

37 Congreso
Nacional
CENTRO DE
CONVENCIONES
INTERNACIONALES

Barcelona
22/25
MAYO 2024

seram
Sociedad Española de Radiología Médica

FERM
FUNDACIÓN ESPAÑOLA DE RADIOLOGÍA MÉDICA

RC | RADIOLEGS
DE CATALUNYA

LA UTILIDAD DE LA INTELIGENCIA ARTIFICIAL EN LA INTERPRETACIÓN DE LA RADIOGRAFÍA DE TÓRAX EN URGENCIAS.

Julia Castaños Cortina¹, Carlos Pacios Cerecedo¹, Jaime López Prieto¹,
Francisco Javier Sanz Carrio¹, Ángels Moreno Gutierrez¹,
Elia Perez Fernández¹, Jose Martel Villagran¹

¹Hospital Universitario Fundación Alcorcón (HUFA)

INTRODUCCIÓN:

La radiografía simple de tórax es la técnica de imagen más utilizada en el estudio de enfermedades torácicas, siendo en la mayoría de casos la primera prueba diagnóstica y muchas veces la única. Su alta rentabilidad es evidente, permite obtener gran cantidad de información con un precio y tiempo de adquisición muy bajo respecto a otras pruebas radiológicas. Basándose en estudios estadísticos actuales, en los hospitales de tercer nivel hasta un 19% de todas las pruebas de imagen realizadas, son radiografías torácicas (1).

Al ser un método insustituible en el examen inicial de las patologías torácicas, juega un papel fundamental en el servicio de urgencias, donde prima la rapidez y eficacia (2).

Sin embargo, para que esta herramienta sea de utilidad requiere una interpretación adecuada, preferiblemente por parte del médico radiólogo especializado. Lamentablemente, debido al incremento de la carga asistencial y la complejidad de los estudios a los que se enfrentan los especialistas de la unidad de diagnóstico por imagen a diario, muchas veces, resulta imposible informar todas las placas de tórax o se hace con retraso, lo que entorpece la actividad en urgencias.

Por otro lado, en los últimos años se está presenciando el auge de una tecnología que está cambiando la manera de entender el mundo, la inteligencia artificial (IA).

Aplicado al campo de la medicina y más concretamente de la radiología, aparece la creación de modelos centrados en el reconocimiento de patrones de imagen.

La implementación de servicios de IA en entornos hospitalarios es una tendencia emergente, pero aún hay pocos estudios que analicen su efectividad en la interpretación de radiografías torácicas.

Además, los resultados de los estudios existentes son contradictorios, algunos sugieren que la IA puede ser útil para este propósito (3-4), mientras que otros cuestionan su precisión (5-7).

Esta discrepancia subraya la necesidad de investigaciones adicionales para comprender mejor el papel de la IA en este campo y su impacto en la práctica clínica.

OBJETIVOS:

Evaluar la capacidad diagnóstica de un sistema de inteligencia artificial en la detección de las principales patologías en radiografías simples de tórax.

MATERIAL Y METODOS:

Diseño:

Se llevó a cabo un **estudio observacional transversal en un solo centro**, el Hospital Universitario Fundación Alcorcón, en el que se revisaron de forma retrospectiva radiografías simples de tórax solicitadas desde el servicio de urgencias sin importar el motivo de petición en un periodo de tres meses, junio, julio y agosto de 2023.

Criterios de inclusión y exclusión:

Con el propósito de asegurar una mayor validez externa se incluyeron todas las radiografías torácicas de pacientes mayores de 18 años.

Únicamente se excluyeron del estudio las pruebas de imagen de mala calidad y las radiografías portátiles.

Una radiografía se considera de mala calidad cuando presenta defectos técnicos que dificultan su adecuada interpretación, como artefactos de movimiento, superposición de estructuras, exposición inadecuada o problemas de contraste y nitidez, en este estudio fue a criterio del radiólogo.

Hallazgos valorados:

El rendimiento de la IA en este estudio se evalúa a través de la capacidad diagnóstica de las siguientes patologías torácicas:

- Nódulo pulmonar
- Derrame pleural
- Síndrome alveolar
- Neumotórax
- Masa mediastínica

Se considera el **informe del radiólogo como Gold estándar (GS)** en el diagnóstico de las enfermedades mencionadas previamente.

Realización de pruebas de imagen e informes correspondientes:

Cuando se solicita una radiografía de tórax desde el servicio de urgencias, el paciente se dirige a la unidad de radiodiagnóstico para realizar el estudio, el cual puede incluir diferentes proyecciones según el contexto clínico del paciente. Una vez obtenidas las imágenes radiográficas, estas son cargadas en el sistema informático PACS (Picture Archiving and Communication System).

Cada radiografía de tórax fue analizada tanto por el radiólogo como por la inteligencia artificial.

El mecanismo de análisis de la IA es el siguiente:

Se utilizó un PACS implementado con un sistema de detección asistida por computadora (CAD) basado en inteligencia artificial conocido como AI-PACS.

Este sistema analiza automáticamente las imágenes y genera un informe que describe la "positividad", "duda" o "negatividad" de cada una de las cinco patologías estudiadas.

El AI-PACS superpone los resultados del CAD basado en IA en las radiografías de tórax poco después de su adquisición, lo que permite localizar los hallazgos detectados. La "positividad" de una lesión se indica mediante un cuadro formado por una línea continua de color naranja, mientras que la "duda" se representa con un cuadro formado por una línea discontinua de color naranja. La "negatividad" se muestra con ausencia de marcado en la imagen.

El mecanismo de análisis del radiólogo es el siguiente:

Un radiólogo examina minuciosamente cada una de las imágenes desde el PACS y elabora un informe radiológico basado en su análisis, evaluando la "presencia" o "ausencia" de cada una de las patologías estudiadas.

En nuestro hospital, las radiografías de tórax se informan, por el servicio de diagnóstico de imagen, tanto en el horario laboral ordinario, como durante los turnos de guardias.

Todos los informes, los generados por la IA y los elaborados por el radiólogo, se almacenan en el RIS (Sistema de Información Radiológica) /PACS, lo que permite su consulta a posteriori.

Recogida de datos:

Tres residentes recogieron los datos de los informes emitidos por los radiólogos y por la IA de forma retrospectiva, para posteriormente asignar a cada serie radiográfica, un verdadero positivo (VP), verdadero negativo (VN), falso positivo (FP) o falso negativo (FN).

En nuestro estudio para un primer análisis de datos, se consideraron como positivos, aquellos hallazgos que el soporte de IA categorizaba como “duda”.

Se define una serie radiográfica como el conjunto de proyecciones complementarias realizadas en un mismo estudio. Por ejemplo, si se llevan a cabo una proyección postero-anterior (PA) y una lateral (L) para evaluar la presencia o ausencia de síndrome alveolar en un paciente, estas dos imágenes se considerarán como una única serie radiográfica.

Cuando en la imagen proporcionada por la IA se detectaban dos o más hallazgos en una misma localización anatómica, únicamente se tuvo en cuenta el de mayor probabilidad de acierto.

Análisis de datos:

Se describe la distribución de la edad mediante media y desviación estándar (DE), y la distribución del sexo y las proyecciones mediante frecuencias absolutas y relativas.

Para responder al objetivo principal se calculan los **índices de validez**, sensibilidad (S) y especificidad (E), y **de utilidad**, valor predictivo positivo (VPP) y valor predictivo negativo (VPN), de la IA, considerando el informe radiológico como GS. Además se evalúa la concordancia entre ambos métodos mediante el índice Kappa. Para interpretar el índice de concordancia se consideran los valores mostrados en la Tabla 1 a continuación:

Concordancia	Valores del índice Kappa
Muy débil	0 – 0,2
Débil	0,21 – 0,4
Moderada	0,41 – 0,6
Buena	0,61 – 0,8
Muy buena	0,81 - 1

Tabla 1. Valores de concordancia (Índice Kappa)

Todos los resultados se acompañan del correspondiente intervalo de confianza al 95% (IC95%).

Para valorar el impacto que los casos marcados como “duda” tienen en la utilidad de la IA, se realiza un análisis de sensibilidad adicional considerando la valoración dudosa como negativa.

RESULTADOS:

Índice de siglas:

- ✓ Inteligencia Artificial (IA)
- ✓ Gold estándar (GS)
- ✓ Picture Archiving and Communication System (PACS)
- ✓ Detección asistida por computadora (CAD)
- ✓ Sistema de Información Radiológica (RIS)
- ✓ Postero-anterior (PA)
- ✓ Antero-posterior (AP)
- ✓ Lateral (L)
- ✓ Desviación estándar (DE)
- ✓ Verdadero positivo (VP)
- ✓ Verdadero negativo (VN)
- ✓ Falso positivo (FP)
- ✓ Falso negativo (FN)
- ✓ Sensibilidad (S)
- ✓ Especificidad (E)
- ✓ Valor predictivo positivo (VPP)
- ✓ Valor predictivo negativo (VPN)
- ✓ Análisis de sensibilidad (AS)
- ✓ Lóbulo inferior derecho (LID)
- ✓ Lóbulo superior derecho (LSD)

Se analizaron **858 series radiográficas de tórax** pertenecientes a **783 pacientes** atendidos en el servicio de urgencias.

De un total de 783 pacientes, 362 eran mujeres, lo que representa un 46,2%. La edad media fue de 65 años (DE 19).

De las 858 pruebas de imagen, 673 correspondieron a proyecciones PA-L, 183 a proyecciones AP y las 9 restantes a proyecciones especiales.

Nódulo pulmonar:

De los estudios que la IA refiere como “Nódulo pulmonar”, **47 son VP** (31 considerados como “positivos” y 16 como “duda”) y **37 FP** (16 “positivos” y 21 “duda”) (Tabla 2).

Tabla de contingencia		Radiólogo (GS)		Total
		Positivo	Negativo	
IA	Positivo + Duda	47	37	84
	Negativo	2	772	774
Total		49	809	858

Tabla 2. Tabla de contingencia nódulo pulmonar.

La IA detecta la presencia de nódulo pulmonar en comparación con el radiólogo con una S del 95,90% y una E del 95,40%, un VPP del 56% y un **VPN del 99,70%** (Tabla 3 y Figura 1).

	Índices	Valor	IC95%	
Nódulo pulmonar	S	95,90%	86%	99,50%
	E	95,40%	93,80%	96,80%
	VPP	56%	44,70%	66,80%
	VPN	99,70%	99,10%	100%

Tabla 3. Resultados nódulo pulmonar.

