

¿Estamos de acuerdo con la calidad técnica de nuestras radiografías de tórax? Un radiólogo, un técnico especialista en radiodiagnóstico y un médico de familia intentan corroborarlo

J.M. Garófano-Jerez^a, F. Quesada-Jiménez^b, J. Cabrerizo-Castro^b, G. Morales-Navarra^b, T. García-Ruiz^b y P. Lardelli-Claret^c

Objetivo. Valorar el grado de concordancia en la interpretación de la calidad técnica de las radiografías de tórax de un centro de salud entre el técnico especialista en radiodiagnóstico (TER), el médico de familia y el radiólogo.

Diseño. Estudio transversal.

Emplazamiento. Atención primaria. Centro de Salud de Cartuja en Granada.

Participantes. Pacientes del Centro de Salud de Cartuja con estudio radiológico simple de tórax durante 2002. Mediante muestreo aleatorio simple se seleccionaron 150 estudios. Se rechazaron 2 por tratarse de exploraciones repetidas del mismo paciente y 7 porque no presentaban una calidad mínima. La muestra final fue de 141 exploraciones.

Mediciones principales. Los observadores cumplimentaron independientemente, para cada exploración, un protocolo de 11 ítems sobre calidad técnica de la imagen. Para cada ítem se calculó el índice kappa entre parejas de observadores, así como el índice kappa global.

Resultados. El 96% de las exploraciones fue realizado con placas radiográficas de gran tamaño (35-43). Sólo en 2 preguntas hubo una concordancia aceptable o buena entre los 3 observadores (kappa = 0,559-0,858). En 5 preguntas, la concordancia fue homogéneamente baja (kappa = 0,034-0,375). En algunas preguntas se apreció una fuerte discrepancia entre la valoración del TER y la realizada por los otros 2 observadores.

Conclusiones. La concordancia general puede considerarse baja, aunque algo mayor entre el radiólogo y el médico de familia que entre cada uno de éstos y el TER. Esto plantea la necesidad de mejorar la formación de los profesionales en cuanto a la valoración de la calidad técnica de las imágenes. Además, se detecta un gasto innecesario de placas radiográficas de gran tamaño.

Palabras clave: Centro de salud. Control de calidad. Diagnóstico por la imagen. Radiografía torácica. Evaluación.

DO WE AGREE ON THE TECHNICAL QUALITY OF OUR CHEST X-RAYS? A RADIOLOGIST, AN EXPERT IN X-RAY DIAGNOSIS, AND A FAMILY DOCTOR TRY TO FIND OUT

Objective. To appraise the degree of concordance in the interpretation of the technical quality of chest x-rays at a health centre between an expert in x-ray diagnosis, a family doctor, and a radiologist.

Design. Transversal study.

Setting. Primary care. Cartuja Health Centre, Granada, Spain.

Participants. Patients at the Cartuja Health Centre who had a simple chest x-ray in 2002. 150 studies were chosen by simple randomised sampling. Two were rejected because they dealt with a repeat examination of the same patient and seven because they did not reach minimum quality. The final sample was 141 x-rays.

Main measurements. The observers filled in independently, for each examination, an 11-item protocol on the technical quality of the images. The kappa index between pairs of observers was calculated for each item, as was the overall kappa index.

Results. 96% of the examinations were conducted with large x-ray plates (35-43). There was only acceptable or good concordance between the 3 observers in 2 questions (kappa, 0.559-0.858). In 5 questions concordance was homogeneously low (kappa, 0.034-0.375). In some questions there was a strong discrepancy between the appraisal of the expert and that of the 2 other observers.

Conclusions. General concordance can be considered low, although it is somewhat greater between the radiologist and the family doctor than between either of these and the expert. This poses the need to improve professional training in evaluation of the technical quality of images. In addition, there was unnecessary expense in large-size x-ray plates.

Key words: Health centre. Quality control. Image diagnosis. Chest x-ray. Evaluation.

