



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

Facultad de  
**MEDICINA**

**VALIDACIÓN DE UNA HERRAMIENTA PARA LA EVALUACIÓN DE LOS  
CONOCIMIENTOS Y LAS DESTREZAS PARA LA INTERPRETACIÓN DE LA  
RADIOGRAFÍA DE TÓRAX**

**VALIDATION OF A TOOL DESIGNED TO EVALUATE THE KNOWLEDGE AND  
DEXTERITY OF CHEST X-RAY INTERPRETATION**

TESIS PARA OPTAR POR EL TÍTULO PROFESIONAL DE MÉDICO CIRUJANO

**AUTORES**

BLAS MIGUEL CORNEJO ESPARZA

JUAN FELIPE VASQUEZ RIOJA

SEBASTIAN RAUL ORBEGOSO ALIAGA

**ASESORES**

KARLA BEATRIZ TAFUR BANCES

**CO-ASESORES**

RAY WILLY TICSE AGUIRRE

LIMA – PERÚ

2024



## **JURADO**

Presidente: Dra. Ana Cecilia Olascoaga Mesia

Vocal: Dra. Meylin Rosa Aphanh Lam De Vilchez

Secretario: Dr. Cristian Paul Leon Rabanal

**Fechas de sustentación:** 16/01/24

**Calificación:** Aprobado

**ASESORES DE TESIS**

**ASESORA**

Dra. Karla Beatriz Tafur Bances

Universidad Peruana Cayetano Heredia

ORCID: 0000-0001-5040-8432

**CO-ASESOR**

Dr. Ray Willy Ticse Aguirre

Universidad Peruana Cayetano Heredia

ORCID: 0000-0002-4568-5723

## **DEDICATORIA**

Dedicado a nuestras familias, amigos y mentores quienes nos acompañaron durante esta larga carrera.

## **AGRADECIMIENTOS**

Agradecemos a nuestros asesores la Dra. Karla Tafur y el Dr. Ray Ticse, a los médicos especialistas validadores y a los internos de medicina de la promoción 2023 quienes dedicaron su tiempo en ayudar al desarrollo de este proyecto.

## **FUENTES DE FINANCIAMIENTO**

El proyecto de investigación fue autofinanciado por los autores.

## **DECLARACIÓN DE CONFLICTO DE INTERÉS**

Los autores de este estudio declaran no tener conflictos de interés.

## RESULTADO DEL INFORME DE SIMILITUD

### VALIDACIÓN DE UNA HERRAMIENTA PARA LA EVALUACIÓN DE LOS CONOCIMIENTOS Y LAS DESTREZAS PARA LA INTERPRETACIÓN DE LA RADIOGRAFÍA DE TÓRAX

#### INFORME DE ORIGINALIDAD

<b>10%</b>	<b>10%</b>	<b>3%</b>	<b>2%</b>
INDICE DE SIMILITUD	FUENTES DE INTERNET	PUBLICACIONES	TRABAJOS DEL ESTUDIANTE

#### FUENTES PRIMARIAS

<b>1</b>	<b>repositorio.upch.edu.pe</b> Fuente de Internet	<b>2%</b>
<b>2</b>	<b>www.researchgate.net</b> Fuente de Internet	<b>1%</b>
<b>3</b>	<b>pesquisa.bvsalud.org</b> Fuente de Internet	<b>1%</b>
<b>4</b>	<b>Submitted to Universidad Peruana Cayetano Heredia</b> Trabajo del estudiante	<b>&lt;1%</b>
<b>5</b>	<b>dspace.utpl.edu.ec</b> Fuente de Internet	<b>&lt;1%</b>
<b>6</b>	<b>catalonica.bnc.cat</b> Fuente de Internet	<b>&lt;1%</b>
<b>7</b>	<b>renatiqa.sunedu.gob.pe</b> Fuente de Internet	<b>&lt;1%</b>
<b>8</b>	<b>issuu.com</b> Fuente de Internet	<b>&lt;1%</b>

## **TABLA DE CONTENIDOS**

<b>RESUMEN.....</b>	<b>.....</b>
<b>ABSTRACT.....</b>	<b>.....</b>
<b>1. INTRODUCCIÓN.....</b>	<b>1</b>
<b>2. OBJETIVOS.....</b>	<b>4</b>
<b>2.1 Objetivo principal.....</b>	<b>4</b>
<b>2.2 Objetivos específicos.....</b>	<b>4</b>
<b>3. MATERIALES Y MÉTODOS.....</b>	<b>5</b>
<b>3.1 Diseño de estudio.....</b>	<b>5</b>
<b>3.2 Población.....</b>	<b>5</b>
<b>3.2.1 Validación de la herramienta.....</b>	<b>5</b>
<b>3.2.2 Confiabilidad de la herramienta.....</b>	<b>6</b>
<b>3.3. Operacionalización de las variables (Anexo 1).....</b>	<b>7</b>
<b>3.4. Procedimientos y Técnicas.....</b>	<b>7</b>
<b>3.5. Consideraciones éticas.....</b>	<b>10</b>
<b>3.6. Análisis estadístico.....</b>	<b>10</b>
<b>4. RESULTADOS.....</b>	<b>11</b>
<b>5. DISCUSIÓN.....</b>	<b>15</b>
<b>6. CONCLUSIONES.....</b>	<b>24</b>
<b>7. REFERENCIAS BIBLIOGRÁFICAS.....</b>	<b>25</b>
<b>8. TABLAS, GRÁFICOS Y FIGURAS.....</b>	<b>31</b>
<b>9. ANEXOS.....</b>	<b>.....</b>

## **RESUMEN**

**Antecedentes:** Existen muchos estudios a nivel mundial donde evalúan los conocimientos y destrezas en la interpretación de radiografías de tórax en médicos en formación; sin embargo, no existe ninguna herramienta validada para ello ni estudios nacionales sobre este tema. Por esto resulta de suma importancia la creación y validación de una herramienta con esta finalidad. **Objetivos:** Diseñar y validar una herramienta para evaluar las destrezas para la interpretación de la radiografía de tórax en estudiantes de pregrado de la carrera de medicina. **Materiales y métodos:** Estudio observacional, descriptivo y de corte transversal. Se diseñó una herramienta basada en los currículos de UPCH y ASPEFAM. Se realizó la validación por 7 médicos expertos del Hospital Cayetano Heredia. Se aplicó dicho cuestionario a un grupo piloto de 107 alumnos, se calculó el índice de confiabilidad y se realizó el análisis de ítems. **Resultados:** Todos los ítems tuvieron CVI  $>0.58$  y herramienta completa obtuvo un CVI de 0.93. Se obtuvo 2 (12,5%), 6 (37.5%) y 8 (50%) ítems con dificultad “muy difícil”, “intermedia” y “muy fácil” respectivamente, con un promedio de 0.64. Se obtuvo 11 (68.8%) ítems con discriminación aceptable ( $\geq 0.2$ ) y con un promedio 0.36. El índice de confiabilidad fue de 0.40 con los ítems iniciales y de 0.51 cuando se retiraron los ítems con peor desempeño. **Conclusiones:** Nuestra herramienta es válida, cuenta con una confiabilidad moderada y muestra índices de discriminación y dificultad acorde a la bibliografía.

**Palabras clave:** radiografías de tórax, interpretación, conocimientos, destrezas, estudiantes de medicina, validación, confiabilidad

## **ABSTRACT**

**Background:** There are many studies worldwide evaluating medical trainees' chest X ray interpretation skills. However, there are no tools validated for this area or local studies. Designing a systematically validated tool is needed to close this knowledge gap. **Objectives:** Create and validate a tool to evaluate final year medical students' chest X ray interpretation skills. **Methodology:** Cross sectional, observational study. A chest X ray interpretation tool was designed using UPCH's and ASPEFAM's syllabus. Validation was performed by 7 medical doctors from Hospital Nacional Cayetano Heredia. The tool was then implemented in a pilot group of 107 final year medical students and a consistency ratio was calculated.

**Results:** All items in the tool got a CVI  $>0.58$  and the tool as a whole got a CVI of 0.93. 2 (12,5%), 6 (37.5%) y 8 (50%) items had a difficulty index of "very hard", "intermediate" and "very easy", respectively, with an average of 0.64. 11 (68.8%) items had an acceptable discrimination index ( $\geq 0.2$ ) with an average score of 0.36. Initially, internal consistency was 0.40 which increased to 0.51 when items with poor performance were removed from the tool. **Conclusions:** The tool designed in this study had appropriate validity, moderate consistency ratios, as well as acceptable item discrimination and difficulty scores.

**Key words:** chest X ray, validation, consistency, medicine students

## **1. INTRODUCCIÓN**

La radiografía de tórax es una herramienta básica para la evaluación de patologías cardiovasculares, mediastinales y, en especial, pulmonares (1). La placa de tórax es de uso cotidiano en servicios de emergencia, hospitalización, quirúrgicos y de cuidados críticos por su rapidez, bajo costo y amplia disponibilidad (2). Por estas razones la correcta interpretación de una radiografía de tórax es una habilidad necesaria, tanto entre los médicos graduados, así como también en los que todavía están en formación para reducir la tasa de error diagnóstico y con ello, una mala práctica profesional.