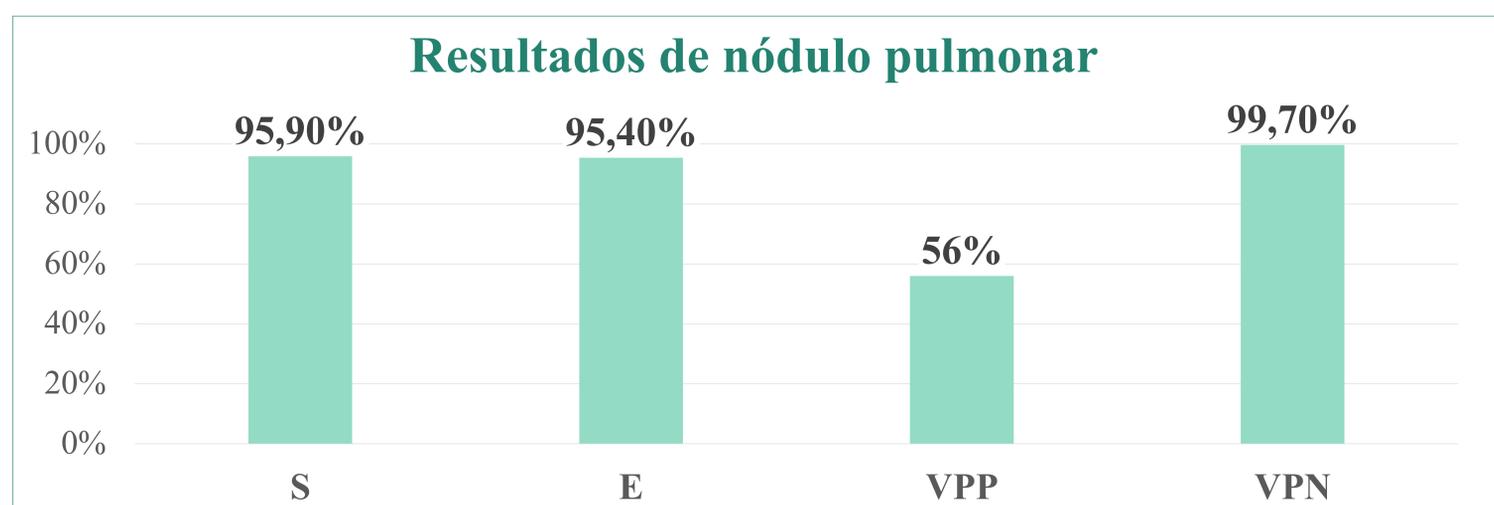


Figura 1. Resultados nódulo pulmonar

El grado de concordancia es **bueno** para el nódulo pulmonar con un índice Kappa de 0,68 (IC95%: 0,59-0,78).

Derrame pleural:

De los estudios que la IA refiere como “Derrame pleural”, **129 son VP y 30 FP**. No se recopiló ninguna placa en la que la IA indicase duda para esta patología (Tabla 4).

Tabla de contingencia		Radiólogo (GS)		Total
		Positivo	Negativo	
IA	Positivo + Duda	129	30	159
	Negativo	4	695	700
Total		133	726	858

Tabla 4. Tabla de contingencia derrame pleural.

La IA detecta la presencia de derrame pleural en comparación con el radiólogo con una S del 97% y una E del 95,90%, un VPP del 81% y un **VPN del 99,40%** (Tabla 5 y Figura 2).

	Índices	Valor	IC95%	
Derrame pleural	S	97,00%	93%	99,20%
	E	95,90%	94,10%	97,20%
	VPP	81%	74,20%	86,90%
	VPN	99,40%	98,50%	100%

Tabla 5. Resultados derrame pleural.

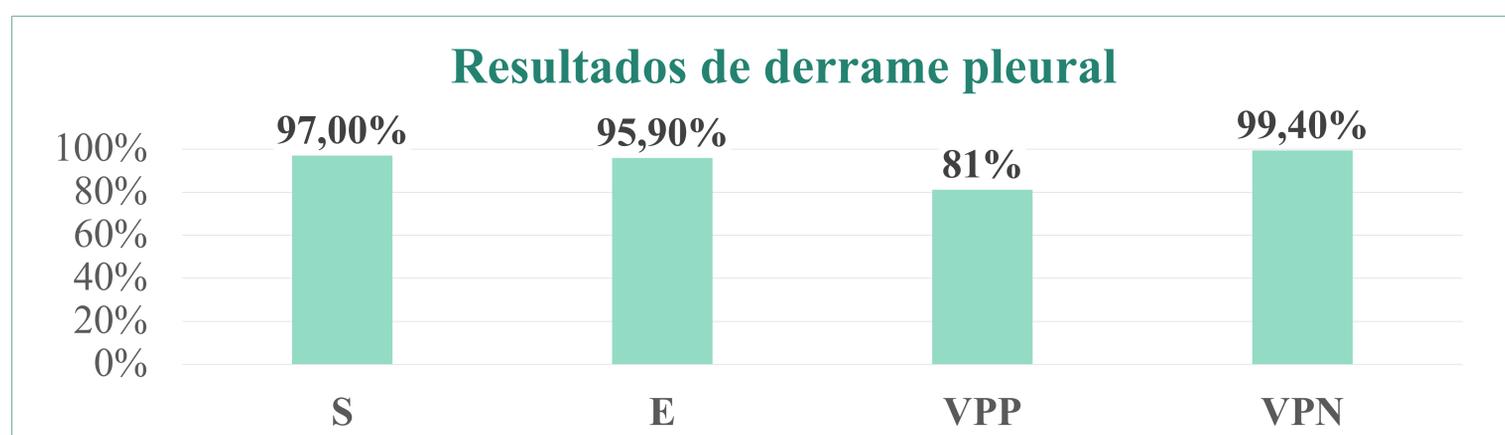


Figura 2. Resultados de derrame pleural.

El grado de concordancia es **muy bueno** para el derrame pleural con un índice Kappa de 0,86 (IC95%: 0,81-0,91).

Síndrome alveolar:

De los estudios que la IA refiere como “Síndrome alveolar”, **88 son VP** (73 “positivos” y 15 “duda”) y **57 FP** (22 “positivos” y 35 “duda”) (Tabla 6).

Tabla de contingencia		Radiólogo (GS)		Total
		Positivo	Negativo	
IA	Positivo + Duda	88	57	145
	Negativo	2	711	713
Total		90	768	858

Tabla 6. Tabla de contingencia síndrome alveolar.

La IA detecta la presencia de derrame pleural en comparación con el radiólogo con una S del 97,80% y una E del 92,60%, un VPP del 61% y un **VPN del 99,70%** (Tabla 7 y Figura 3).

	Índices	Valor	IC95%	
Síndrome alveolar	S	97,80%	92%	99,70%
	E	92,60%	90,50%	94,30%
	VPP	61%	52,20%	68,70%
	VPN	99,70%	99,00%	100%

Tabla 7. Resultados síndrome alveolar

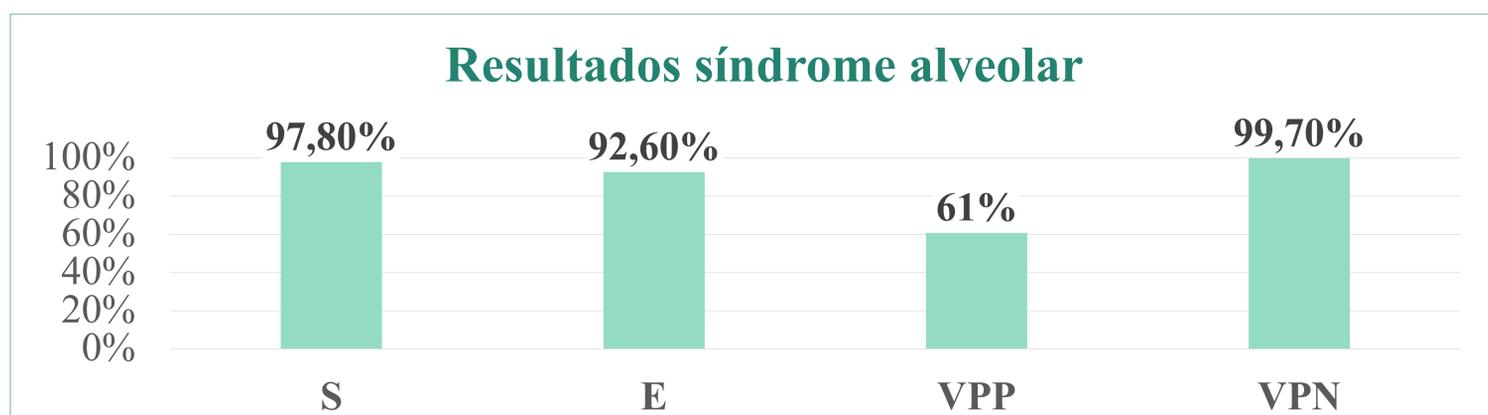


Figura 3. Resultados síndrome alveolar.

El grado de concordancia es **bueno** para el síndrome alveolar con un índice Kappa de 0,71 (IC95%: 0,64-0,78).

Neumotórax:

De los estudios que la IA refiere como “Neumotórax”, **5 son VP** (4 “positivos” y 1 “duda”) y **10 FP** (6 “positivos” y 4 “duda”) (Tabla 8).

Tabla de contingencia		Radiólogo (GS)		Total
		Positivo	Negativo	
IA	Positivo + Duda	5	10	15
	Negativo	0	843	843
Total		5	853	858

Tabla 8. Tabla de contingencia neumotórax.

La IA detecta la presencia de derrame pleural en comparación con el radiólogo con una S del 100% y una E del 98,80%, un VPP del 33,3% y un VPN del 100% (Tabla 9 y Figura 4).

Neumotórax	Índices	Valor	IC95%	
	S	100,0%	47,8%	100,0%
E	98,8%	97,9%	99,4%	
VPP	33,3%	11,8%	61,6%	
VPN	100,0%	99,6%	100,0%	

Tabla 9. Resultados neumotórax.

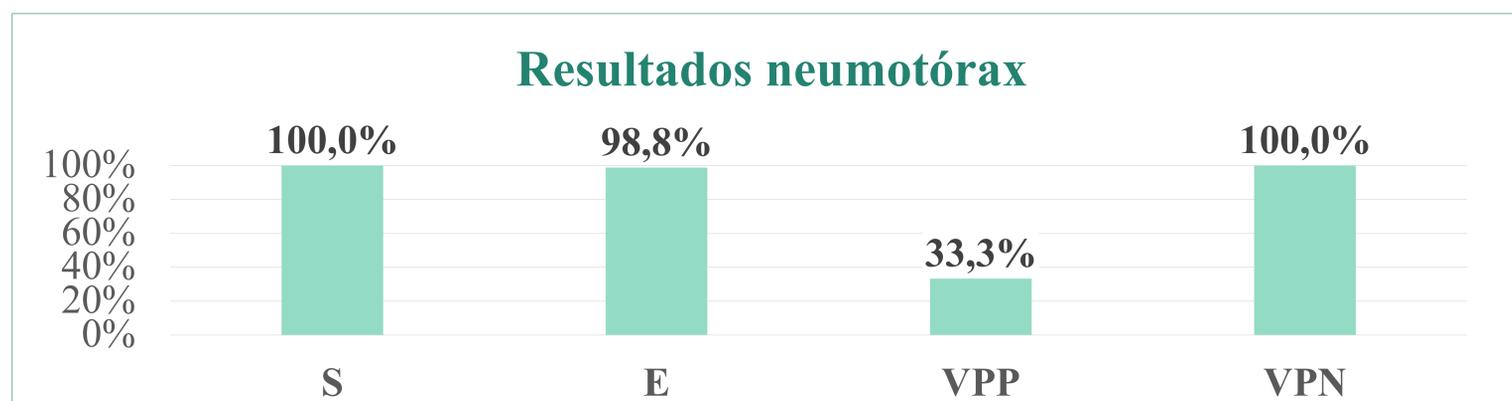


Figura 4. Resultados neumotórax.