^aServicio de Radiodiagnóstico. Hospital Universitario San Cecilio. Granada. España.

^bCentro de Salud de Cartuja. Granada. España.

^cDepartamento de Medicina Preventiva y Salud Pública de la Universidad de Granada. Granada. España.

Correspondencia:
F. Quesada Jiménez.
Concepción, 30. 18009 Granada.
España.
Correo electrónico:
fqesadaj@meditex.es

Manuscrito recibido el 5 de julio de 2004.

Manuscrito aceptado para su publicación el 14 de marzo de 2005.

Introducción

El diagnóstico por imagen es un complemento importante, aunque no indispensable, en el proceder asistencial en atención primaria. La radiografía del tórax es la exploración radiológica más utilizada y representa alrededor del 25-40% de éstas¹⁻³. Sigue vigente a pesar del tiempo transcurrido desde su incorporación al arsenal diagnóstico. Es una técnica simple de realizar, requiere poco tiempo, es accesible y barata, y la dosis de radiación recibida es mínima. Por tanto, cualquier clínico debería estar capacitado para interpretarla correctamente. Sin embargo, en la práctica clínica es habitual que no se pueda aprovechar del todo este recurso debido a que el estudio presenta una deficiente calidad técnica (éste es un comentario generalizado entre los profesionales), sobre todo si tenemos en cuenta que la radiografía de tórax no suele ser informada al médico de familia.

Es en el entorno técnico en el que con más frecuencia se ve comprometida la calidad radiográfica, tanto en su vertiente intrínseca o técnica como en su apariencia o percepción. Además, es el momento en el que, de forma más clara, se pone en juego la pericia del personal de un servicio de radiodiagnóstico⁴. Los programas de investigación sobre Protección Radiológica de la Comisión Europea han abordado en numerosos proyectos la evaluación de los riesgos del radiodiagnóstico. En particular, se han estudiado las características específicas que debe poseer una imagen radiológica concreta para satisfacer los requisitos diagnósticos, así como la dosis de radiación requerida usualmente para conseguirla. Estos estudios han cristalizado en un documento comunitario sobre «Criterios de calidad de las imágenes en radiodiagnóstico»⁵, cuya importancia es fundamental para poner en perspectiva la calidad del acto radiológico en relación con la protección radiológica del paciente. La evaluación de la calidad de las imágenes por métodos físicos puede sustituirse con ventajas por la observación de estructuras anatómicas apropiadas y detalles de imagen característicos, en la que el propio facultativo juzga imágenes de pacientes reales⁶.

Para el cumplimiento de las recomendaciones de la Comisión Internacional de Protección Radiológica, y de acuerdo con lo establecido en la legislación de la Comunidad Europea (Directiva 84/466/EURATOM) y el Real Decreto 783/2001 sobre protección radiológica de las personas sometidas a exámenes o tratamientos médicos, se exige que todas las exposiciones a radiaciones en un acto médico estén médicamente justificadas, se realicen con la dosis más baja posible y que se lleven a cabo bajo la responsabilidad de un médico. En el caso de exámenes directamente relacionados con la enfermedad de un paciente, la justificación de la prueba puede ser responsabilidad tanto del médico que la prescribe como

del que la realiza. En cualquier caso, es imperativo que la decisión se base en una evaluación correcta de las indicaciones del examen, del rendimiento que se espera de éste y de la probabilidad de que los resultados afecten al diagnóstico y al posterior tratamiento⁷.

Por otra parte, desde hace años, muchos estudios han ido aportando información acerca de la variabilidad entre patrones de práctica médica⁸⁻¹¹. Esta variabilidad también afecta a las observaciones de dos o más médicos cuando examinan a un mismo paciente¹²⁻¹⁵. La interpretación de los resultados de una prueba diagnóstica afecta al pronóstico y al tratamiento del paciente, e influye en la decisión de solicitar otras exploraciones complementarias o interconsultas con otros profesionales de distinto ámbito asistencial.