Existen diversos patrones radiológicos claves que pueden ser identificados en placas de tórax, los cuales fueron descritos por la sociedad de Fleischner en el 2008. Entre ellos están los consolidados donde se encuentran la neumonía lobar y el edema agudo de pulmón, el infiltrado intersticial, donde se encuentran la fibrosis pulmonar y el enfisema, la enfermedad pleural, caracterizada por neumotórax y efusión pleural, y finalmente las masas pulmonares donde se encuentran los tumores, nódulos, micronódulos y abscesos. (3)

Los estudiantes de último año de medicina de la Universidad Peruana Cayetano Heredia (UPCH) han recibido entrenamiento en lectura de radiografías de tórax en diferentes momentos durante la carrera, desde el segundo hasta el último año en cursos tales como anatomía, introducción a la clínica, clínica médica I y clínica de especialidades pediátricas (4, 5, 6, 7). Uno de los objetivos planteados en el plan curricular de la carrera profesional de medicina es que los alumnos egresados tengan la capacidad de saber cuándo solicitar e interpretar adecuadamente pruebas diagnósticas de imágenes, entre ellos radiografía de tórax (8).

Para poder evaluar las destrezas de los alumnos de medicina para interpretar radiografías de tórax, primero se debe validar la herramienta que busca esta finalidad. El proceso de validación se divide en dos componentes principales: la validez y la confiabilidad (9). La validez se define como el grado en que una prueba mide la información diseñada con esta finalidad. Se puede dividir en tres tipos: la validez de contenido, la cual mide qué tanto las preguntas del instrumento reflejan el conocimiento específico que se busca evaluar; la validez de constructo, la cual mide la congruencia entre la información recolectada y el objetivo del autor; y finalmente la validez predictiva, la cual mide la capacidad predictiva de los resultados de la herramienta con respecto a una variable externa, ya sea exámenes de universidad, residentado médico u otro tipo de evaluaciones.

El segundo componente de la validación es la confiabilidad, la cual se define como la precisión con la que la herramienta mide la variable en cuestión; es decir, una herramienta confiable es aquella que mide con el menor error posible el valor real de dicha variable. Adicionalmente, cuando se diseña una herramienta con preguntas de conocimiento, se puede analizar cada pregunta (o “ítem”) con el índice de dificultad y el índice de discriminación. El índice de dificultad es la proporción de respuestas correctas para cada ítem; este tiene valores de 0.0 a 1.0 y refleja que tan fácil es una pregunta, de manera que mientras más cerca de la unidad, más fácil es la pregunta (10, 11). Por otro lado, el índice de discriminación evalúa a capacidad que tiene un ítem para diferenciar entre alumnos que obtuvieron bajos puntajes versus altos puntajes en la herramienta; este toma valores 0.0 a 1.0 y mientras más cerca de la unidad, más discriminador es el ítem (10, 11).

Los estudios realizados por *Jeffrey et al.* (2002), *Eisen et al.* (2006) y *Erguera et al.* (2018) evaluaron la destreza en la interpretación de radiografías de tórax entre estudiantes de medicina y médicos residentes (12, 13, 14). No obstante, ninguno propone una herramienta que haya sido validada previamente para tal fin. Además, no se han realizado investigaciones sobre este tema a nivel nacional. Por estas razones, es necesario un estudio que valide dicha herramienta para nuestro contexto, ya que la radiografía de tórax es una prueba de bajo costo, esencial en el diagnóstico de patologías cardiorrespiratorias y que todo médico debería saber interpretar. Esto es especialmente cierto en Perú, país con alta prevalencia de enfermedades respiratorias tales como tuberculosis pulmonar, uno de los más afectados mundialmente por la pandemia del virus SARS-CoV-2 y con recursos limitados. Este estudio busca la construcción y validación de un instrumento para evaluación de los conocimientos y destrezas en la interpretación de radiografías de tórax de alumnos de pregrado de la carrera de medicina, ya que no existe una herramienta semejante validada en la literatura para la población peruana.

## **2. OBJETIVOS**

### **2.1 Objetivo principal**

Diseñar y validar una herramienta para evaluar las destrezas para la interpretación de la radiografía de tórax en estudiantes de pregrado de la carrera de medicina.

### **2.2 Objetivos específicos**

1. Determinar la validez de contenido de la herramienta para la evaluación de lectura de radiografía de tórax.
2. Determinar la confiabilidad de la herramienta para la evaluación de lectura de radiografía de tórax.
3. Evaluar la correlación entre el nivel de certeza en la respuesta de la herramienta con la nota obtenida en la misma.
4. Determinar el índice de dificultad y el índice de discriminación de las preguntas planteadas en la herramienta.

### **3. MATERIALES Y MÉTODOS**

#### **3.1 Diseño de estudio**

Estudio observacional, descriptivo y de corte transversal.

#### **3.2 Población**

##### **3.2.1 Validación de la herramienta**

Médicos docentes neumólogos, radiólogos e internistas del Hospital Cayetano Heredia (HNCH).

##### **Criterios de inclusión**

Se seleccionó a médicos expertos en la lectura de radiografías. Para ser incluidos en nuestro estudio, debían cumplir con los siguientes criterios: (a) 2 años de experiencia como docente, (b) contar con correo institucional UPCH activo y (c) tener 5 años de experiencia como médico especialista.

##### **Criterio de exclusión**

Se excluyeron a los médicos expertos que (a) contestaron el correo electrónico en un plazo mayor a 2 semanas y (b) entregaron un formato de validación incompleto o mal llenado.

##### **Muestra**

Se seleccionó a tres médicos neumólogos del servicio de CENEX, dos médicos radiólogos del servicio de Radiología y dos médicos internistas del servicio de Medicina “A” y “B” que cumplieran los criterios de inclusión y exclusión del estudio.

##### **Muestreo**

La unidad muestral fueron médicos neumólogos, médicos radiólogos o médicos internistas. El tipo de muestreo fue no probabilístico por conveniencia. El tamaño de muestra fue 7 (3 médicos neumólogos, 2 médicos radiólogos y 2 médicos internistas).

### **3.2.2 Confiabilidad de la herramienta**

Estudiantes del séptimo año académico de la carrera de medicina de la UPCH

#### **Criterios de inclusión**

Se seleccionó a estudiantes del séptimo año de la carrera de medicina de la Universidad Peruana Cayetano Heredia (UPCH). Se incluyeron aquellos que cumplieran los siguientes criterios: (a) Alumnos de séptimo año de la carrera de medicina de la UPCH, (b) contar con una matrícula activa en la universidad y (c) contar con correo institucional UPCH activo.

#### **Criterio de exclusión**

Se excluyeron a (a) alumnos que tengan entrenamiento extracurricular de lectura de radiografía de tórax, (b) alumnos que tengan carrera de tecnología médica y (c) Respuestas que hayan llegado con menos de 1 minuto de diferencia entre una y otra para evitar plagio.

#### **Muestra**

Se seleccionó a todos los alumnos de último año de la carrera de medicina de la UPCH que contesten el cuestionario y cumplan con los criterios de inclusión y exclusión.

#### **Muestreo**

La unidad muestral fue el alumno de pregrado de la carrera de medicina de la UPCH. El tipo de muestreo fue no probabilístico por conveniencia, se le envió el

cuestionario a toda la promoción 2023 de la UPCH y se tomó en cuenta para la muestra piloto todas las respuestas obtenidas. El tamaño de muestra calculado fue de 104 alumnos de pregrado de la carrera de medicina. Se calculó el tamaño muestral mediante la fórmula propuesta por *Bonett et al* utilizada para estudios de evaluación de la confiabilidad de una herramienta, teniendo en cuenta el coeficiente de confiabilidad mínimo de 0.7. (15) Esto concuerda con lo propuesto por *Imasuen K*, donde se demuestra que el tamaño muestral mínimo para obtener un coeficiente de confiabilidad aceptable es de 100 participantes. (16)

### **3.3. Operacionalización de las variables (Anexo 1)**

**Validación de Contenido:** Se mide mediante el llenado de un formato de validación por 7 médicos expertos los cuales evaluarán la calidad técnica de la radiografía, la coherencia con la respuesta, la inducción a la respuesta y la claridad de la redacción de la viñeta clínica para cada uno de los ítems de la herramienta.

**Confiabilidad:** Precisión con la cual cada ítem de la herramienta recolecta información acerca de las habilidades y destrezas en la interpretación de las radiografías de tórax, evaluada aplicando la herramienta a un grupo piloto de estudiantes y utilizando las notas obtenidas para calcular un índice de confiabilidad mediante el método de Kuder-Richardson 20. Un índice de confiabilidad de 0.01 a 0.20 corresponde a una confiabilidad muy baja, de 0.21 a 0.4 baja, de 0.41 a 0.6 moderada, de 0.61 a 0.8 alta y de 0.81 a 1 muy alta.