El grado de concordancia es **moderado** para el neumotórax con un índice Kappa de 0,50 (IC95%: 0,23-0,76).

Masa mediastínica:

De los estudios que la IA refiere como “Masa mediastínica”, **9 son VP** y **67 FP** (20 “positivos” y 47 “duda”) (Tabla 10).

Tabla de contingencia		Radiólogo (GS)		Total
		Positivo	Negativo	
IA	Positivo + Duda	9	67 (47)	76
	Negativo	1	781	782
Total		10	848	858

Tabla 10. Tabla de contingencia masa mediastínica.

La IA detecta la presencia de derrame pleural en comparación con el radiólogo con una S del 90% y una E del 92,10%, un VPP del 11,8% y un **VPN del 99,9%** (Tabla 11 y Figura 5).

	Índices	Valor	IC95%	
Masa mediastínica	S	90,0%	55,5%	99,7%
	E	92,1%	90,1%	93,8%
	VPP	11,8%	5,6%	21,3%
	VPN	99,9%	99,3%	100,0%

Tabla 11. Resultados masa mediastínica

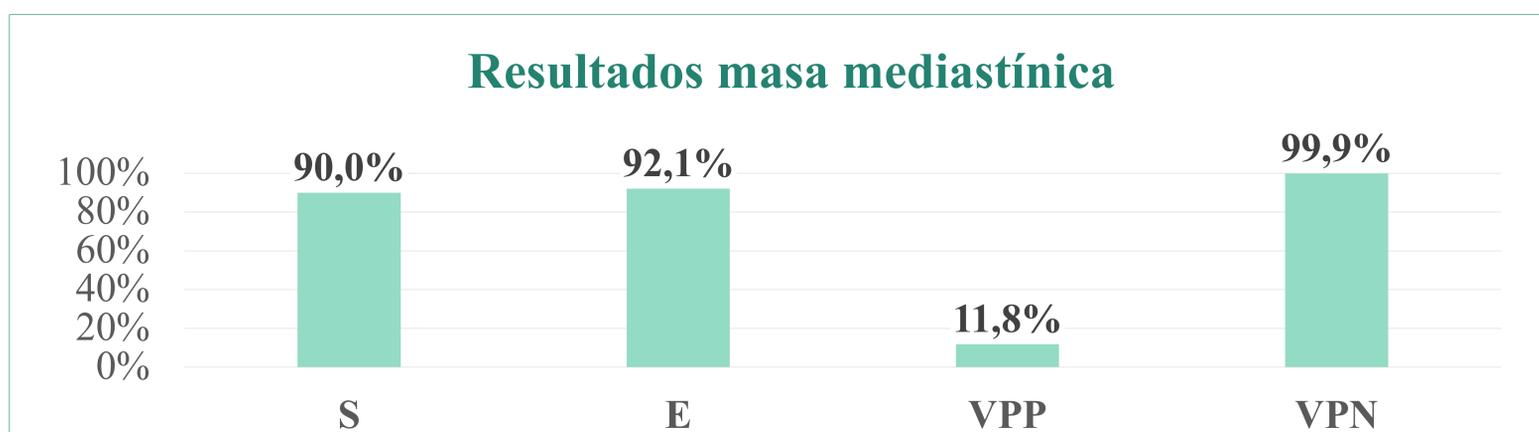


Figura 5. Resultados masa mediastínica.

El grado de concordancia es **muy débil** para la masa mediastínica con un índice Kappa de 0,20 (IC95%: 0,08-0,30).

Resumen de las cinco categorías:

La IA presenta **VPN similares** en todas las patologías con valores superiores al **99,4%**, sin embargo, los **VPP son más heterogéneos**, siendo el VPP del derrame pleural el más elevado, del 81%. (Figura 6)

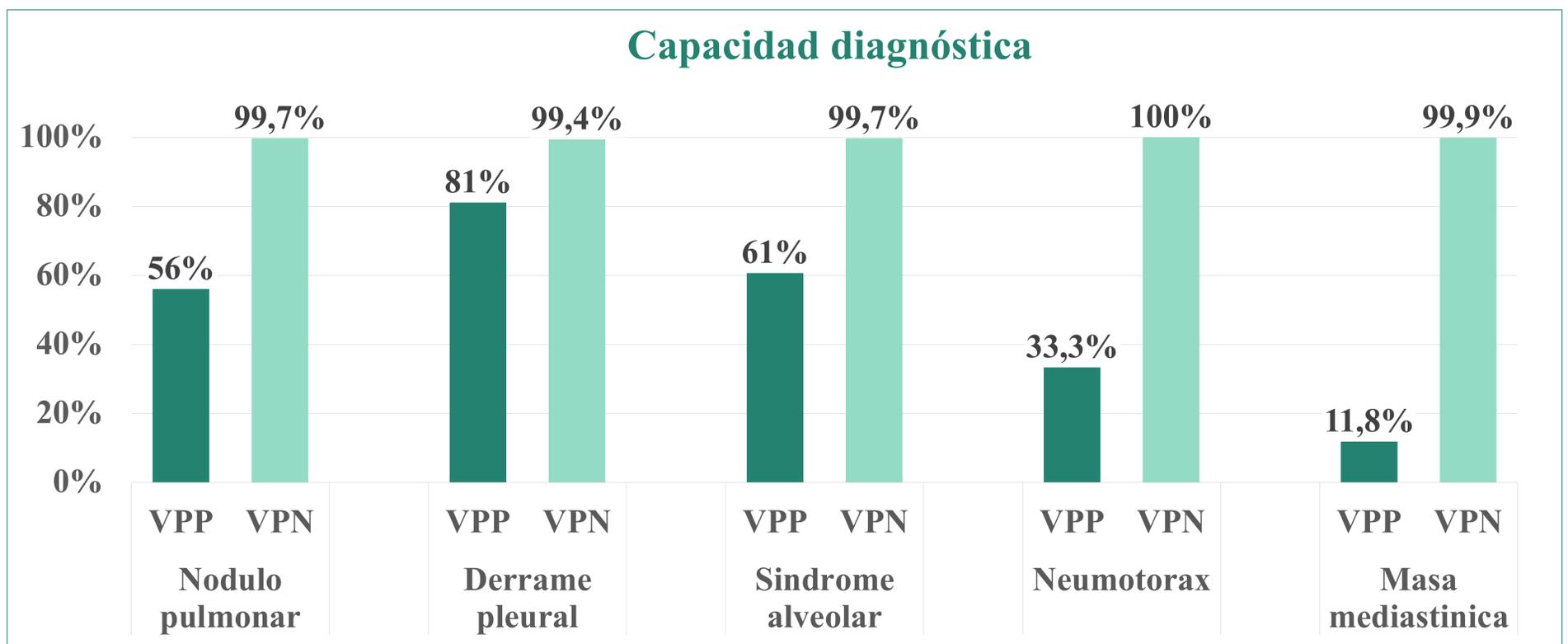


Figura 6. Resumen de los índices de utilidad de cada una de los hallazgos estudiados.

Análisis de sensibilidad (AS):

En el caso del **derrame pleural** no se recopiló **ninguna placa** en la que la IA **indicase duda** para esta patología.

En cuanto al **nódulo pulmonar**, de los 37 marcados como dudosos, **16 fueron VP** y **21 FP**. Al tener en cuenta los dudosos como negativos, la S es 67,30%, la E 98,00%, el VPP 67%, y el VPN 98,00%. El grado de concordancia es bueno, con un índice Kappa de 0,65 (IC95%: 0,54-0,77).

En el caso del síndrome alveolar, de los 50 marcados como dudosos, **15 fueron VP y 35 FP**. Al tener en cuenta los dudosos como negativos, la S es 78,50%, la E 97,10%, el VPP 77%, y el VPN 97,40%. El grado de concordancia es bueno, con un índice Kappa de 0,75 (IC95%: 0,68- 0,82).

Para el neumotórax, de los 5 dudosos, **1 fue VP y 4 FP**. Al tener en cuenta los dudosos como negativos, la S es 80,0%, la E 99,3%, el VPP 40,0%, y el VPN es 99,9%. El grado de concordancia es moderado, con un índice Kappa de 0,53 (IC95%: 0,22- 0,84).

Comparando con los resultados previos, se observa que el índice más afectado es **la sensibilidad, que disminuye de media un 20%**, debido a que los sujetos dudosos que previamente se clasificaban en VP ahora son FN. Por otro lado, se evidencia una reducción de los FP y en consecuencia un aumento de los VN.

El análisis de sensibilidad muestra resultados distintos para la masa mediastínica, debido a que **todos los casos detectados como “duda”, 47 en total, resultaron ser FP**, de modo que **augmentó la E y VPP** sin alterar la S, ni el VPN. Adicionalmente, el grado de concordancia **pasó de ser muy débil a moderado** con un índice Kappa de 0,45 (0,26-0,64).

Los resultados se muestran a continuación en la Tabla 12.

	Índices de validez	Considerando la duda como positivo	Considerando la duda como negativo (AS)
Nódulo pulmonar	S	95,90%	67,30%
	E	95,40%	98,00%
	VPP	56%	67%
	VPN	99,70%	98,00%
	Kappa	0,684	0,654
Síndrome alveolar	S	97,80%	78,50%
	E	92,60%	97,10%
	VPP	61%	77%
	VPN	99,70%	97,40%
	Kappa	0,712	0,749
Neumotórax	S	100,0%	80,0%
	E	98,8%	99,3%
	VPP	33,3%	40,0%
	VPN	100,0%	99,9%
	Kappa	0,496	0,53
Masa mediastínica	S	90,0%	90,0%
	E	92,1%	97,6%
	VPP	11,8%	31,0%
	VPN	99,9%	99,9%
	Kappa	0,193	0,452

Tabla 12. Resultados considerando la duda como positivo (primera opción de análisis) y duda como negativo (análisis secundario o análisis de sensibilidad).

DISCUSIÓN:

Se considera el **VPN** como el **parámetro de mayor relevancia** al evaluar la utilidad de un sistema de inteligencia artificial en el servicio de urgencias, debido a que la falta de detección de una patología podría ocasionar consecuencias negativas para el paciente.