El conocimiento de las causas de las discrepancias puede contribuir a mejorar la calidad de la práctica médica. Sin embargo, no hemos hallado trabajos realizados en nuestro país con el objeto de estudiar la variabilidad en la interpretación de la calidad técnica del estudio radiológico simple de tórax entre el radiólogo, el técnico especialista en radiología (TER) y el médico de familia. El objetivo de este estudio es valorar el grado de concordancia en la interpretación de la calidad técnica de las radiografías de tórax realizadas en el centro radiológico de referencia de un centro de salud, entre los profesionales implicados con más frecuencia: TER, médico de familia y médico especialista en radiodiagnóstico.

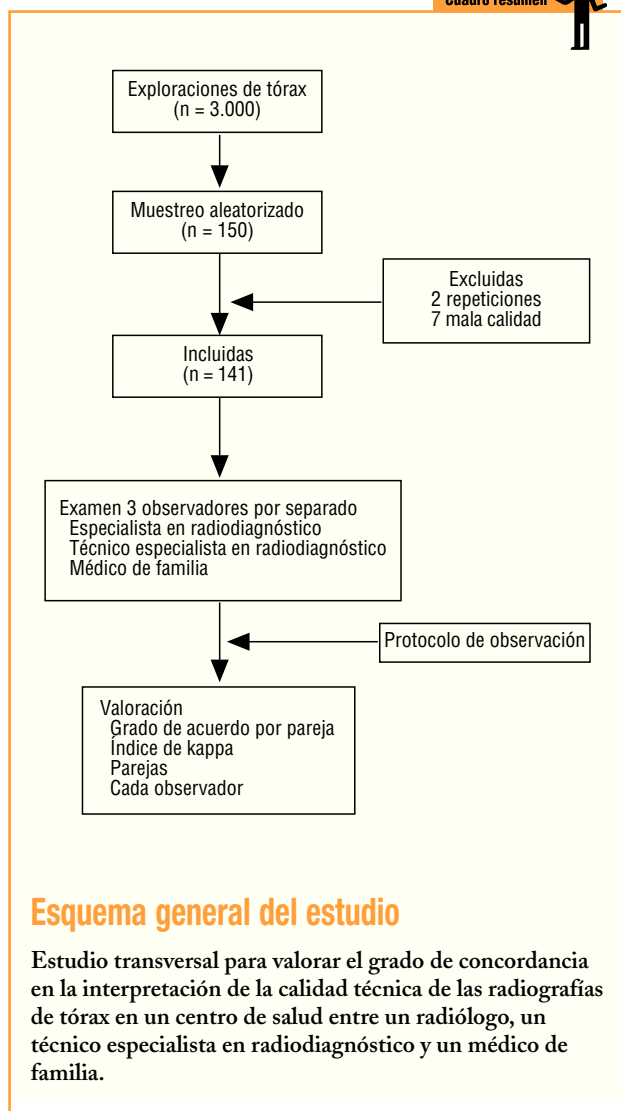
Material y método

Se ha realizado un estudio de concordancia basado en la serie de estudios radiológicos de tórax constituida por las placas realizadas para el Centro de Salud de Cartuja, cuya población de referencia está constituida por unas 17.000 personas que residen en el barrio de Cartuja de Granada.

La población elegible fue cualquier persona adscrita al citado centro de salud a la que se le hubiera realizado un estudio radiológico simple de tórax, con placas anteroposterior (AP) y lateral, durante el año 2002. Dichas exploraciones se realizaron en el servicio de radiodiagnóstico de referencia del centro de salud del que procedían los pacientes. De las cerca de 3.000 exploraciones de tórax recibidas por el centro de salud en el año 2002, mediante muestreo aleatorio simple se seleccionaron 150 estudios radiográficos de tórax. Tras rechazar 9 estudios (2 correspondientes a exploraciones repetidas en los mismos pacientes y 7 a estudios sin una calidad mínima), la muestra final del estudio quedó constituida por 141 exploraciones.