### **3.4. Procedimientos y Técnicas**

La herramienta fue diseñada en modo de un cuestionario de 16 preguntas. Cada pregunta cuenta con la misma estructura: una viñeta clínica pequeña con la

descripción del caso y una radiografía de tórax patológica (**Anexo 2**). Los alumnos respondieron a las preguntas con el diagnóstico que creían que era el más adecuado, luego, el alumno contestó cuán seguro se encuentra en su respuesta mediante una escala Likert que va del 1 al 4, siendo 1 muy inseguro y 4 muy seguro.

La elección de las patologías a evaluar fue basada en las competencias fijadas en los sílabos de cursos clínicos de la UPCH (4, 5, 6, 7) y en los tópicos evaluados en el Examen Nacional de Medicina (ENAM) por ASPEFAM (17, 18). La lista final de patologías fue: neumonía lobar, edema agudo de pulmón, fibrosis pulmonar, neumotórax, derrame pleural, atelectasia, efusión pericárdica, absceso pulmonar, tuberculosis, tuberculosis miliar, masa pulmonar, nódulo pulmonar solitario, hidroneumotórax y bronquiectasias. Las radiografías fueron extraídas del archivo del HCH por los autores y evaluadas por un médico neumólogo del mismo hospital, para determinar que sean representaciones clásicas de la enfermedad que se busca evaluar y no exista alguna otra patología mayor en la radiografía.

Para la validación de la herramienta, primero se realizó la validez de contenido y luego la confiabilidad de la misma. Se envió un formato de validación (**Anexo 3**) adaptado de estudios previos (9, 19) y la herramienta a 7 médicos: 3 médicos neumólogos, 2 médicos internistas y 2 médicos radiólogos. En el formato de validación, cada evaluador calificó con un puntaje del 1 al 4 las siguientes categorías: calidad técnica de la radiografía, coherencia entre radiografía y respuesta, inducción a la respuesta, redacción de la pregunta y finalmente si estaban o no de acuerdo con el listado de respuestas correctas posibles. Además, los evaluadores indicaron sus observaciones para cada ítem de la herramienta. Se

obtuvieron 7 formatos de validación (1 por cada experto) y se construyó una base de datos en Microsoft Excel versión 17.0. Se calculó el ratio de validez de contenido (CVR) para cada uno de los 4 criterios por ítem y luego se calculó el índice de validez de contenido (CVI) por ítem, promediando los CVR de los cuatro criterios. Luego, se calculó el índice de validez de contenido de todo el cuestionario, promediando los CVI de los 15 ítems.

Una vez resueltas las observaciones recomendadas por los expertos en el formato de evaluación, se le enseñaron los cambios realizados a los expertos evaluadores y se re-calcularon los CVR y CVI de cada ítem, se procedió a la aplicación de la misma en un grupo piloto de alumnos de último año de medicina de la UPCH para obtener la confiabilidad. Se envió el consentimiento informado, el compromiso de no plagio y la herramienta a toda la promoción de internos 2023 de la UPCH y se tomaron en cuenta todos los alumnos en responder que cumplían con los criterios de inclusión y exclusión establecidos. Se aplicó la herramienta con un tiempo límite de 16 minutos, correspondiente a 1 minuto por pregunta. Se eliminaron 8 respuestas que no cumplían con los criterios de inclusión y exclusión. Posterior a ello, se construyó una base de datos en Microsoft Excel versión 17.0 con las respuestas obtenidas de Google Forms y con esta base se realizaron los cálculos de Índice de confiabilidad, Índice de Dificultad e Índice de Discriminación.

Finalmente, se evaluó la correlación entre el nivel de seguridad que tenían los alumnos en su respuesta y si es que estos acertaban a dicha respuesta.

### **3.5. Consideraciones éticas**

El presente estudio está registrado en el Sistema Descentralizado de Información y Seguimiento a la Investigación (SIDISI) de la UPCH (N° 211514) y cuenta con la aprobación del Comité Institucional de Ética de la UPCH. Previo al desarrollo del cuestionario, los alumnos firmaron un consentimiento informado para formar parte del estudio (**Anexo 4**).

### **3.6. Análisis estadístico**

Se calculó el ratio de Validez de Contenido (CVR) para los 4 criterios de cada ítem utilizando el método cuantitativo de Lawshe (20). Luego, para evaluar la confiabilidad de la herramienta, se utilizó el método de Kuder-Richarson 20 (KR20) para respuestas dicotómicas con una sola aplicación del cuestionario en el grupo piloto. (21,22). Para realizar en análisis psicométrico de los ítems, se calculó el índice de dificultad y de discriminación de cada ítem de la herramienta (10).

Finalmente, se calculó el coeficiente de correlación de Pearson entre el número de veces que el participante asignó una puntuación 3 o 4 en el nivel de seguridad en su prueba (es decir “seguro” o “totalmente seguro”) y el puntaje total que obtuvo (23).

#### **4. RESULTADOS**

Los 7 expertos evaluaron los 15 ítems bajo los siguientes 4 criterios: calidad técnica de radiografía, coherencia entre la radiografía y la respuesta, inducción a la respuesta y redacción de la viñeta clínica (**ANEXO 3**).

Para el criterio de calidad técnica de la radiografía los expertos puntuaron todos los ítems como “aceptable” o “buena calidad” excepto los ítems 2, 3, 9 y 14, (neumonía, edema agudo de pulmón, absceso pulmonar y masa pulmonar respectivamente) donde uno de los 7 expertos le dio una puntuación de “mala calidad” y el ítem 7 (atelectasia) donde dos expertos le dieron una puntuación de “mala calidad.” (**TABLA 1**)

Para el criterio de coherencia entre radiografía escogida y respuesta correcta, todos los expertos puntuaron la radiografía de tórax escogida como “coherente” o “muy coherente con la respuesta” excepto en el ítem 2 (neumonía) donde uno de los expertos le dio una puntuación de “poco coherente”, y el ítem 11 (nódulo pulmonar solitario) donde uno de los expertos le dio una puntuación de “muy poco coherente”. (**TABLA 2**)

Para el criterio de viñeta clínica induce a la respuesta, hubo un consenso entre los expertos para los ítems 3, 4, 6, 7, 8 y 11 (edema agudo de pulmón, fibrosis pulmonar, derrame pleural, atelectasia masiva, derrame pericárdico y nódulo pulmonar solitario respectivamente) estos fueron puntuados como “respuesta difícilmente inducida por la viñeta” y “respuesta no se induce por la viñeta”. Sin embargo, en los ítems 2, 5, 9, 10 y 12 (neumonía, neumotórax, absceso pulmonar, patrón miliar y bronquiectasias respectivamente) uno de los expertos evaluadores les dio una puntuación de “respuesta podría ser inducida por la viñeta”, para los

ítems 13, 14 y 15 (tuberculosis, masa pulmonar e hidroneumotórax respectivamente) dos expertos les dieron una puntuación de “respuesta inducida fácilmente por la viñeta” o “respuesta podría ser inducida por la viñeta”. Finalmente, para el ítem 1 (radiografía normal) 3 expertos le dieron una puntuación de “respuesta podría ser inducida por la viñeta”. **(TABLA 3)**

Para el cuarto criterio de redacción de la viñeta clínica, todos los ítems recibieron una puntuación de “redacción buena” o “redacción aceptable” excepto los ítems 7 (atelectasia) y 11 (nódulo pulmonar solitario) los cuales recibieron una puntuación de “redacción mala” por parte de uno de los expertos validadores. **(TABLA 4)**

También se les preguntó a los expertos si estaban de acuerdo con el listado de respuestas correctas posibles. En primera instancia los 7 expertos estuvieron de acuerdo con las respuestas propuestas, siempre y cuando se modifiquen una o dos respuestas en todos los ítems menos en el ítem 7 (atelectasia).

Basado en los resultados mostrados en la **TABLA 5**, se realizaron las modificaciones planteadas por los expertos y se volvió a validar la herramienta con los 7 expertos validadores. Las viñetas clínicas modificadas obtuvieron CVR perfectos en la revalidación y el CVI total del instrumento aumentó de 0.88 a 0.93. **(TABLA 6)**

Una vez aplicada la herramienta al grupo piloto, se obtuvieron los siguientes resultados de 107 respuestas, resumidos en la **TABLA 6**. El 62.6% de alumnos que respondió el cuestionario fue de sexo femenino. La media de edad fue de 24.5 años, la mediana y moda de edad fue de 24 años. 99 alumnos participantes no han recibido formación adicional en lectura de radiografías de tórax (92.5%), los 8 (7.5%) restantes refieren usar videos académicos en Internet. 70 (65.4%) de los

alumnos participantes reportaron las prácticas clínicas como el método de aprendizaje más útil para aprender a interpretar radiografías de tórax, en segundo lugar, están las clases teóricas (12.1%) y luego los ECOE, TBLs y SGPs y otros (7.5%, 6.5%, 5.6% y 2.8% respectivamente). **(TABLA 7)**

82 (76.6%) alumnos aprobaron la herramienta de interpretación de radiografías de tórax **(Gráfico 1)**. La media, moda y mediana de puntaje obtenido fue 10.18, 11 y 10 respectivamente de un total de 16 puntos. Se calculó el coeficiente de confiabilidad para los 16 ítems el cual fue 0.40.