Este estudio demuestra el **alto VPN** de la IA en la interpretación de las cinco patologías estudiadas siendo, en todas ellas, superior al 99,4% (IC95%: 99 - 100). En consecuencia, se confirma que el sistema permite identificar con **notoria fiabilidad la presencia de radiografías torácicas libres de las enfermedades analizadas**. Nuestros resultados son comparables a otros estudios, como el multicéntrico de Hyunsuk Yoo(8) en que se observó que un sistema de triaje basado en IA logró excluir el 42.9% (2354/5482) de las radiografías normales a expensas de la eliminación errónea de tan solo 3.5% (14/405) y 1.8% (4/227) de las imágenes con anomalías visibles y claramente visibles, respectivamente.

No obstante, se observa que los **VPP varían considerablemente** entre los cinco hallazgos evaluados, oscilando desde un 81% para el derrame pleural (IC95%: 74,2-86,9) hasta un 11% para la masa mediastínica (IC95%: 5,6-21,3). Estos resultados indican la presencia de un **elevado número de falsos positivos**.

En el contexto de la práctica clínica, especialmente en el servicio de urgencias, los FP pueden **generar incertidumbre en el clínico** y, por consiguiente, esto puede llevar a un aumento del número de interconsultas, realización de pruebas de imagen adicionales y prescripción de medicamentos innecesarios.

Por consiguiente, se sugiere que el software de IA sea utilizado como una **herramienta complementaria**, con el propósito de asistir en el diagnóstico o en la priorización de las órdenes de lectura por parte del radiólogo, en lugar de confiar ciegamente los resultados que proporciona.

A través del análisis de sensibilidad se objetiva que al considerar los hallazgos “dudosos” como “negativos” el **VPP aumenta** en todos los casos, a expensas de una **importante disminución de la sensibilidad**, de hasta el 28% en el caso del nódulo pulmonar.

El aumento de los FN que esto conlleva, no es asumible en el servicio de urgencias, lo que refuerza la opción del primer análisis, donde se interpretan los dudosos como positivos.

Como dato a destacar, este estudio demuestra que la interpretación de la herramienta de IA para la masa mediastínica es deficiente, con un VPP de 11% y una concordancia muy débil respecto al radiólogo, índice Kappa de 0,20 (IC95%: 0,08-0,30).

Sin embargo, en el análisis de sensibilidad, los resultados mejoran, ya que el total de los casos marcados como dudosos resultaron ser FP, aumentando el VPP de 11% a 31% y la concordancia pasando a ser moderada con un índice Kappa de 0,45 (IC95%: 0,26-0,64).

La mayoría de falsos positivos hacen referencia **hilios prominentes de origen vascular o lipomatosis mediastínica**, que el radiólogo no considera como hallazgo equivalente a “masa mediastínica” al no tener relevancia clínica y no requerir estudios complementarios.

Se sugiere reconsiderar el término "masa mediastínica" y evaluar la posibilidad de utilizar en su lugar "ensanchamiento mediastínico", ya que podría ser más apropiado para describir los hallazgos detectados.

Nuestro estudio no está exento de limitaciones. Al obtener directamente el análisis de la IA en el PACS junto con la prueba de imagen original y realizar la recogida de datos de forma retrospectiva, **no fue factible llevar a cabo un estudio ciego** con respecto al radiólogo evaluador.

Estudios como el de Michael H. Bernstein(7) han observado que cuando la IA proporciona resultados incorrectos, las tasas de FN y FP entre los radiólogos aumentan, de modo que los resultados de la IA podrían estar sobrestimados. De cara a estudios futuros se debería implementar un sistema de ciego al radiólogo o un grupo control sin-IA para minimizar dicho posible sesgo.

No se tuvo en cuenta el grado de experiencia del radiólogo lector de la prueba diagnóstica.

Aunque no se han incluido en nuestro estudio las técnicas de mala calidad, se observó durante la recogida de datos que la tasa de fallos en estas técnicas es considerablemente mayor. Por ello, una recomendación adicional para mejorar la interpretación de los resultados de la IA es que el sistema identifique cuando la prueba se ha realizado de manera deficiente, de modo que el clínico sea consciente de la posible falta de fiabilidad en los resultados obtenidos.

CONCLUSIÓN:

El presente estudio pone en manifiesto que la integración de un sistema de IA para la interpretación de radiografías torácicas en el servicio de urgencias puede colaborar en la **mejora de la capacidad diagnóstica del clínico**, al **identificar con alta fiabilidad la ausencia de patología**.

No obstante, cabe destacar, que **el número de FP no es despreciable**, de modo que las radiografías consideradas como potencialmente patologías por dicha herramienta **requieren una revisión posterior por un especialista en radiología** para limitar el sobrediagnóstico y así la implementación de intervenciones adicionales innecesarias.

CASOS DE EJEMPLO:

Caso 1: Verdadero positivo para nódulos pulmonares. Falso positivo para derrame pleural.

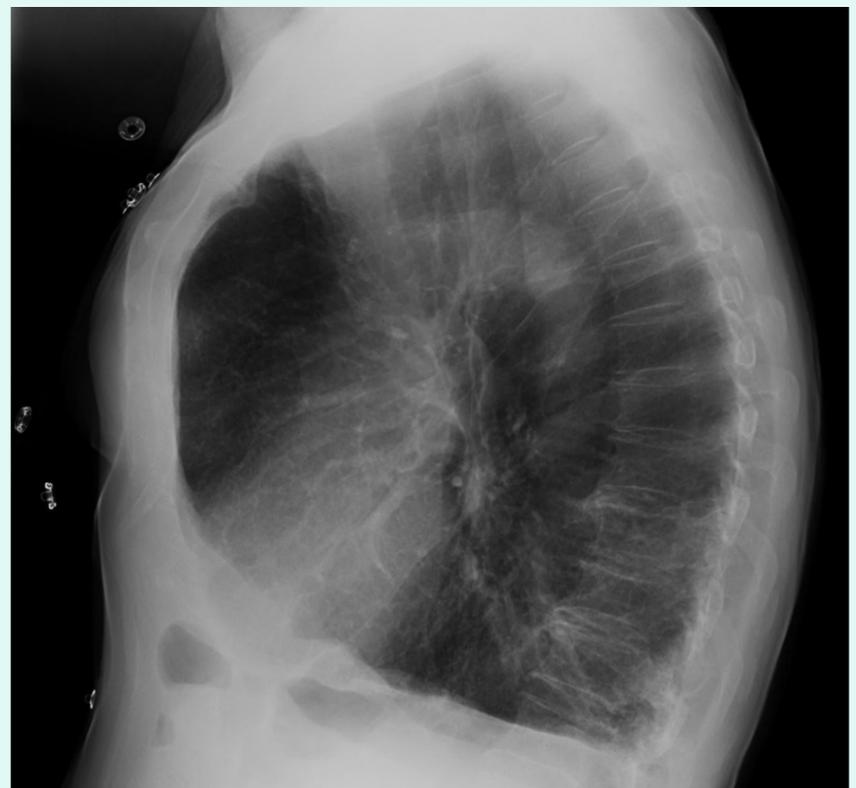
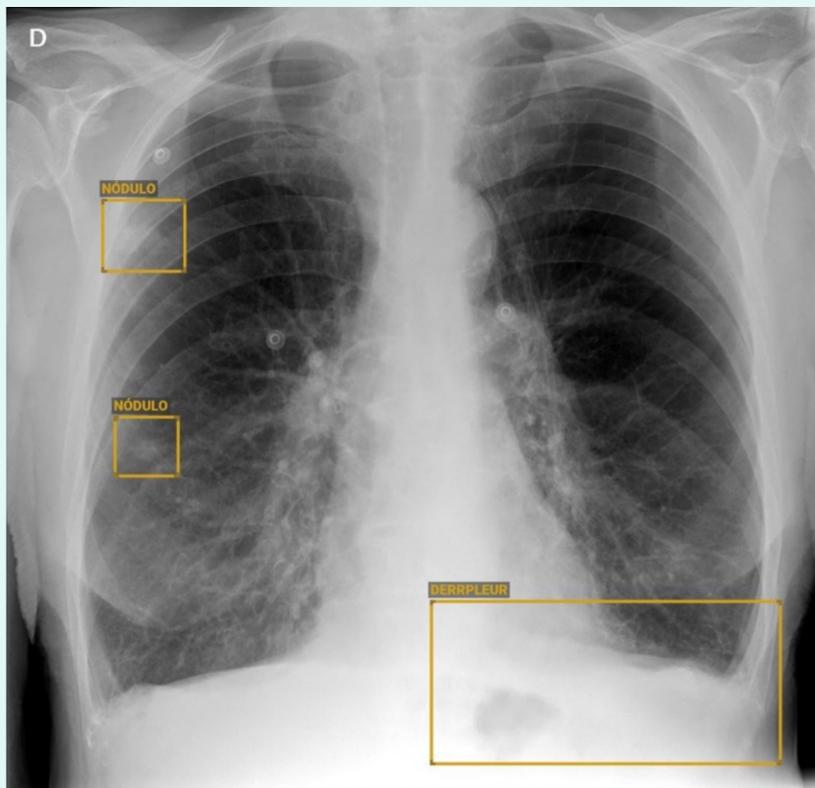


Figura 7 y 8: **Radiografía PA-L analizada por la IA:** marca “positividad” para nódulo pulmonar en campo superior y medio derecho, y para derrame pleural en base izquierda. **Informe del radiólogo:** Hiperinsuflación pulmonar con signos de enfisema. Nódulos pulmonares que en el contexto del paciente son sugestivos de metástasis pulmonares.

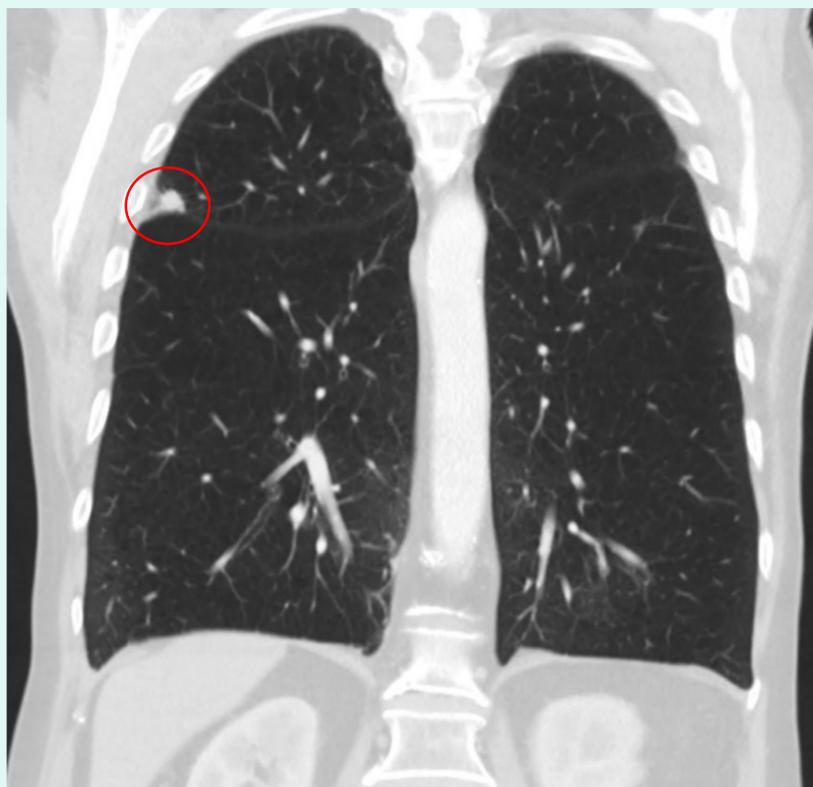


Figura 9 y 10. **TC de tóraco-abdominal con contraste.** El TC refuerza el informe del radiólogo, donde los nódulos pulmonares marcados por la IA corresponden a metástasis pulmonares de adenocarcinoma de colon (VP), pero el derrame pleural ha sido confundido con la horizontalización diafragmática que se produce en la hiperinsuflación pulmonar (FP).