Las radiografías fueron examinadas de forma independiente por 3 observadores entre enero y junio de 2003: un especialista en radiodiagnóstico (con una experiencia de 12 años de servicio), un TER con 4 años de experiencia profesional y un médico de familia (con 16 años de experiencia, más 4 años como alumno interno en el servicio de radiodiagnóstico). Los 3 observadores conocían su participación en el estudio en el momento de evaluar las placas y con anterioridad se había realizado una reunión pre-

Material y métodos
Cuadro resumen



via en la que se pusieron en común los criterios de valoración. El examen se realizó de acuerdo con un protocolo de observación, modificado por el Servicio de Radiofísica del Hospital Universitario San Cecilio a partir del inicialmente desarrollado por una comisión de la Comunidad Europea sobre criterios de calidad para imágenes radiográficas⁵. Dicho protocolo, que se presenta en el anexo 1, constaba de 11 preguntas de respuesta dicotómica (sí/no): 8 para valorar la calidad de la proyección AP y 3 para la lateral.

Para cada pregunta del protocolo se valoró inicialmente el grado de acuerdo entre parejas de observadores, mediante la obtención de las proporciones de acuerdo (número de coincidencias entre el total de observaciones). Con posterioridad se calculó el índice kappa, tanto para cada pareja de observadores como para los 3 evaluadores conjuntamente¹⁶. El índice kappa proporciona una estimación del grado de acuerdo descontando el que se obtendría simplemente por azar¹⁷. Los índices < 0,4 sugieren un grado de acuerdo bajo, mientras que los > 0,7 son indicativos de un buen grado de concordancia. Como variable adicional se registró el tamaño de la placa radiográfica utilizada para cada proyección.

El análisis de los datos se realizó con el paquete estadístico Stata¹⁸.

Resultados

En relación con el tamaño de placa (tabla 1), el 96% de las exploraciones fue de tamaño 35 43. Tan sólo el 3% de las exploraciones fue de tamaño 35 35. Las dos proyecciones de una exploración (el 1%) fueron realizadas con placas de tamaño 30 40. En la tabla 2 se presentan, para cada pregunta y para cada par de observadores, el porcentaje de acuerdo global observado y el índice kappa, así como el índice de kappa global obtenido conjuntamente para los 3 observadores. Con respecto a los valores de los índices kappa obtenidos para cada par de observadores, se aprecian grandes discrepancias que dependen tanto de la pregunta como del par considerado. De forma global se pueden diferenciar los siguientes patrones:

1. Hay 2 preguntas, la 3 (escápula fuera de campos pulmonares) y, en menor medida, la 8 (visualización de la zona retrocardíaca), para las que se aprecia una buena concordancia en todas las parejas, con valores que oscilan entre 0,559 y 0,858. En el extremo opuesto, las preguntas 1, 2 y 9 (inspiración profunda en proyección posteroanterior [PA], simetría del tórax e inspiración profunda en proyección lateral) muestran una concordancia muy escasa en todas las parejas (entre 0,034 y 0,369). Una concordancia algo mayor a la anterior, pero también baja (por debajo de 0,4 en las 3 parejas) se aprecia para las preguntas 5 y 11 (reproducción del sistema vascular pulmonar y reproducción visualmente nítida del contorno posterior del corazón, la aorta, el mediastino, la tráquea, el diafragma, el esternón y la columna vertebral).

2. Mención especial merecen las preguntas 4 y 10 (reproducción de toda la caja torácica y brazos levantados por encima del tórax). En ellas se aprecia una fuerte discrepancia entre los bajos valores del índice kappa de las 2 parejas en las que está incluido el TER (entre -0,039 y 0,273), con respecto a la restante pareja, que muestra unos índices aceptables o buenos (0,530 y 0,732).

