambos grupos es similar y significativamente bajo. El coeficiente alfa de Cronbach fue de 0,84.

Conclusiones

De acuerdo a los resultados de nuestro estudio, los médicos de atención primaria no presentaron diferencias en su competencia diagnóstica en la interpretación de las imágenes radiológicas según sean mostradas en placa o en monitor. Este hallazgo coincide con los estudios realizados con observadores de otras especialidades¹⁻⁵. Se refuerza así la idea de que las imágenes digitales son iguales de útiles que las radiografías impresas para el diagnóstico clínico, pero además ofrecen la posibilidad de modificar de forma interactiva los niveles de contraste de la imagen, y permiten que la misma imagen sea vista simultáneamente en distintos lugares.

Por otra parte, el instrumento diseñado para evaluar la competencia en la interpretación de radiografías ha mostrado poseer validez de contenido (debido a su construcción), validez de constructo (debido a que se utilizó el método de grupos extremos¹⁰ y que hubo diferencias muy significativas entre dichos grupos extremos de médicos de atención primaria y especialistas en radiología) y una adecuada confiabilidad (ya que muestra un coeficiente alfa de Cronbach de 0,84) en esta población en estudio. Hasta nuestro conocimiento no había sido desarrollado un instrumento de evaluación de la competencia con estas características. Como se ha revelado válido y confiable, será posible aplicarlo en futuras evaluaciones para valorar cambios en la competencia diagnóstica de radiografías luego de intervenciones educativas que estén destinadas a mejorarla.

Este estudio cuenta como fortaleza haber sido realizado de forma pragmática, es decir, respetando las condiciones reales del trabajo de los médicos de Atención Primaria, simulando situaciones habituales de consultorio, a través de la presentación de viñetas clínicas reales y la utilización de los recursos habituales de consultorio (iluminación, monitor, negatoscopio, etc.).

Con respecto a los posibles sesgos (de contaminación y de recuerdo), se intentó minimizarlos utilizando la mayor cantidad de radiografías posibles, generando una secuencia aleatoria de examen para cada participante y separando las 2 exposiciones al examen por al menos 8 semanas.

En cuanto a las debilidades de este estudio debe mencionarse que la baja puntuación obtenida por los médicos de familia y el moderado grado de acuerdo entre ambos métodos (monitor y digital) pueden influir en la falta de poder para detectar diferencias significativas entre ambos métodos de visualización. Por otro lado, si bien los residentes de imágenes lograron una mejor puntuación y un coeficiente kappa más elevado, el nivel de acuerdo con ambos métodos es de todas formas moderado, dando cuenta la falta de acuerdo en el diagnóstico del mismo observador utilizando uno u otro método.

Este estudio contribuye a la conclusión de que la competencia diagnóstica de los médicos de atención primaria

es similar utilizando indistintamente radiografías impresas o digitales. Además aporta un instrumento de evaluación válido y confiable para médicos de atención primaria, que permite su utilización para futuros estudios y evaluaciones de desempeño de esta población.

Conflicto de intereses

Ninguno.

Agradecimientos

Al Dr. Andrés D'Ippolito por su participación en el desarrollo del protocolo de investigación. Al Dr. Osvaldo Velan por su asesoría en imágenes. A los Dres. Sergio Terrasa y Fernando Rubinstein por su asesoría en el análisis de resultados.

Bibliografía

1. Konen E, Greenberg I, Rozenman J. Visibility of normal thoracic anatomic landmarks on storage phosphor digital radiography versus conventional radiography. *IMAJ*. 2005;8:495-7.
2. Garmer M, Hennigs S, Jäger H, Schrick F, van de Loo T, Jacobs A, et al. Digital radiography versus conventional radiography in chest imaging. *AJR Am J Roentgenol*. 2000;174:75-80.
3. Thompson MJ, Kubicka RA, Smith C. Evaluation of cardiopulmonary devices on chest radiographs: Digital vs analog radiographs. *AJR Am J Roentgenol*. 1989;153:1165-8.
4. Elam EodA, Rehm K, Hillman BJ, Maloney K, Fajardo LL, McNeill K. Efficacy of digital radiography for the detection of pneumothorax: comparison with conventional chest radiography. *AJR Am J Roentgenol*. 1992;158:509-14.
5. Thaete FL, Fuhrman CR, Oliver JH, Britton CA, Campbell WL, Feist JH, et al. Digital radiography and conventional imaging of the chest: a comparison of observer performance. *AJR Am J Roentgenol*. 1994;162:575-81.
6. Mac Mahon H, Vybomy CJ, Metz CE, Britton CA, Campbell WL, Feist JH, et al. Digital radiography of subtle pulmonary abnormalities: An ROC study of the effect of pixel size on observer performance. *Radiology*. 1986;1:21-6.
7. Goodman LR, Foley WD, Wilson CR, Rimm AA, Lawson TL. Digital and conventional chest images: observer performance with film digital radiography system. *Radiology*. 1986;158:27-33.
8. Chakraborty DP, Breatnach ES, Yester MV, Soto B, Barnes GT, Fraser RG, et al. Digital and conventional chest imaging: a modified ROC study of observer performance using simulated nodules. *Radiology*. 1986;1:35-9.
9. Fraser RG, Breatnach E, Barnes GT. Digital radiography of the chest: Clinical experience with a prototype unit. *RadioGraphics*. 1985;5:415-28.
10. Merritt CRB, Tutton RH, Bell KA, Bluth EI, Kalmar JA, Matthews CC, et al. Clinical application of digital radiography: Computed radiographic imaging. *RadioGraphics*. 1985;5:397-414.