Caso 2: Falso positivo para nódulo pulmonar.

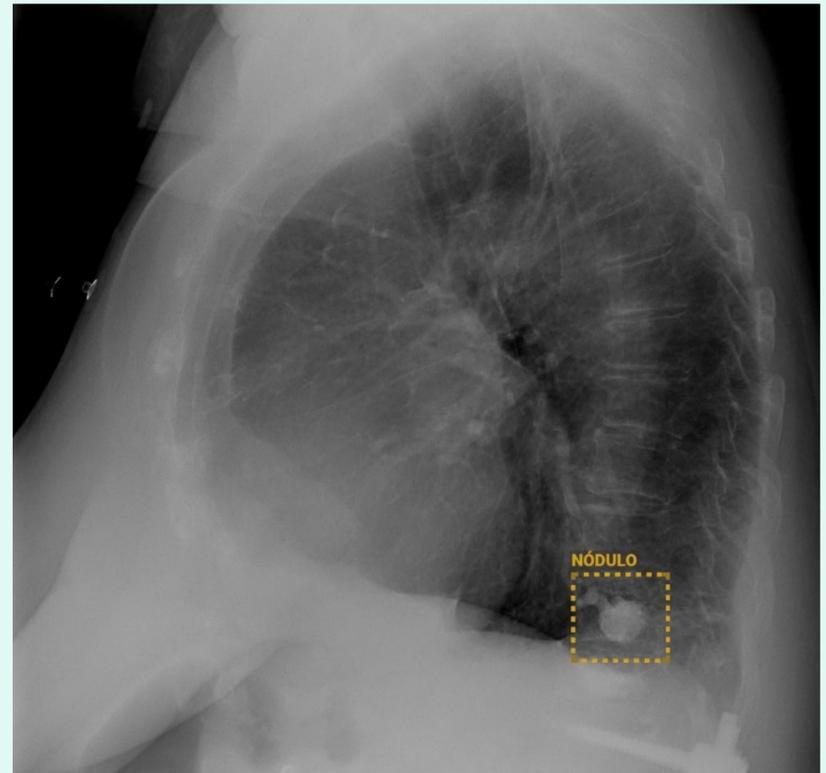


Figura 11 y 12: **Radiografía PA-L analizada por la IA:** marca “duda” para nódulo pulmonar campo pulmonar inferior en la proyección lateral

Informe del radiólogo: Vertebroplastia en cuerpos vertebrales dorsales bajos (T11y T12)

Caso 3: Verdadero positivo para derrame pleural izquierdo.

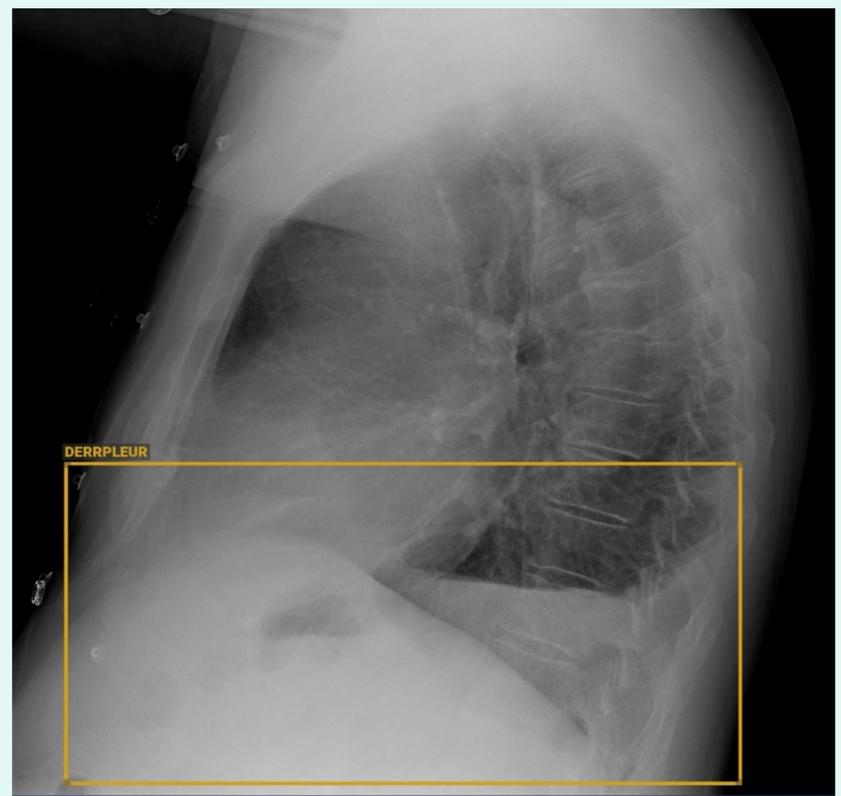
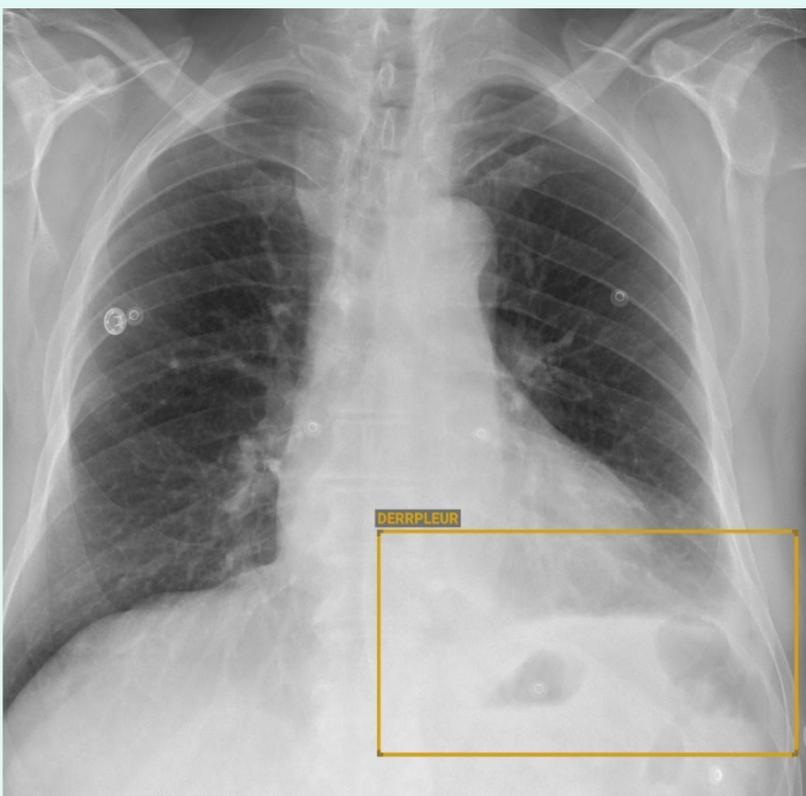


Figura 13 y 14: **Radiografía PA-L analizada por la IA:** marca “positividad” para derrame pleural izquierdo.

Informe del radiólogo: Pinzamiento del seno costofrénico lateral izquierdo en relación con derrame pleural.

Caso 4: Verdadero positivo para síndrome alveolar en las dos primeras imágenes, y posteriormente (mismo paciente tras resolución del cuadro) verdadero negativo para las dos siguientes imágenes.

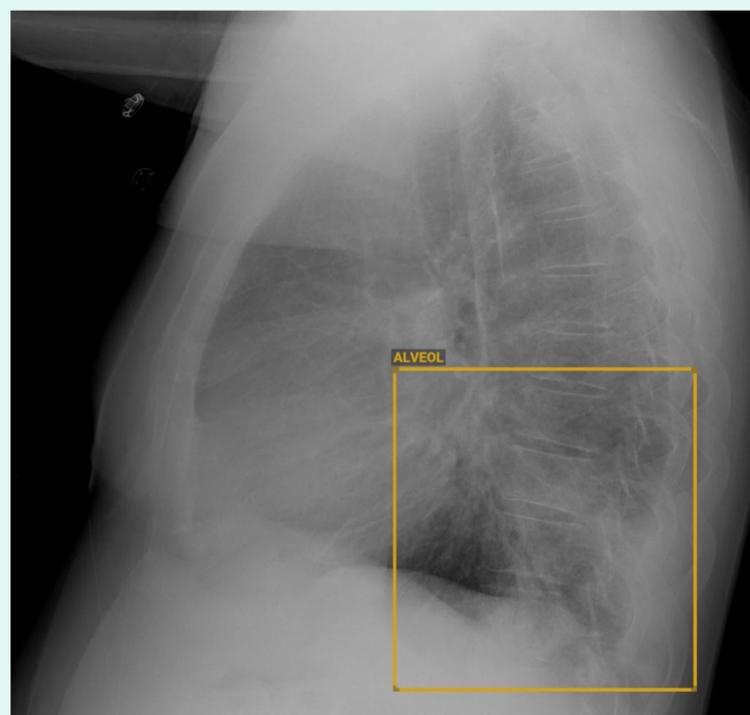
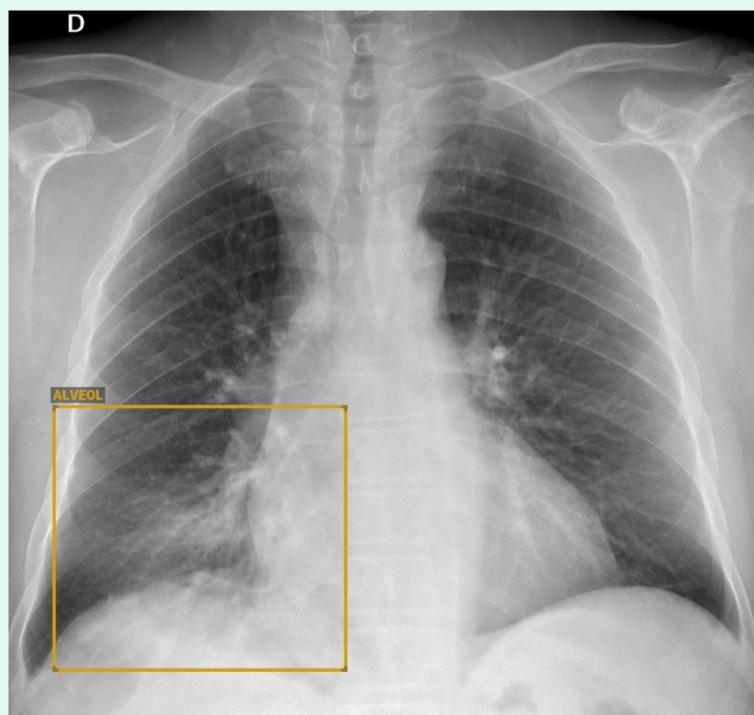


Figura 15 y 16: **Radiografía PA-L analizada por la IA:** marca “positividad” para síndrome alveolar en lóbulo inferior derecho (LID).