Con respecto a los ítems del cuestionario, los 3 que tuvieron una mayor frecuencia de acierto fueron el 1.2 (identificación de sexo), 6 (efusión pleural), 5 (neumotórax) y 1 (radiografía normal) con 105, 100, 96 y 91 alumnos respectivamente. Los ítems con mayor grado de error fueron el 12 (bronquiectasias), 11 (nódulo pulmonar solitario) y 7 (atelectasia) con una frecuencia de acierto de 12, 21 y 42 alumnos de un total de 107 respectivamente.

### **(Gráfico 2)**

Luego, se calculó el índice de dificultad para cada ítem de la herramienta **(TABLA 8)**. Se obtuvo 8 ítems (ítems 1, 1.2, 2, 3, 5, 6, 8, 13) con índice de dificultad  $>0.7$  (o “muy fácil”), 6 ítems (ítems 4, 7, 9, 10, 14 y 15) con índice de dificultad entre 0.3-0.7 (o “intermedio”), 2 ítems (ítems 11 y 12) con índice de dificultad  $< 0.3$  (o “muy difícil”) (11) **(Gráfico 3)**. Los ítems 1.2 (identificar sexo), 6 (efusión pleural) y 5 (neumotórax) tuvieron los mayores índices de dificultad, siendo estos 0.98, 0.93, 0.90 respectivamente. Los ítems con menor índice de dificultad fueron el ítem 12 (bronquiectasias) y 11 (nódulo pulmonar

solitario) obteniendo 0.11 y 0.20 respectivamente. El promedio del Índice de Dificultad de la herramienta fue 0.64.

Se calculó también el índice de discriminación para cada ítem de la herramienta (**TABLA 8**). Se obtuvo 5 ítems (ítems 3, 9, 13, 14, 15) con índice de discriminación  $>0.39$  (o “excelente”), 5 ítems (ítems 2, 4, 7, 8, 10) con índice de discriminación entre 0.3 - 0.39 (o “buena”), 1 ítem (ítem 11) con índice de discriminación entre 0.2 - 0.29 (“regular”), 5 ítems (ítems 1, 1.2, 5, 6, 12) con índice de discriminación entre 0.0-0.20 (“pobre”) y 0 (ningún) ítem con índice de discriminación  $<0.20$  (“pésimo”) (10) (**Gráfico 4**). Los ítems con mayor índice de discriminación fueron el 3 (edema agudo de pulmón), el 9 (absceso pulmonar) y el 14 (masa pulmonar), con índice de discriminación de 0.65, 0.62 y 0.62 respectivamente. Los ítems 12 (bronquiectasia) y 1.2 (identificación de sexo) tuvieron los índices de discriminación mas bajos: 0.04 y 0.12 respectivamente. El promedio de Índice de Discriminación de la herramienta fue 0.36.

Finalmente, se calculó el coeficiente de correlación entre el número de veces que el participante asignó una puntuación 3 o 4 en el nivel de seguridad en su test (es decir “seguro” o “totalmente seguro”) y el puntaje total que obtuvo cuyo resultado fue de 0.06.

## **5. DISCUSIÓN**

En este estudio se diseñó y validó una herramienta mediante un método sistemático y estandarizado, la cual obtuvo índices de validación y confiabilidad aceptables. Existen otros estudios tales como los conducidos por *Cheung et al (2018)*, *Eisen et al (2006)* y *Jeffrey et al (2002)* donde se diseña una herramienta para evaluar la interpretación de radiografías de tórax por alumnos de medicina, sin embargo, estos autores no utilizaron un método cuantitativo para evaluar la validez de contenido y no evaluaron la confiabilidad. (24, 14, 12) El método de validación de los ítems del instrumento utilizado en estos estudios únicamente fue la presentación de las radiografías con sus respectivas viñetas clínicas a 3 médicos, validando dicho ítem si los 3 coincidían en el diagnóstico, en ninguno se evaluó el resto de los criterios utilizados en nuestro estudio.

La herramienta se diseñó teniendo en mente la idea de simular lo más cercanamente posible la práctica clínica diaria. Para conseguir esto se decidió: uno, agregar una breve viñeta clínica y dos, que la modalidad de respuesta sea de respuesta abierta. Se agregó la viñeta ya que, a pesar de que las radiografías de tórax pueden ser interpretadas sin necesidad de datos clínicos, en la práctica diaria siempre se conocen y utilizan los datos clínicos y epidemiológicos. Se optó por que la viñeta clínica sea breve para ayudar a discriminar entre los diagnósticos diferenciales posibles sin inducir a una respuesta. Esta decisión se fundamenta en dos estudios realizados: una revisión sistemática por *Loy et al (2004)* sobre la interpretación de pruebas diagnósticas con o sin datos clínicos donde concluye que la información clínica mejora la exactitud de dichos exámenes y un estudio por *Schreiber (1963)*, quien concluye que la información clínica aumenta el

desempeño en interpretación de radiografías (25, 26). Se utilizó la modalidad de respuesta abierta en lugar de opción múltiple debido a que, de haber utilizado la segunda alternativa, habría existido la posibilidad de que los participantes adivinen la respuesta o, en su defecto, se les habría facilitado llegar a esta al leer las alternativas posibles. Este enfoque resulta más cercano a la realidad ya que, en la práctica diaria, los médicos deberán afrontar pacientes reales y diagnosticar sin opción a elegir entre un número limitado de alternativas.

La decisión de emplear la herramienta de forma virtual se basó en los hallazgos del estudio realizado por *Miranda et al.* (2019) (27). En este estudio, se comparó el rendimiento de un grupo que realizó la prueba de forma presencial con otro que la realizó a través del correo electrónico. Los resultados indicaron que el método de administración de la prueba (presencial o electrónico) no modificó radicalmente el desempeño en la encuesta realizada electrónicamente. Además, en un estudio realizado por *Chan et al* de la Universidad de Washington donde se compararon notas de 2000 alumnos de 18 diferentes cursos se concluye que existe una relación fuerte entre la nota de exámenes presenciales y virtuales, a pesar de que estos últimos no cuenten con vigilancia. (28) Concluyen también que el plagio no aumenta las notas obtenidas en exámenes virtuales de una forma efectiva. También se consideró que, al no ser una prueba calificada, los alumnos no tienen un incentivo real para cometer plagio ya que no están bajo la presión de tener que obtener un buen desempeño en la prueba. Otra razón por la cual consideramos que el plagio en la aplicación de la herramienta sería mínimo es que las preguntas planteadas no son de naturaleza memorística, sino de razonamiento, ya que no es posible buscar la respuesta a la pregunta ya que se tendría que utilizar a una

persona más experimentada en lectura de radiografías que el alumno tomando la prueba. Sin embargo, aplicamos 3 medidas para evitar el plagio: la primera fue un compromiso de no plagio previo a la resolución de la herramienta, donde el alumno debía firmar un documento asegurando que no respondería la prueba utilizando recursos externos. La segunda medida fue incluir un tiempo límite para la prueba de 16 minutos, dando 1 minuto por radiografía, limitando el tiempo en el que el alumno podría utilizar recursos externos para completar la prueba. Finalmente, consideramos que el método principal de plagio sería la resolución de la herramienta en pareja, por lo que luego de haber obtenido las respuestas, filtramos y eliminamos aquellas que se registraron con menos de 1 minuto de diferencia entre una y otra.. Por lo tanto, se concluyó que la implementación de la herramienta en un formato virtual sería una opción viable y efectiva.

Respecto a la validación de contenido, el grupo de expertos estuvo conformado por médicos neumólogos, internistas y radiólogos. Fue organizado de esta manera para que el instrumento tenga la mayor validez posible: los médicos radiólogos aportarían mayor conocimiento sobre la calidad técnica e interpretación de las radiografías de tórax, los médicos internistas aportarían la correlación clínico-radiológica para mejorar las viñetas clínicas y finalmente los médicos neumólogos tendrían una visión más amplia, sintetizando ambos aportes previamente mencionados.

Al determinar la validez de la herramienta, se tomaron en consideración las observaciones realizadas por los expertos. La literatura recomienda que para conservar los ítems estos deben tener el CVR  $\geq 0.58$ . y CVI  $> 0.58$  (29). A pesar de lo argumentado anteriormente, para evitar que la viñeta clínica induzca la

respuesta como se demuestra en *Babatunde et al (2020)* (30), se incorporó el criterio de “inducción de respuesta por la viñeta”. En una primera instancia de revisión por los validadores, los ítems de radiografía normal, tuberculosis pulmonar, masa pulmonar e hidroneumotórax no cumplían con este criterio. Se modificaron en base a las observaciones planteadas y luego de la re-evaluación, dichos ítems fueron validados con un CVR de 1 por lo que permanecieron en el instrumento final.