TABLA 1

Tamaño de las placas obtenidas para cada una de las dos proyecciones de las radiografías de tórax

Proyección	Lateral				Total
	Tamaño placa		30	40	
PA	30	40	1	–	1
	35	35	–	3	3
	35	43	–	3	134
Total			1	6	141

PA: posteroanterior.

TABLA 2 Valores de kappa para la comparación de los ítems del cuestionario entre pares de observadores y entre los 3 observadores conjuntamente

Pregunta*	RX contra MF		RX contra TER		MF contra TER		Global	
	Kappa	IC del 95%	Kappa	IC del 95%	Kappa	IC del 95%	Kappa	p
1	0,18	0,08 0,28	0,26	0,11 0,42	0,34	0,21 0,48	0,22	< 0,001
2	0,29	0,09 0,41	0,11	0,02 0,21	0,03	-0,08 0,15	0,08	0,058
3	0,86	0,69 1,02	0,83	0,67 0,99	0,75	0,59 0,91	0,81	< 0,001
4	0,73	0,57 0,90	-0,04	-0,20 0,12	0,1	0 0,17	0,20	< 0,001
5	0,35	0,19 0,50	0,31	0,15 0,48	0,24	0,10 0,39	0,29	< 0,001
6	0,27	0,11 0,43	0,53	0,37 0,70	0,43	0,26 0,59	0,40	< 0,001
7	0,33	0,20 0,46	0,30	0,15 0,45	0,56	0,42 0,75	0,41	< 0,001
8	0,68	0,51 0,84	0,71	0,55 0,87	0,56	0,40 0,71	0,65	< 0,001
9	0,24	0,13 0,35	0,37	0,21 0,52	0,29	0,15 0,43	0,26	< 0,001
10	0,53	0,37 0,69	0,27	0,13 0,42	0,20	0,08 0,32	0,34	< 0,001
11	0,34	0,18 0,51	0,37	0,22 0,53	0,37	0,23 0,51	0,35	< 0,001

RX: radiólogo; MF: médico de familia; TER: técnico especialista en radiodiagnóstico; IC del 95%: intervalo de confianza del 95%.

*Ver anexo 1.

Discusión

Antes de comentar los resultados más relevantes que se desprenden del presente estudio es necesario subrayar algunas limitaciones metodológicas de su diseño que deben tenerse en cuenta. En primer lugar, es evidente que estos resultados no pueden ser directamente extrapolados a todos los miembros de las 3 categorías profesionales representadas por cada uno de los 3 participantes en el estudio. Sobre todo la experiencia profesional de los observadores, muy superior en el médico de familia (16 años) y en el radiólogo (12 años) con respecto al TER (4 años), creemos que puede haber condicionado de forma manifiesta las concordancias observadas entre ellos.

En este punto conviene comentar la posible influencia que puede tener en los resultados el hecho de que el médico de familia fuera alumno interno de radiodiagnóstico durante 4 años. Creemos que este hecho debería haber condicionado una mayor concordancia con los otros observadores respecto a la que cabía encontrar únicamente por ser médico de familia. Por tanto, consideramos que habla aún más a favor de la hipótesis de baja concordancia entre observadores.

Un segundo aspecto que se debe tener en cuenta es que previamente al estudio no se realizó ninguna otra reunión entre los 3 participantes con el objeto de discutir y acordar los criterios de respuesta a cada pregunta del cuestionario. Ello se debió a que, en principio, se planteó un protocolo de estudio previamente validado para valorar una dimensión en teoría muy objetiva. Aunque indudablemente la concordancia entre observadores podría haber aumentado con el establecimiento de criterios consensuados antes del estudio, es muy probable que nuestros resultados sobrestimen la concordancia que se hubiera obtenido en relación

con la valoración de parámetros de interpretación más subjetivos, como son los referidos al diagnóstico.