Informe del radiólogo: Consolidación pulmonar en el LID sin signos de derrame pleural.



Figura 17, 18 y 19: **Radiografía PA-L analizada por la IA:** La IA no marca ningún hallazgo sobre la serie radiográfica (17 y 18) y aporta una imagen adicional anotando la ausencia de patologías estudiadas (19).

Informe del radiólogo: Resolución de la consolidación pulmonar en el LID.

Caso 5: Falso positivo para síndrome alveolar.

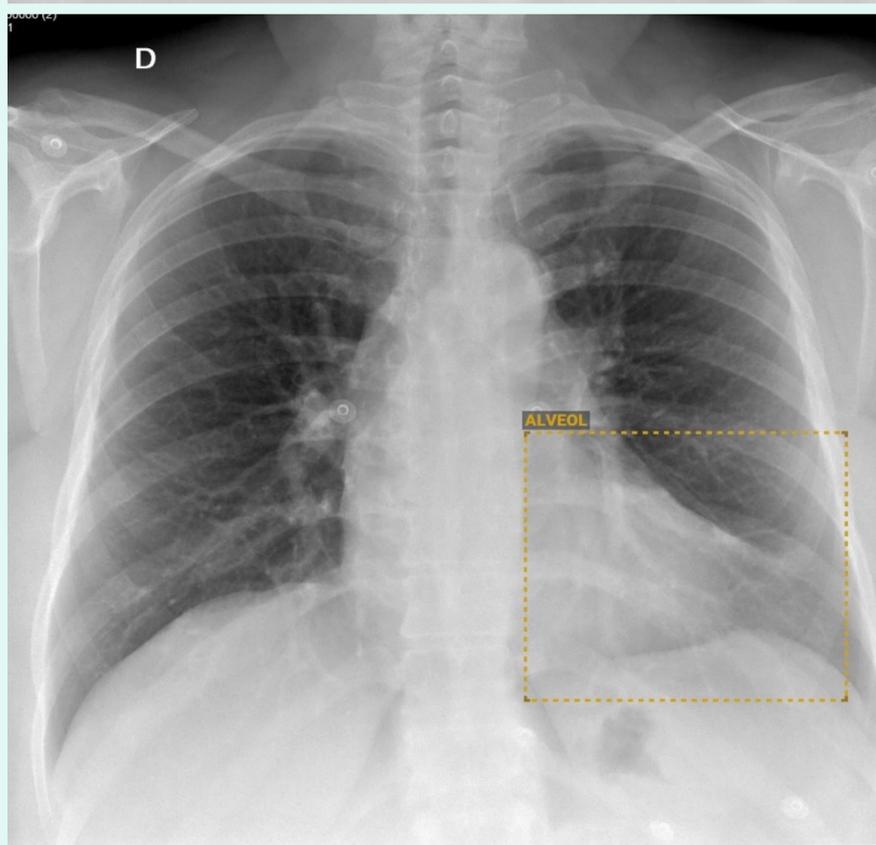


Figura 20, 21, 22 y 23:

Radiografía PA-L original, sin ser procesada por la IA (Figura 20 y 21).

Radiografía PA-L analizada por la IA (Figura 22 y 23): se marca “duda” para síndrome alveolar en campo inferior izquierdo.

Informe del radiólogo: Almohadilla de grasa pericárdica.

Caso 6: Verdadero positivo para neumotórax.



Figura 24 y 25: **Radiografía AP original (24). Radiografía AP analizada por la IA (25):** marca “positividad” para neumotórax en ápex derecho.

Informe del radiólogo: Fractura de 3° a 8° arco costal posterior derecho con pequeño neumotórax asociado ipsilateral.



Figura 26 y 27. **TC torácico.** El TC refuerza el informe del radiólogo, donde el neumotórax corresponde realmente a un pequeño neumotórax derecho de un espesor máximo de 18 mm.

Las fracturas óseas no son valoradas por la IA en radiografías torácicas en nuestro hospital.

Caso 7: Falso negativo de masa pulmonar.

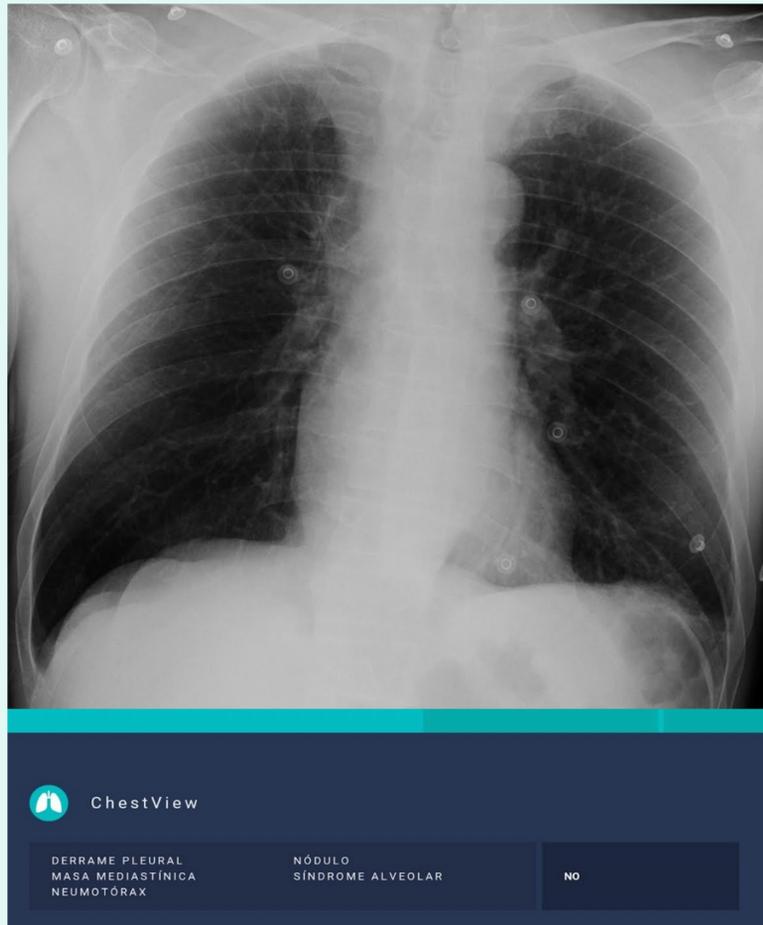


Figura 28 y 29: **Radiografía AP analizada por la IA** como “normal”.

Informe del radiólogo: Se aprecia opacidad una suprahiliar derecha y desplazamiento craneal del hilio derecho, no presentes en estudio previo de agosto 2021, se realiza TC torácico para caracterizar la lesión. (Respecto a la figura 28)



Figura 30: **Radiografía PA original** del mismo paciente **hace dos años** (agosto 2021).

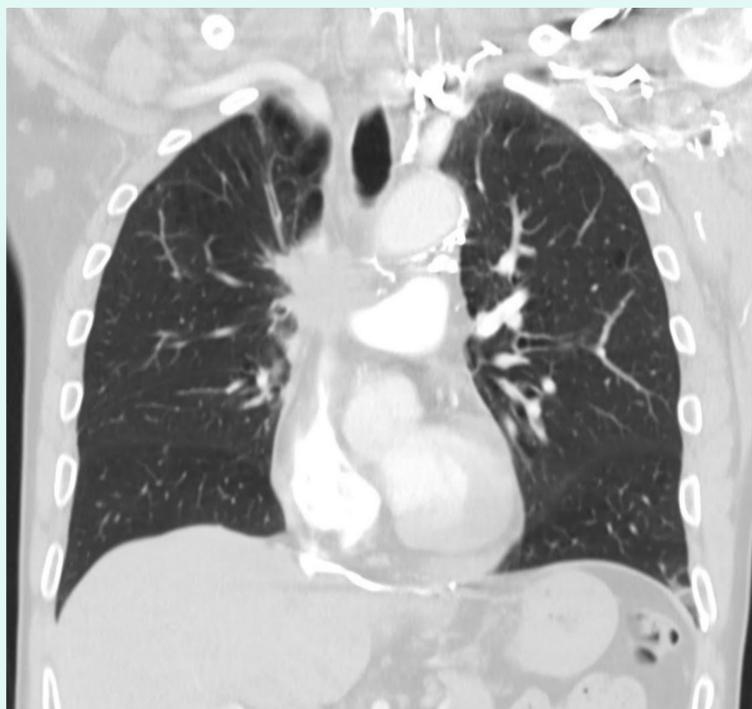


Figura 31 y 32: **TC torácico, corte coronal en ventana pulmón (31) y ventana mediastino (32)**. Lesión tumoral centrada en el hilio derecho y en la región suprahiliar del lóbulo superior derecho (LSD) con invasión del mediastino (paratraqueal derecho, subcarinal) y vascular (arterial y venosa).

Caso 8: Verdadero positivo para nódulo pulmonar y masa mediastínica.