Por otro lado, se decidió no retirar el caso de atelectasia por dos razones: uno, a pesar de contar con un CVR menor a lo recomendado para el criterio de calidad técnica, el CVI del ítem fue superior al rango planteado. Dos, los médicos radiólogos, quienes tienen mayor experiencia en evaluar la calidad técnica de la radiografía, calificaron la placa utilizada en este ítem como “aceptable” considerando que es una patología de atención médica inmediata por lo que se requiere una radiografía portátil donde la calidad técnica no siempre es óptima.

Respecto a la confiabilidad, para evaluaciones formativas que no pretenden clasificar a los alumnos por orden de mérito, se propone como índice de confiabilidad aceptable un valor de  $\geq 0.7$  (11). En nuestro estudio, con los 15 ítems iniciales, se obtuvo un índice de confiabilidad de 0.40 o “moderado” (9) el cual subió a 0.51, persistiendo dentro del mismo rango de una confiabilidad moderada luego de retirar aquellos que tuvieron peor desempeño (ítems de radiografía normal, identificación de sexo, efusión pleural y neumotorax, bronquiectasias). Este incremento es esperable ya que los ítems con poco poder discriminativo disminuyen la precisión de la herramienta (31).

Algo a resaltar es la discrepancia entre la alta validez y la baja confiabilidad del instrumento creado. El método empleado para el cálculo del índice de confiabilidad asume que el instrumento es unidimensional, es decir que la herramienta solo mide un concepto (en este caso la habilidad de interpretación de radiografías de tórax), y el modelo de tau equivalente, es decir que cada ítem de la herramienta tenga la misma relación con el constructo que se está queriendo evaluar. (32) Sin embargo, es importante notar que el valor del coeficiente de confiabilidad no solo es afectado por estos dos conceptos sino también puede verse afectado por la cantidad de ítems en la herramienta. (32) Al esta herramienta haber obtenido un índice de validez de contenido alto de 0.93 y solo tener 16 ítems, es probable que el bajo coeficiente de confiabilidad obtenido se deba a la poca cantidad de ítems y no a que la herramienta no cumpla con el modelo de tau equivalente o el de unidimensionalidad. Agregando nuevos ítems que evalúen las mismas patologías escogidas e incluyéndolos en la herramienta (es decir, tener más de un ítem que evalúe cada una de las patologías; 3 de neumonía, 3 de edema agudo de pulmón, etc.) probablemente llevaría a un aumento del coeficiente de confiabilidad.

En nuestro estudio piloto, los participantes obtuvieron una puntuación mediana de 12.5 sobre 20 puntos. Mientras que en el estudio realizado por *Jeffrey et al* (2002) donde se evaluó la interpretación de radiografías de tórax a 52 estudiantes de medicina de último año en el Reino Unido, la puntuación fue de 12.5 sobre 20 puntos (12). En dicho estudio, se presentaron 10 radiografías seleccionadas por los investigadores, cada una con un único diagnóstico posible, y revisadas por dos expertos radiólogos. Al igual que en nuestro estudio, los participantes fueron

voluntarios y las respuestas fueron abiertas. Sin embargo, a diferencia del nuestro, a los estudiantes no se les proporcionó información clínica previa y no hubo límite de tiempo para responder. En *Eisen et al* (2006), utilizaron una metodología similar a la anterior pero se evaluó a estudiantes de medicina y médicos residentes como parte de su rotación en medicina interna, se obtuvo una mediana de 11 de 20 puntos (14). Cabe destacar que este estudio se realizó de forma presencial con las imágenes en formato físico y se otorgó puntuación parcial a las respuestas que consideraron poco específicas. Por lo tanto, la nota obtenida en nuestra muestra piloto es comparable e incluso mayor a la mediana obtenida por estudiantes y residentes en el resto del mundo.

En lo que refiere al análisis de los ítems, *Andre De Champlain et al* (2010) recomiendan que los ítems de una herramienta tengan un índice de dificultad entre 0.3 - 0.7 y un índice de discriminación  $\geq 0.2$  (11). Nuestra herramienta en general tuvo dificultad “intermedia” y discriminación “buena”.

Aunque nuestra herramienta en promedio tuvo dificultad “intermedia”, su valor se alejaba del valor ideal de 0.5 (11). Esto es porque la mayoría de los ítems fueron “muy fáciles”, lo que contrasta con otros estudios en los que se analiza tales índices en preguntas de respuesta múltiple, en los cuales se encontró que alrededor del 80% de las preguntas tienen índices de dificultad entre 0.3-0.7 (33). Esta proporción de preguntas “muy fáciles” en nuestro estudio se puede deber a que las patologías seleccionadas para la herramienta son enseñadas con énfasis en la facultad.

Con respecto al índice de discriminación, la mayoría de nuestros ítems estuvieron en el rango recomendando y el promedio de la herramienta completa estuvo

encima del rango aceptable de 0.2 (11) o incluso de 0.25 como lo propuesto en otros estudios (31).

De los ítems que de acuerdo al índice de dificultad se denominaron “muy difíciles”, la dificultad del NPS puede deberse a que la lesión era muy pequeña (aprox. 4 mm) y por tanto, poco distinguible por alumnos con mínimo entrenamiento, por lo que las respuestas erróneas más frecuentes fueron radiografía normal y tuberculosis. Por otro lado, la dificultad del ítem de bronquiectasias puede deberse a que estas son un patrón difícil de diagnosticar en radiografías simples y que requieren de una tomografía para confirmar, siendo así las respuestas más erróneas frecuentes fibrosis pulmonar y tuberculosis miliar. En contraste con el ítem de nódulo pulmonar solitario, que obtuvo un índice de discriminación “regular”, el ítem de bronquiectasias obtuvo un índice de 0.04. Este valor significa que la pregunta es respondida correctamente con la casi misma frecuencia en el grupo con mejor desempeño y en el grupo con peor desempeño, aportando así ningún poder discriminativo. Por tanto, se debería considerar retirar o cambiar este ítem.

Con respecto a los ítems que según el índice de dificultad se denominaron “muy fáciles”, la “facilidad” del ítem de identificación de sexo, pudo deberse a que la silueta mamaria se marcaba claramente en la radiografía proporcionada. Con respecto a la “facilidad” de los ítems de radiografía normal, neumonía lobar, neumotórax, derrame pleural y tuberculosis pulmonar, esta es comparable a otras investigaciones (24, 34, 35), ya que en estos estudios tales patologías son las que con mayor frecuencia son respondidas correctamente por los médicos en formación. La “facilidad” de los ítems de edema agudo de pulmón y derrame

pericárdico no es comparable con la de otras investigaciones ya que estas patologías tuvieron en otros estudios dificultad intermedia y muy difícil, respectivamente (24, 36).

El ítem de radiografía de tórax normal y de identificación de sexo obtuvieron índices de discriminación muy bajos, por tanto se debería considerar retirar o cambiar ambos ítems, ya que ítems con índice discriminación negativo o cercano a cero disminuyen la validez de la herramienta (31). Con respecto a los ítems neumotórax y derrame pleural, a pesar de tener ambos índices de discriminación fuera de los rangos recomendados, estos fueron limítrofes y se sugiere conservarlos para garantizar que estas áreas se evalúen ya que son patologías que se enseñan con énfasis en la facultad.

Finalmente, la proporción de los ítems “ideales” según los parámetros planteados es concordante con varios estudios que evalúan el desempeño de preguntas de respuesta múltiple, donde se encontró que la proporción de preguntas “ideales” fue de aprox. 15% (37) o 30% (31). Estos ítems podrían presentar esa dificultad ya que son patologías que no son frecuentes y no se enseñan con énfasis, y que por esa misma razón permite discriminar con buen poder entre alumnos con buen desempeño y mal desempeño . De esta manera, debido a que nuestros ítems son “muy fáciles” en su mayoría y en promedio discriminación “buena”, nuestra herramienta tendría mayor utilidad para evaluaciones formativas.

Es interesante que exista una correlación nula entre la seguridad que el participante tiene en sus respuestas y su nota final. Esto concuerda con lo hallado en el estudio de *Jeffrey et al* (2002) donde se comparó la tasa de acierto vs. la seguridad que los alumnos tenían en sus respuestas ( $r=-0.022$ ) (12). En el estudio

de *Eisen et al.* (2006), los estudiantes de medicina obtuvieron seguridad diagnóstica promedio de menos del 25%, la cual aumentaba en proporción directa al nivel de entrenamiento (14). Debido a que mayoría de preguntas fueron “muy fáciles”, esta falta de correlación se puede deber a que se asignaba bajo nivel seguridad a las preguntas respondidas correctamente, lo que a su vez se debe a falta de experiencia en los alumnos y a que eran preguntas de respuesta abierta.