Dentro de la baja concordancia general, hay un mayor grado de acuerdo entre el radiólogo y el médico de familia que entre cada uno de éstos y el TER. Una posible explicación podría ser la disposición del médico de familia y del radiólogo para valorar la utilidad clínica diagnóstica de las radiografías, mientras que el TER sólo se fijaría en su calidad técnica. Sería interesante conocer cómo se modificaría la concordancia si se unificaran los criterios de clasificación según la formación y la experiencia previa.

No nos atrevemos a imaginar lo que ocurriría si intentáramos valorar la concordancia entre el médico de familia y el radiólogo en la evaluación diagnóstica de estas radiografías.

Aunque aumentaría de manera importante la complejidad del estudio, sus resultados serían de gran valor, especialmente ahora que se pretende maximizar la calidad de la asistencia médica y se trata de evaluar la capacidad del profesional en el desarrollo de sus competencias. No obstante, este paso ya ha sido planteado en España, pues se ha realizado un estudio para investigar el acuerdo en la interpretación diagnóstica de radiografías PA del tórax entre médicos de atención primaria de los centros de salud docentes de Elche, y radiólogos del Servicio de Radiodiagnóstico del Hospital de Elche. En este estudio se ha encontrado que, aunque el acuerdo global era aceptable y el observado en los médicos formados por vía MIR era bueno, el obtenido por los otros grupos no sobrepasaba un umbral mínimo de aceptabilidad, por lo que es necesaria una formación que eleve la competencia clínica, tal y como lo hace el período MIR¹⁵.

El abuso indiscriminado, e injustificado en la gran mayoría de las exploraciones, de placas radiográficas de gran tamaño (35 43) conlleva un alto gasto, lo que indica una mala praxis y una escasa formación del que realiza la radiografía (¡que no se me «corte»!). Además, con frecuencia el cuerpo del paciente no se encuentra centrado con la placa y, por tanto, con el haz de rayos X, lo que perjudica la calidad del estudio radiográfico.

La conclusión fundamental es que no hay una buena concordancia en la valoración de la calidad técnica de las radiografías de tórax por parte del radiólogo, el médico de familia y el TER de nuestro estudio, lo que nos hace sospechar la necesidad de una formación adecuada en calidad técnica.

Tenemos que volver a insistir en el gasto innecesario por el uso de placas radiográficas de gran tamaño.

Discusión
Cuadro resumen



Lo conocido sobre el tema

- La interpretación de los resultados de una prueba diagnóstica afecta al pronóstico y al tratamiento del paciente, e influye en la decisión de solicitar otras exploraciones complementarias o interconsultas con otros profesionales de distinto nivel asistencial.
- Es importante valorar el grado de concordancia en la interpretación de la calidad técnica de las radiografías de tórax en los profesionales más frecuentemente implicados: técnico especialista en radiología, médico de familia y médico especialista en radiodiagnóstico.
- No hemos encontrado trabajos realizados en nuestro país que estudien la variabilidad en la interpretación de la calidad técnica del estudio radiológico simple de tórax entre radiólogo, técnico especialista en radiología y médico de familia.

Qué aporta este estudio

- El conocimiento de las causas de las discrepancias puede contribuir a mejorar la calidad de la práctica médica.
- A pesar de que se detecta una baja concordancia general, hay mayor grado de acuerdo entre el radiólogo y el médico de familia que entre cada uno de éstos y el técnico especialista en radiología.
- El uso indiscriminado e injustificado, en la gran mayoría de las exploraciones, de placas radiográficas de gran tamaño (35 x 43) origina un elevado gasto, lo que implica una mala praxis y una escasa formación del que realiza la radiografía.