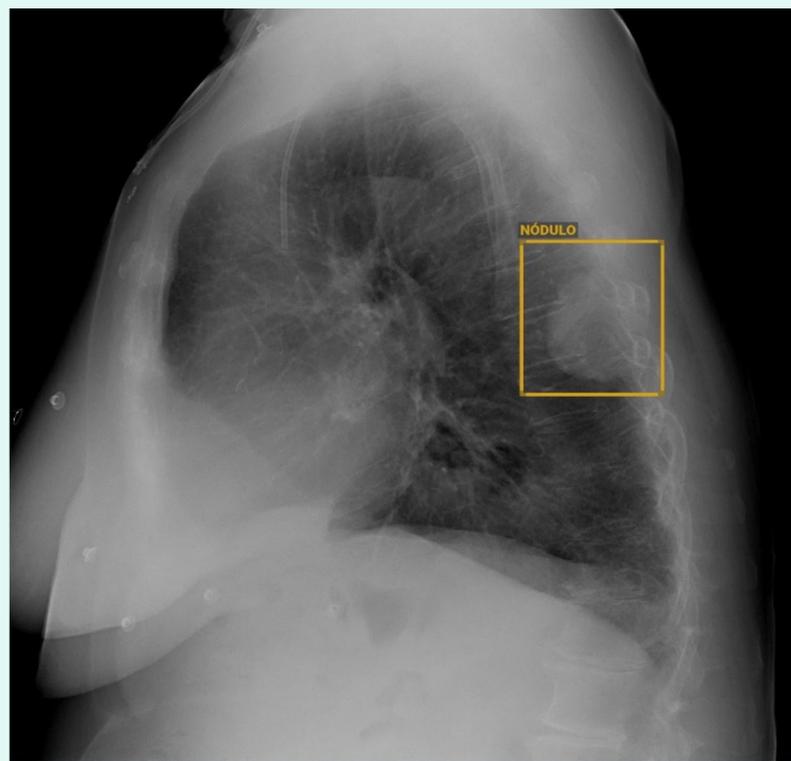
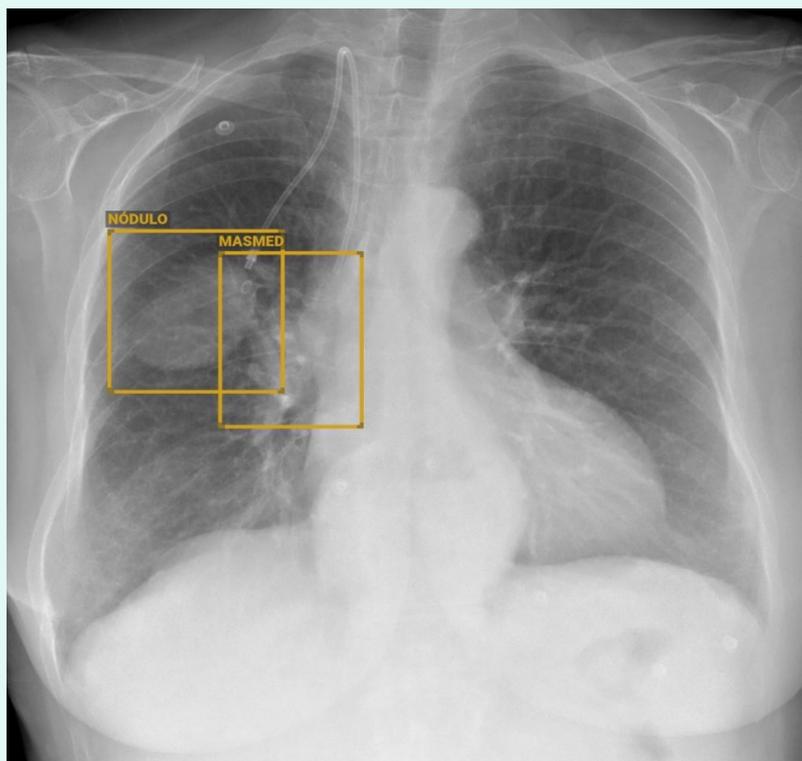


Figura 33 y 34: **Radiografía PA-L analizada por la IA:** marca “positividad” para nódulo pulmonar en LID y para masa mediastínica en hilio ipsilateral.

Informe del radiólogo: Masa pulmonar en LID, en el contexto del paciente, probable metástasis. Aumento del tamaño hilar derecho que sugiere adenopatía. Hernia de hiato. Catéter central con extremo en VCS.



Figura 35 y 36: **TC torácico, corte coronal y sagital en ventana mediastino.** Adenopatía hiliar derecha que presenta 36 mm de diámetro transversal. Lesión nodular pulmonar de 38 mm en LID. LOEs hepáticas.

Caso 9: Falso positivo para masa mediastínica.

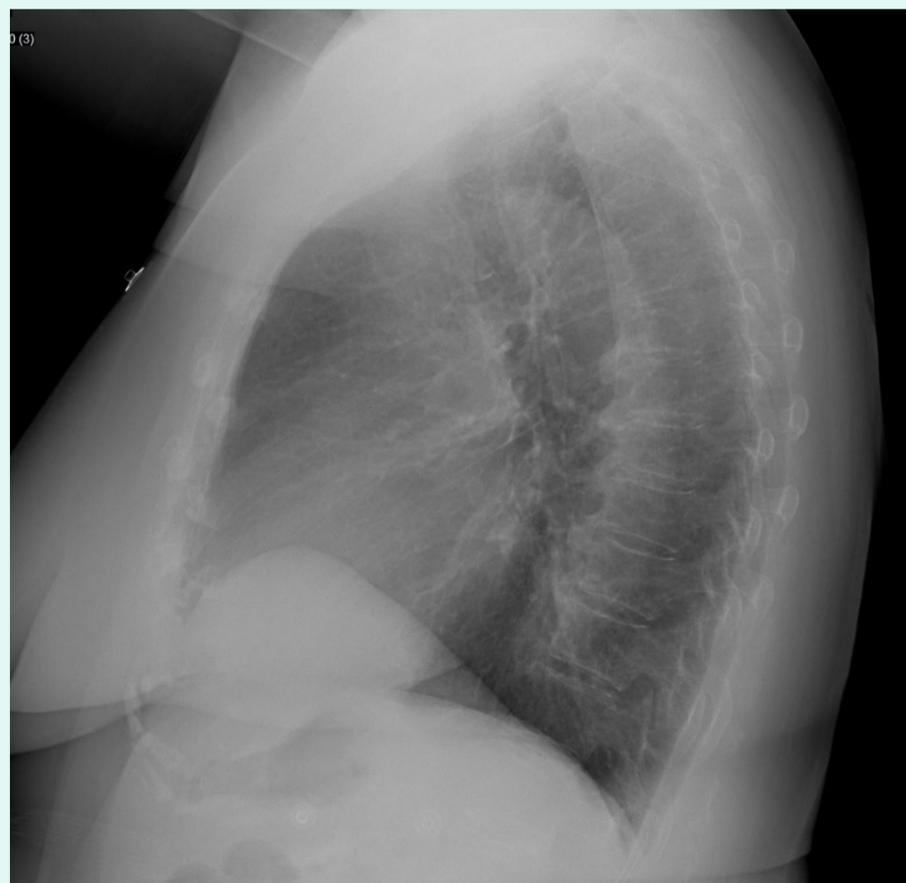
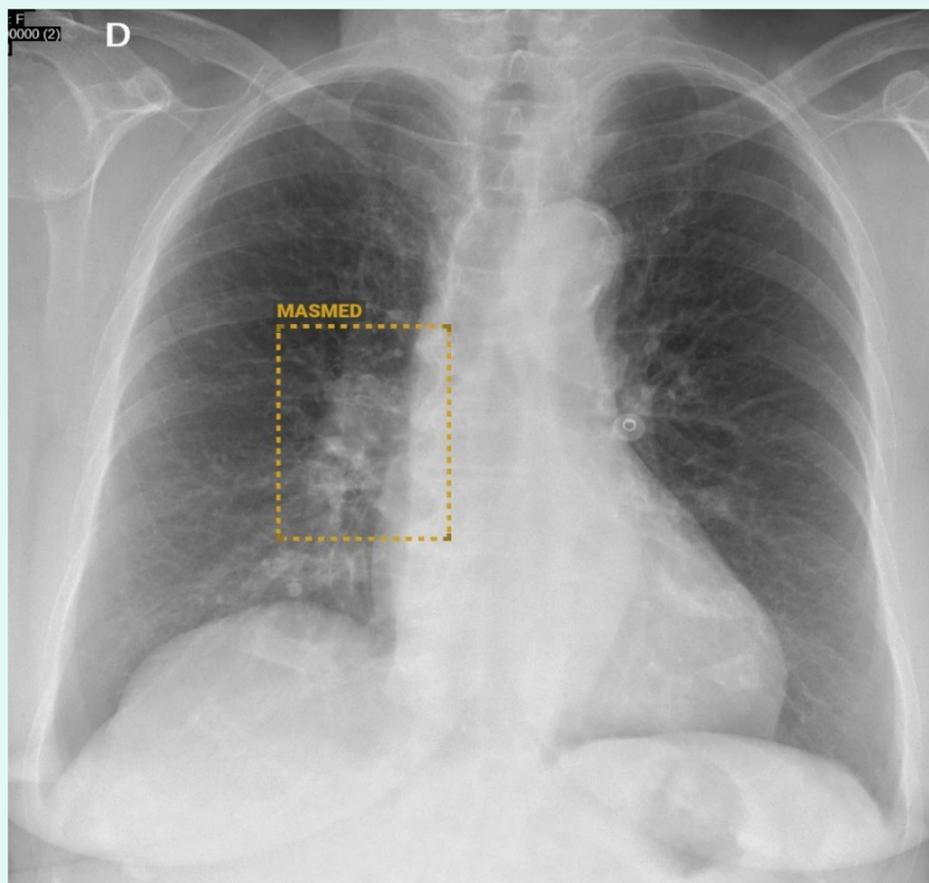


Figura 37 y 38: **Radiografía PA-L analizada por la IA:** marca “duda” para masa mediastínica.

Informe del radiólogo: Hilios prominente de origen vascular.