### **Limitaciones**

A pesar de su importancia, nuestro estudio tiene algunas limitaciones. La primera de ellas es que no se estandarizaron los dispositivos (y por tanto resolución, brillo, etc.) en los cuales los alumnos resolvieron la herramienta (por ejemplo, Tablet vs. PC vs. Móvil). Esto puede haber afectado su rendimiento ya que la calidad de la radiografía podría variar de acuerdo con el dispositivo usado. Segundo, al ser una prueba no supervisada, los participantes podrían haber usado recursos externos o compartido respuestas, factores que podrían haber aumentado el puntaje final. Esta limitación puede no ser tan significativa por los hallazgos del estudio de *Miranda et al* (2019) y *Chan et al* (2023), discutido previamente. Por último, el método de Kuder-Richardson 20 que se utilizó para evaluar la confiabilidad de la herramienta asume que todos los ítems tienen la misma dificultad, lo que no podemos asegurar que ocurre en nuestra herramienta. Esto puede explicar por qué la herramienta obtuvo un índice de confiabilidad moderado y no mayor.

## **6. CONCLUSIONES**

El instrumento diseñado en este estudio basado en malla curricular y ASPEFAM obtuvo un índice de validez de contenido aceptable. Obtuvo una confiabilidad moderada para la herramienta completa, la cual se mantuvo en el mismo rango al retirar los ítems con peor desempeño. No hubo correlación entre el nivel de certeza en la respuesta de la herramienta con la nota obtenida en la misma. Finalmente, la herramienta tuvo dificultad “intermedia” y discriminación “buena”, con más de la mitad de ítems con discriminación excelente o buena. Por ello, proponemos que esta herramienta sirva como base para un repositorio virtual de preguntas de carácter formativo, donde los alumnos puedan resolverlo para autoevaluar sus habilidades y destrezas en la interpretación de radiografías de tórax, donde se vayan añadiendo ítems que evalúen las 16 patologías escogidas y así aumentar la confiabilidad del instrumento.

## **7. REFERENCIAS BIBLIOGRÁFICAS**

1. American College of Radiology. ACR standard for the performance of pediatric and adult chest radiography. American College of Radiology: United States, 2001.
2. Speets, A. M., van der Graaf, Y., Hoes, A. W., Kalmijn, S., Sachs, A. P., Rutten, M. J., Gratama, J. W., Montauban van Swijndregt, A. D., & Mali, W. P. (2006). Chest radiography in general practice: indications, diagnostic yield and consequences for patient management. *The British journal of general practice: the journal of the Royal College of General Practitioners*, 56(529), 574–578.
3. Hansell DM, Bankier AA, MacMahon H, McLoud TC, Müller NL, Remy J. Fleischner Society: Glossary of terms for Thoracic Imaging. *Radiology*. 2008;246(3):697–722. doi:10.1148/radiol.2462070712
4. Departamentos Académicos de Clínicas Médicas y Quirúrgicas. Silabo Introducción a la Clínica. Lima: Universidad Peruana Cayetano Heredia; 2019.
5. Departamento Académico de Ciencias Preclínicas y de Apoyo. Silabo Anatomía y Anatomía Radiológica De Tórax, Abdomen y Pelvis. Lima: Universidad Peruana Cayetano Heredia; 2017.
6. Departamento Académico de Clínica Médica. Silabo Clínica Médica I. Lima: Universidad Peruana Cayetano Heredia; 2019.
7. Álvarez Novoa A. Silabo Clínica de Especialidades Pediátricas. Lima: Universidad Peruana Cayetano Heredia;

8. Universidad Peruana Cayetano Heredia. Plan curricular carrera profesional de medicina. Lima: Vicerrectorado Académico de la Universidad Peruana Cayetano Heredia. 2016: p 29.
9. Corral Y. Validez y confiabilidad de los instrumentos de investigación para la recolección de datos. *Revista Ciencias de la Educación*. 2009; 19: 228-247.
10. Backhoff E, Larrazolo N, Rosas M. The Difficulty Level and Discrimination Power of the Basic Knowledge and Skills Examination. 2000.p 14-16.
11. De Champlain AF. A primer on classical test theory and item response theory for assessments in medical education. *Medical Education* [Internet]. 1 de enero de 2010;44(1):109-17. Disponible en: <https://doi.org/10.1111/j.1365-2923.2009.03425.x>
12. D.R. Jeffrey, P.R. Goddard, M.P. Callaway, R. Greenwood. *Chest Radiograph Interpretation by Medical Students*. 2002.
13. Erguera L, De los Santos M, Sierra G. Evaluación de identificación de signos en radiografía de tórax en estudiantes de medicina. 2017.
14. Eisen LA, Berger JS, Hegde A, Schneider RF. Competency in chest radiography. A comparison of medical students, residents, and fellows. *J Gen Intern Med*. 2006 May;21(5):460-5. doi: 10.1111/j.1525-1497.2006.00427.x.
15. Bonett D. Sample Size Requirements for Testing and Estimating Coefficient Alpha. *Journal of Educational and Behavioral Statistics*

- [Internet]. 2002 [citado 27 octubre 2022];( 27(4):335–340. Disponible en:  
<https://doi.org/10.3102/10769986027004335>
16. Imasuen K. Sample size determination in test-retest and cronbach alpha reliability estimates. *British Journal of Contemporary Education*. 2022;(1):17–29.
  17. Salcedo Espinoza, C. Competencias esenciales a lograr en el Internado de Medicina. 2019.
  18. Asociación Peruana de Facultades de Medicina.Tabla de Especificaciones. 2023.P 2-3.
  19. Herrera D, Tamashiro J, Záldivar J. Validación de una herramienta de evaluación de conocimientos y destrezas en el análisis de gases arteriales. 2018.
  20. Lawshe CH. A quantitative approach to content validity. *Personnel psychology*. 1975 Dec 1;28(4):563-75.
  21. Kuder GF, Richardson MW. The theory of the estimation of test reliability. *Psychometrika*. 1937;2(3):151–60. doi:10.1007/bf02288391
  22. Zimmerman DW. Test reliability and the Kuder-Richardson formulas: Derivation from probability theory. *Educational and Psychological Measurement*. 1972;32(4):939–54. doi:10.1177/001316447203200408
  23. De Winter JC, Gosling SD, Potter J. Comparing the Pearson and Spearman correlation coefficients across distributions and sample sizes: A tutorial using simulations and empirical data. *Psychological Methods*. 2016 Sep;21(3):273.

24. Cheung T, Hrianato H, Spanger M, Young A, Wadhwa , Low accuracy and confidence in chest radiograph interpretation amongst junior doctors and medical students. *Internal Medicine Journal*.
25. Loy CT, Irwig L. Accuracy of Diagnostic Tests Read With and Without Clinical Information: A Systematic Review. *JAMA*. 2004;292(13):1602–1609. doi:10.1001/jama.292.13.1602
26. Schreiber MH. The Clinical History as a Factor in Roentgenogram Interpretation. *JAMA*. 1963;185(5):399–401. doi:10.1001/jama.1963.03060050077027
27. Miranda, A.C., Monteiro, C.C., Pires, M.L., & Miranda, L.E. (2019). Habilidades em Diagnóstico Radiológico do Interno em Medicina.
28. Chan JCK, Ahn D. Unproctored online exams provide meaningful assessment of student learning. *Proc Natl Acad Sci U S A*. 2023 Aug;120(31):e2302020120. doi: 10.1073/pnas.2302020120.
29. Tristán-López A. Modificación al modelo de Lawshe para el dictamen cuantitativo de la validez de contenido de un instrumento objetivo. *Av En Medición*. 2008;6:37–4
30. Jimah BB, Appiah AB, Sarkodie BD, Anim D. Competency in chest radiography interpretation by junior doctors and final year medical students at a teaching hospital. *Radiology Research and Practice* [Internet]. 6 de noviembre de 2020;2020:1-7. Disponible en: <https://doi.org/10.1155/2020/8861206>

31. Gajjar S, Sharma R, Kumar P, Rana M. Item and test analysis to identify quality Multiple Choice Questions (MCQS) from an assessment of medical students of Ahmedabad, Gujarat. *Indian Journal of Community Medicine* [Internet]. 1 de enero de 2014;39(1):17. Disponible en: <https://doi.org/10.4103/0970-0218.126347>
32. Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011 Jun 27;2:53-55. doi: 10.5116/ijme.4dfb.8dfd.
33. Kumar D, Jaipurkar R, Shekhar A, Sikri G, Srinivas V. Item Analysis of multiple choice questions: A quality assurance test for an assessment tool. *Medical journal, Armed Forces India* [Internet]. 1 de febrero de 2021;77:S85-9. Disponible en: <https://doi.org/10.1016/j.mjafi.2020.11.007>
34. Satia I, Bashagha S, Bibi A, Ahmed R, Mellor S, Zaman F. Assessing the accuracy and certainty in interpreting chest X-rays in the medical division. *Clin Med (Lond)*. 2013 Aug;13(4):349-52. doi: 10.7861/clinmedicine.13-4-349. PMID: 23908502; PMCID: PMC4954299.
35. Christiansen JM, Gerke O, Karstoft J, Andersen PE. Poor interpretation of chest X-rays by junior doctors. *Dan Med J*. 2014 Jul;61(7):A4875. PMID: 25123122.
36. Mehdipoor G, Salmani F, Arjmand Shabestari A. Survey of practitioners' competency for diagnosis of acute diseases manifest on chest X-ray. *BMC*

Med Imaging. 2017 Aug 18;17(1):49. doi: 10.1186/s12880-017-0222-8.  
PMID: 28821227; PMCID: PMC5563017.