Bibliografía

- Casanova R, Rodríguez R, Arrazola J. El diagnóstico por imagen de la urgencia torácica: un problema actual. MDI. 1987;2:9-16.
- Aberle DR, Hausell D, Hang HK. Current status of digital projection radiography of the chest. J Thorac Imag. 1990;5:10.
- Delgado Nicolás MA, Peces Morate FJ. Análisis de la utilización de la radiología en atención primaria. Aten Primaria. 1996;17:52-6.
- Alastuey F, Barranco C, Marco C, Pérez C, Sánchez J, Pardo J, et al. Programas de calidad integral para servicios de radiodiagnóstico. Radiología. 1993;35:303-7.
- Comisión of European Communities (CEC). Quality criteria for diagnostic radiographic images. Working document XII-173-90. 2nd edition. Luxemburg: CEC; 1990.
- Documento de la Comisión Europea sobre criterios de calidad de imagen en radiodiagnóstico. XII/173/90. Radiología. 1994;36:A1-12.
- Consejería de Salud de la Junta de Andalucía. Técnicas de imagen: proceso de soporte. Sevilla: Consejería de Salud; 2004.
- Cabedo García VR, Cantos Escudero R, Corbella y Jané A, Saturno PJ. Evaluación de la calidad en la solicitud de radiología de tórax en patología crónica en un centro de salud. Aten Primaria. 1994;13:80-4.
- Rockette HE, King JL, Medina JL, et al. Imaging systems evaluation: effect of subtle cases on the design and analysis of receiver operating characteristic studies. AJR. 1995;165:697-83.
- Kobayashi T, Xu XW, MacMahon H, et al. Effect of a computer-aided diagnosis scheme on radiologist' performance in detection of lung nodules on radiographs. Radiology. 1996;199:843-8.
- Giger M, MacMahon H. Imaging processing and computer-aided diagnosis. Radiol Clin North Am. 1996;34:565-96.
- Bergus GR, Franken EA Jr, Koch TJ, et al. Radiologic interpretation by family physicians in an office practice setting. J Fam Pract. 1995;41:352-6.
- Kimme-Smith C, Aberle DR, Sayre JM, et al. Effects of reduced exposure on computed radiography: comparison of nodule detection accuracy with conventional and asymmetric screen-film radiographs of a chest phantom. AJR. 1995;165:269-73.
- Ely JW, Berbaum KS, Bergus GR, et al. Diagnosis left lower lobe pneumonia: usefulness of the «spine sign» on lateral chest radiographs. J Fam Pract. 1996;43:242-8.
- Simó Miñana J, Riquelme Miralles DA. Variabilidad en la interpretación de la radiografía de tórax entre una comunidad médica de atención primaria y sus radiólogos de referencia. Aten Primaria. 1998;21:599-606.
- Fleiss JL. Statistical methods for rates and proportions. 2nd ed. New York: John Wiley & Sons; 1981.
- Hernández I, Porta M, Miralles M, García F, Bolívar F. La cuantificación de la variabilidad en las observaciones clínicas. Med Clin (Barc). 1990;95:424-9.
- StataCorp. Stata Statistical Software: release 8.0. College Station: Stata Corporation; 2003.

ANEXO 1

Criterios de calidad para radiografías de tórax

Posteroanterior tórax

Inspiración profunda: se valorará por la posición de las costillas sobre el diafragma; bien 6 anteriormente o 10 posteriormente

Reproducción simétrica del tórax

Escápula fuera de los campos pulmonares

Reproducción de toda la caja torácica por encima del diafragma

Reproducción del sistema vascular en todo el pulmón, en especial de los vasos periféricos

Reproducción visualmente nítida de la tráquea y bronquios proximales, el contorno del corazón y la aorta

Reproducción nítida del diafragma y los ángulos costofrénicos

Visualización de la zona retrocardíaca del pulmón y del mediastino

Lateral tórax

Realización en inspiración profunda

Brazos levantados por encima del tórax

Reproducción visualmente nítida del contorno posterior del corazón, la aorta, el mediastino, la tráquea, el diafragma, el esternón y la columna vertebral