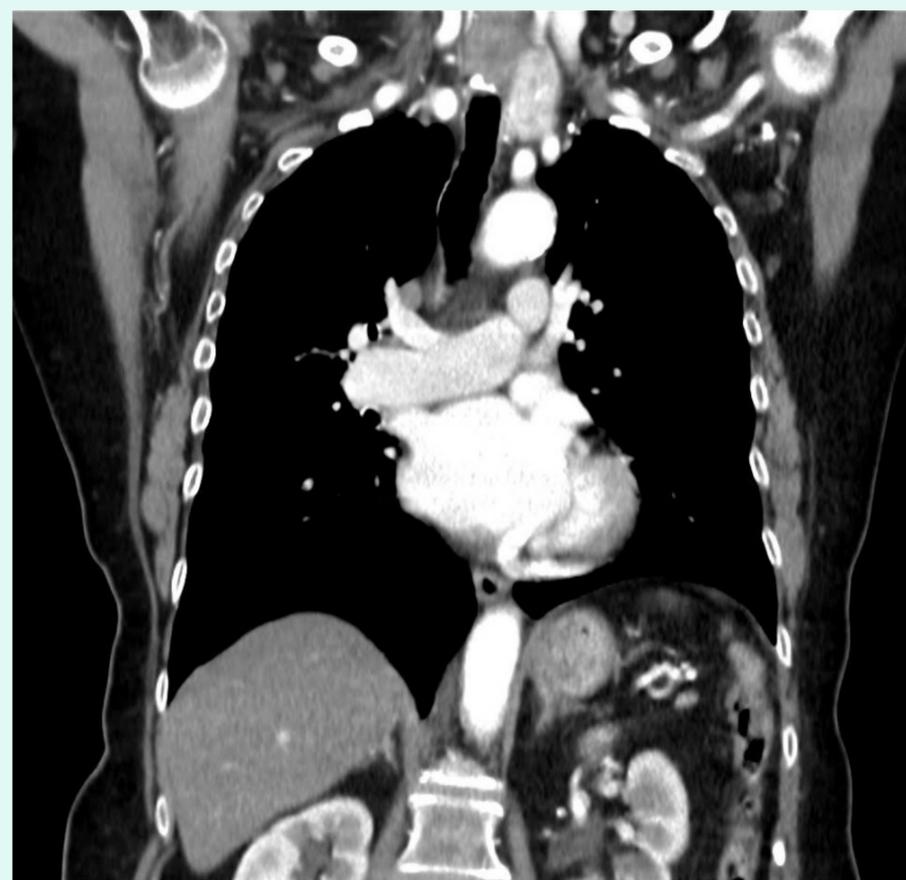
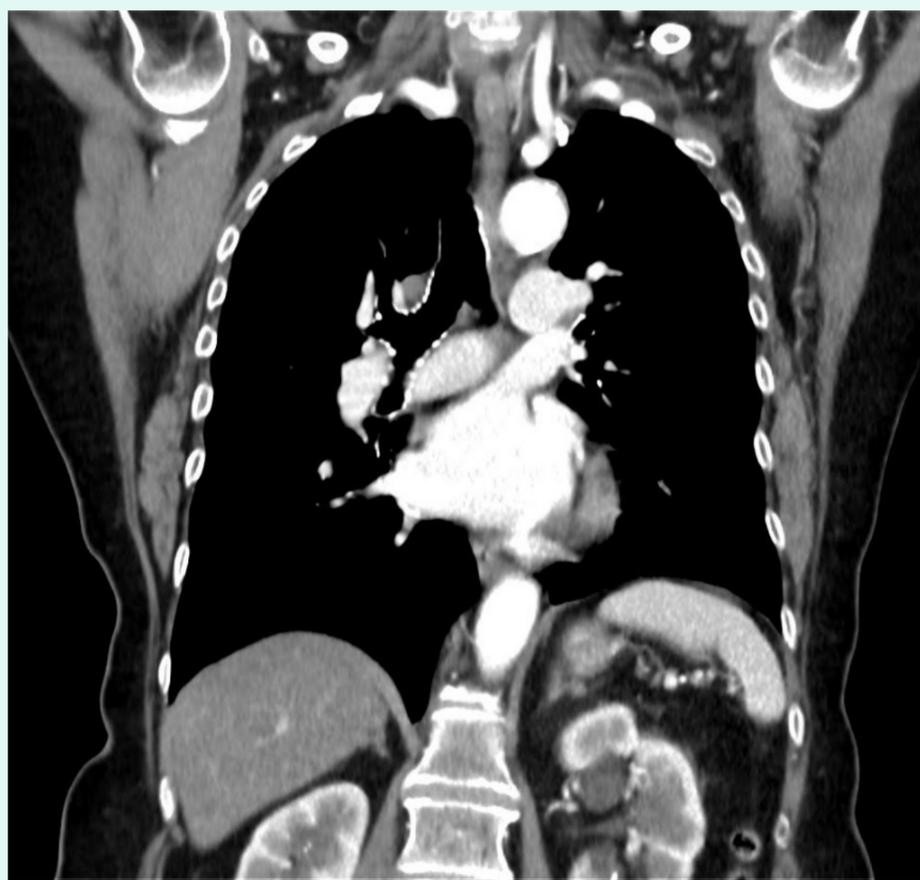


Figura 39 y 40: **TC torácico corte coronal en ventana mediastino:** Hilio prominente de origen vascular.

Caso 10: Falso positivo para masa mediastínica.

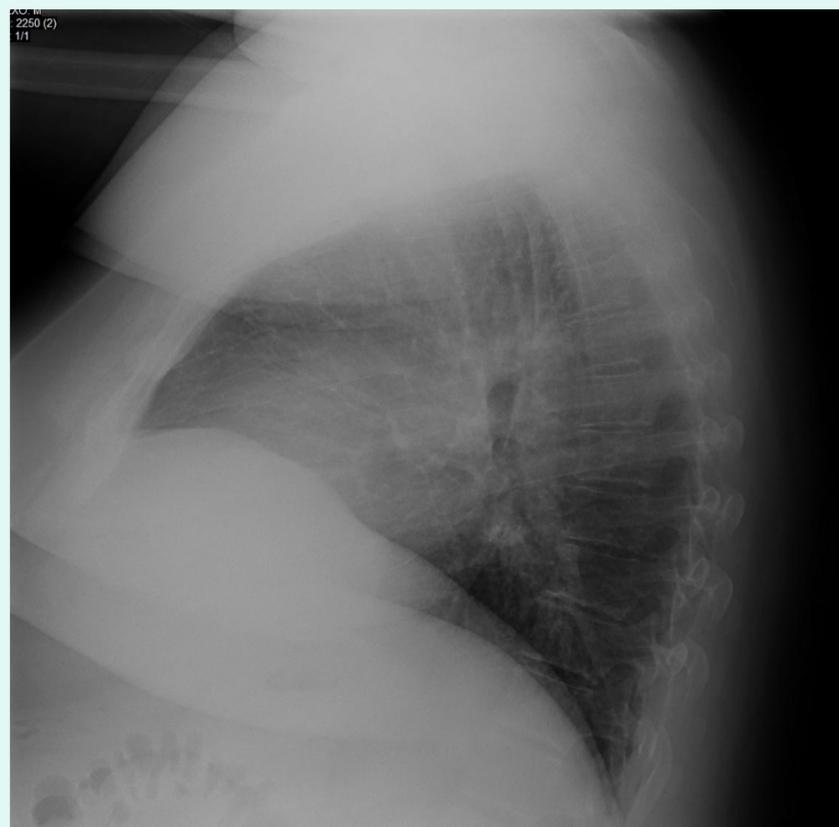
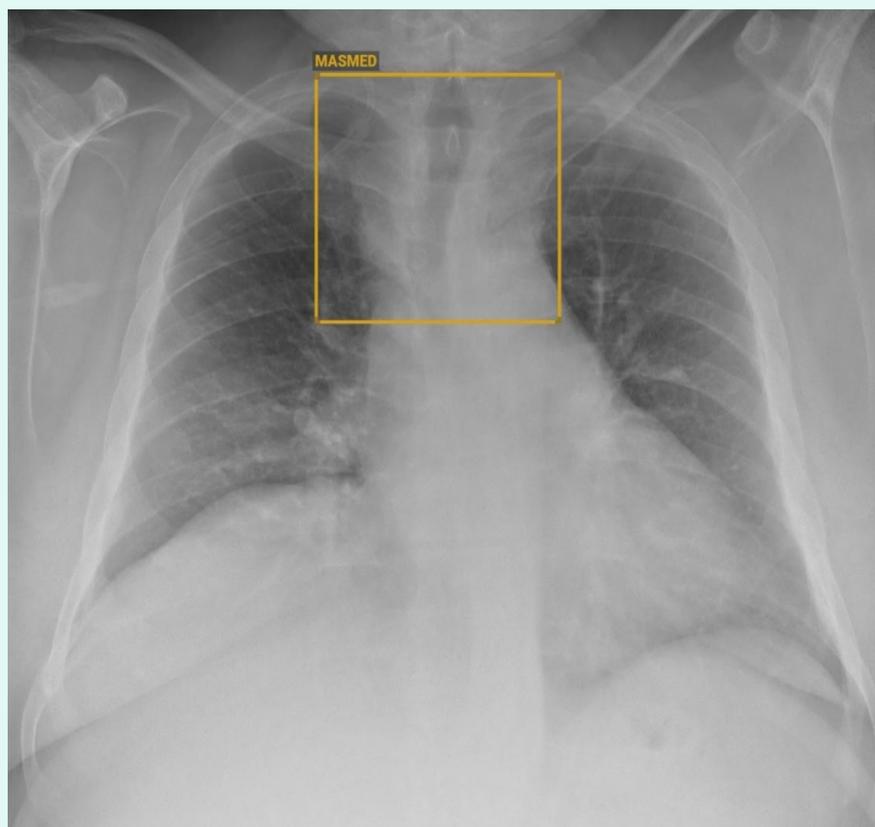


Figura 39 y 40: **Radiografía PA-L analizada por la IA:** marca “positivo” para masa mediastínica.

Informe del radiólogo: no hizo referencia al mediastino, al considerarlo normal.



Figura 41 y 42: **TC torácico, corte coronal en ventana mediastino (41) y en ventana pulmón (42).** El TC confirma la ausencia de masa mediastínica. Su discreto ensanchamiento se debe a lipomatosis mediastínica y a los troncos supraaórticos.

BIBLIOGRAFIA:

1. Cáceres J. La radiografía de tórax. *Imagen Diagnóstica*. 1 de enero de 2012;3(1):1-2.
2. Villoslada B, F J. La radiografía de tórax en los Servicios de Urgencia y Unidades de Corta Estancia: ¿Es necesario el informe del radiólogo? *An Med Interna*. septiembre de 2005;22(9):407-8.
3. Kim EY, Kim YJ, Choi WJ, Jeon JS, Kim MY, Oh DH, et al. Concordance rate of radiologists and a commercialized deep-learning solution for chest X-ray: Real-world experience with a multicenter health screening cohort. *PloS One*. 2022;17(2):e0264383.
4. Nam JG, Hwang EJ, Kim J, Park N, Lee EH, Kim HJ, et al. AI Improves Nodule Detection on Chest Radiographs in a Health Screening Population: A Randomized Controlled Trial. *Radiology*. abril de 2023;307(2):e221894.
5. Hwang EJ, Goo JM, Nam JG, Park CM, Hong KJ, Kim KH. Conventional Versus Artificial Intelligence-Assisted Interpretation of Chest Radiographs in Patients With Acute Respiratory Symptoms in Emergency Department: A Pragmatic Randomized Clinical Trial. *Korean J Radiol*. 1 de marzo de 2023;24(3):259-70.
6. Rudolph J, Schachtner B, Fink N, Koliogiannis V, Schwarze V, Goller S, et al. Clinically focused multi-cohort benchmarking as a tool for external validation of artificial intelligence algorithm performance in basic chest radiography analysis. *Sci Rep*. 27 de julio de 2022;12(1):12764.
7. Bernstein MH, Atalay MK, Dibble EH, Maxwell AWP, Karam AR, Agarwal S, et al. Can incorrect artificial intelligence (AI) results impact radiologists, and if so, what can we do about it? A multi-reader pilot study of lung cancer detection with chest radiography. *Eur Radiol*. noviembre de 2023;33(11):8263-9.
8. Yoo H, Kim EY, Kim H, Choi YR, Kim MY, Hwang SH, et al. Artificial Intelligence-Based Identification of Normal Chest Radiographs: A Simulation Study in a Multicenter Health Screening Cohort. *Korean J Radiol*. octubre de 2022;23(10):1009-18.