37. Mahjabeen W, Alam S, Hassan U, Zafar T, Butt R, Konain S, et al.  
Difficulty Index, Discrimination Index and Distractor Efficiency in  
Multiple Choice Questions. Annals of PIMS. 1 de enero de  
2017;13(4):310-5.

## 8. TABLAS, GRÁFICOS Y FIGURAS

**Tabla 1.** Puntuación de ítems para criterio 1: calidad técnica de la radiografía

Ítem	Criterio 1						
	V1	V2	V3	V4	V5	V6	V7
Ítem 1	4	4	4	4	4	3	4
Ítem 2	4	3	3	3	3	3	2
Ítem 3	4	4	2	3	4	3	3
Ítem 4	4	4	3	3	4	3	4
Ítem 5	4	4	4	4	4	3	4
Ítem 6	4	4	3	3	4	3	4
Ítem 7	4	3	2	3	3	4	2
Ítem 8	4	4	3	3	4	4	3
Ítem 9	4	4	4	3	2	3	3
Ítem 10	4	4	3	3	3	4	3
Ítem 11	4	4	3	3	4	4	3
Ítem 12	4	4	3	3	3	3	4
Ítem 13	4	4	3	3	3	3	4
Ítem 14	4	3	3	3	3	3	2
Ítem 15	4	4	3	3	3	4	3

Vn: Validador n. Celda en rojo indica ítem no aprobado por validador

**Tabla 2.** Puntuación de ítems para criterio 2: radiografía coherente con respuesta

Ítem	Criterio 2						
	V1	V2	V3	V4	V5	V6	V7
Ítem 1	4	4	4	4	4	4	4
Ítem 2	4	3	3	4	4	3	2
Ítem 3	4	4	4	4	4	4	4
Ítem 4	4	4	3	4	4	4	4
Ítem 5	4	4	3	4	4	4	4
Ítem 6	4	4	4	4	4	4	4
Ítem 7	4	4	4	4	4	3	4
Ítem 8	4	4	4	4	4	4	4
Ítem 9	4	4	4	4	4	4	4
Ítem 10	4	4	4	3	4	3	4
Ítem 11	4	4	3	4	4	1	4
Ítem 12	4	4	3	3	4	3	4
Ítem 13	4	4	4	3	4	4	4
Ítem 14	4	4	4	4	4	4	4
Ítem 15	4	4	4	4	4	4	4

Vn: Validador n. Celda en rojo indica ítem no aprobado por validador

**Tabla 3.** Puntuación de ítems para criterio 3: viñeta clínica induce a la respuesta

Ítem	Criterio 3						
	V1	V2	V3	V4	V5	V6	V7
Ítem 1	4	3	4	4	2	2	2
Ítem 2	4	3	4	4	3	3	2
Ítem 3	4	4	4	4	3	4	4
Ítem 4	4	4	3	4	4	4	4
Ítem 5	4	3	4	4	4	4	2
Ítem 6	4	4	4	4	4	4	4
Ítem 7	4	4	4	4	4	3	3
Ítem 8	4	4	4	4	4	4	4
Ítem 9	4	4	3	4	4	4	2
Ítem 10	4	4	2	4	4	4	4
Ítem 11	4	4	4	4	3	4	4
Ítem 12	4	4	2	4	4	3	4
Ítem 13	4	4	1	4	2	4	3
Ítem 14	4	4	2	4	2	4	3
Ítem 15	4	4	2	4	3	2	3

Vn: Validador n. Celda en rojo indica ítem no aprobado por validador

**Tabla 4.** Puntuación de ítems para criterio 4: redacción de la viñeta clínica

Ítem	Criterio 4						
	V1	V2	V3	V4	V5	V6	V7
Ítem 1	4	4	4	4	4	4	4
Ítem 2	4	3	3	4	4	4	4
Ítem 3	3	4	4	4	4	3	4
Ítem 4	4	4	4	4	4	4	4
Ítem 5	4	4	4	4	4	4	4
Ítem 6	4	4	4	4	4	4	3
Ítem 7	4	2	4	4	4	4	3
Ítem 8	4	4	4	4	4	4	3
Ítem 9	4	4	4	4	4	4	3
Ítem 10	4	4	4	4	4	4	4
Ítem 11	4	4	4	4	4	2	4
Ítem 12	4	4	4	4	4	4	4
Ítem 13	4	4	4	4	4	4	4
Ítem 14	4	4	4	4	4	4	4
Ítem 15	4	4	4	4	4	4	4

Vn: validado n. Celda en rojo indica ítem no aprobado por validador

**Tabla 5.** CVR y CVI por ítem y del instrumento pre-revalidación

	CVR criterio 1	CVR criterio 2	CVR criterio 3	CVR criterio 4	CVI del ítem
Ítem 1	1.00	1.00	-0.14	1.00	0.71
Ítem 2	0.71	0.71	0.71	1.00	0.79
Ítem 3	0.71	1.00	1.00	1.00	0.93
Ítem 4	1.00	1.00	1.00	1.00	1.00
Ítem 5	1.00	1.00	0.71	1.00	0.93
Ítem 6	1.00	1.00	1.00	1.00	1.00
Ítem 7	0.43	1.00	1.00	0.71	0.79
Ítem 8	1.00	1.00	1.00	1.00	1.00
Ítem 9	0.71	1.00	0.71	1.00	0.86
Ítem 10	1.00	1.00	0.71	1.00	0.93
Ítem 11	1.00	0.71	1.00	0.71	0.86
Ítem 12	1.00	1.00	0.71	1.00	0.93
Ítem 13	1.00	1.00	0.43	1.00	0.86
Ítem 14	0.71	1.00	0.43	1.00	0.79
Ítem 15	1.00	1.00	0.43	1.00	0.86
<b>CVI del instrumento</b>					<b>0.88</b>

Celda en rojo indica ítem con un CVR no aprobado

**Tabla 6.** CVR y CVI por ítem y del instrumento post-revalidación

	CVR criterio 1	CVR criterio 2	CVR criterio 3	CVR criterio 4	CVI del ítem
Ítem 1	1.00	1.00	1.00	1.00	1.00
Ítem 2	0.71	0.71	0.71	1.00	0.79
Ítem 3	0.71	1.00	1.00	1.00	0.93
Ítem 4	1.00	1.00	1.00	1.00	1.00
Ítem 5	1.00	1.00	0.71	1.00	0.93
Ítem 6	1.00	1.00	1.00	1.00	1.00
Ítem 7	0.43	1.00	1.00	0.71	0.79
Ítem 8	1.00	1.00	1.00	1.00	1.00
Ítem 9	0.71	1.00	0.71	1.00	0.86
Ítem 10	1.00	1.00	0.71	1.00	0.93
Ítem 11	1.00	0.71	1.00	0.71	0.86
Ítem 12	1.00	1.00	0.71	1.00	0.93
Ítem 13	1.00	1.00	1.00	1.00	1.00
Ítem 14	0.71	1.00	1.00	1.00	0.93
Ítem 15	1.00	1.00	1.00	1.00	1.00
<b>CVI del instrumento</b>					<b>0.93</b>

Celda en verde indica ítem con un CVR re-evaluado

**Tabla 7.** Características del grupo piloto

Características	n = 107	
	n	%
Edad (años)	24	
Sexo		
Femenino	67	62.6%
Masculino	40	37.4%
Año de estudio		
7mo año	107	100.0%
Formación adicional en lectura de radiografías de tórax		
Sí	8	7.6%
No	99	92.4%

**Tabla 8.** Mejor método de aprendizaje para interpretación de radiografías de tórax

Método de aprendizaje	n = 107	
	n	%
Prácticas clínicas	70	65.4%
Clases	13	12.1%
ECOE	8	7.5%
SGP	7	6.5%
TBL	6	5.6%
Otros	3	2.8%

**Tabla 9.** Índice de dificultad e Índice de discriminación de cada ítem

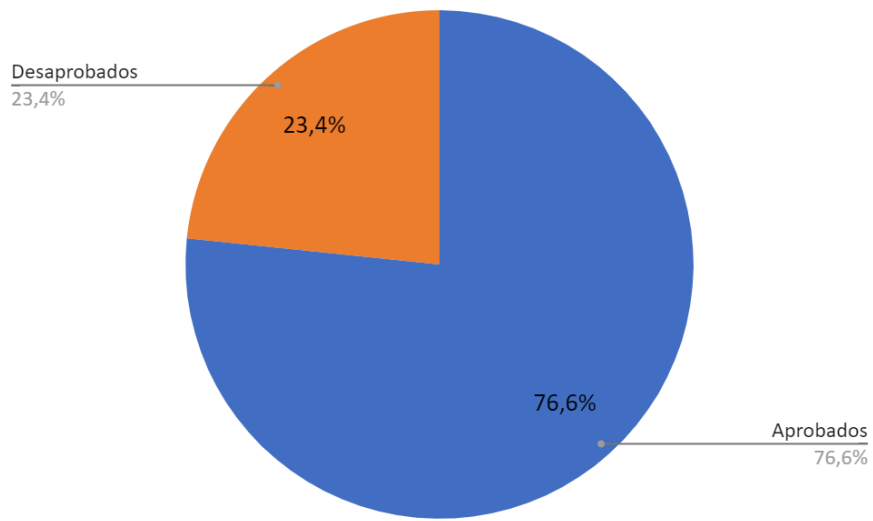
Ítem	Patología	Índice de Dificultad	Índice de Discriminación
1	Normal	0.85	0.19
1.2	Identificar sexo	0.98	0.12
2	Neumonía	0.71	0.31
3	EAP	0.76	0.65
4	Fibrosis pulmonar	0.67	0.38
5	Neumotórax	0.90	0.19
6	Efusión pleural	0.93	0.19
7	Atelectasia masiva	0.39	0.35
8	Derrame pericárdico	0.72	0.38
9	Absceso pulmonar	0.45	0.62
10	TBC miliar	0.63	0.38
11	NPS	0.20	0.27
12	Bronquiectasias	0.11	0.04
13	TBC pulmonar	0.72	0.54
14	Masa pulmonar	0.69	0.62
15	Hidroneumotórax	0.44	0.58

EAP: Edema Agudo de Pulmón, TBC: Tuberculosis, NPS: Nódulo Pulmonar Solitario

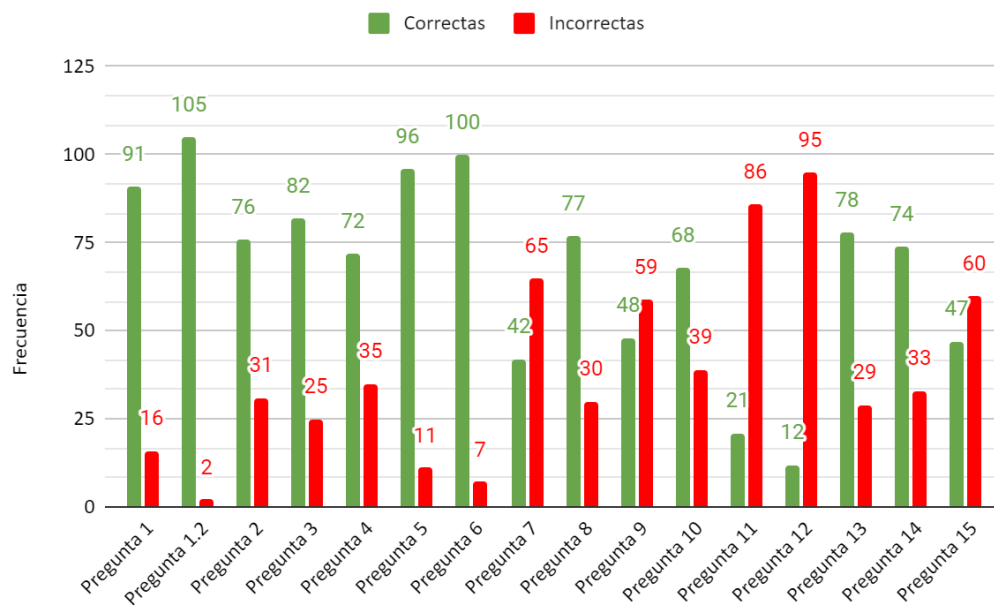
**Tabla 10.** Índice de Dificultad versus Índice de Discriminación

Índice de Dificultad	Índice de Discriminación (V1)				
	>0.39 (Excelente)	0.30-0.39 (Buena)	0.2-0.29 (Regular)	0.0 - 0.2 (Pobre)	<-0.01 (Pésima)
>0.7 (“muy fácil”)	3, 13	2,8		1, 1.2, 5, 6	
0.3-0.7 (“Intermedio”)	9,14,15	4,7,10			
<0.30 (“muy difícil”)			11	12	

**Gráfico 1.** Porcentaje de aprobados

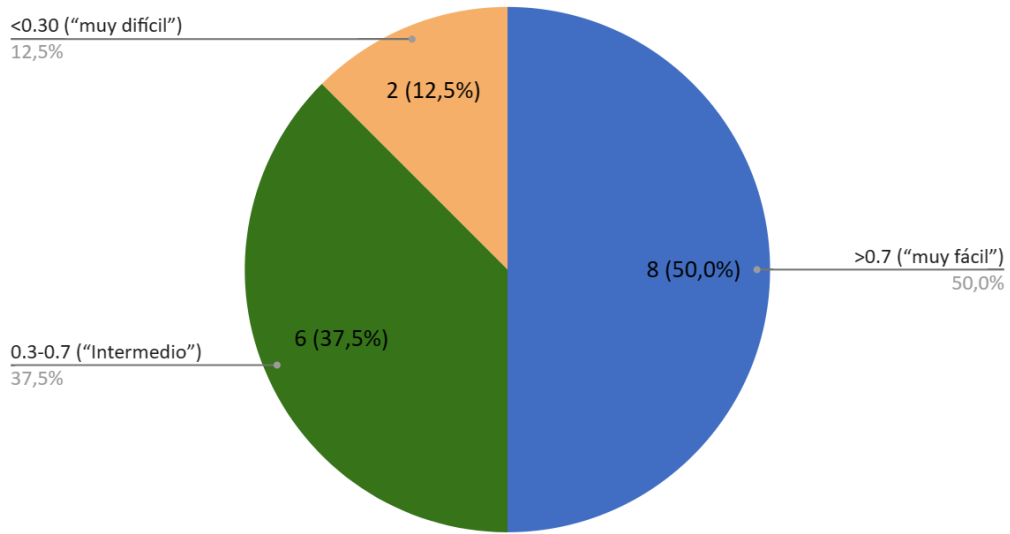


**Gráfico 2.** Frecuencia de correctas e incorrectas por pregunta



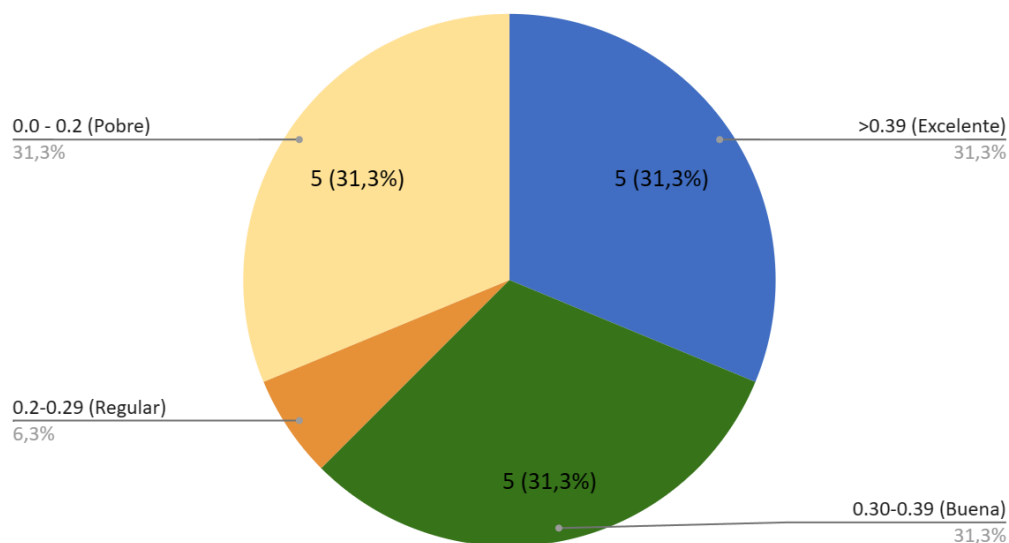
**Gráfico 3.** Distribución de preguntas según Índice de dificultad

## Distribución de Preguntas según Índice de Dificultad



**Gráfico 4.** Distribución de preguntas según Índice de Discriminación

## Distribución de Preguntas según Índice de Discriminación



## 9. ANEXOS

### Anexo 1. Tabla de operacionalización de variables

Variable	Definición operacional
Validación	Se mide mediante el llenado de un formato de validación por 7 médicos expertos los cuales evaluarán la calidad técnica de la radiografía, la coherencia con la respuesta, la inducción a la respuesta y la claridad de la redacción de la viñeta clínica para cada uno de los ítems de la herramienta.
Confiabilidad	Precisión con la cual cada ítem de la herramienta recolecta información acerca de las habilidades y destrezas en la interpretación de las radiografías de tórax, evaluada

	<p>aplicando la herramienta a un grupo piloto de estudiantes y utilizando las notas obtenidas para calcular un índice de confiabilidad mediante el método de Kuder-Richardson 20.</p> <p>Un índice de confiabilidad de 0.01 a 0.20 corresponde a una confiabilidad muy baja, de 0.21 a 0.4 baja, de 0.41 a 0.6 moderada, de 0.61 a 0.8 alta y de 0.81 a 1 muy alta.</p>
--	---

**Anexo 2.** Versión final de la herramienta presentada al grupo piloto

**Pregunta 1**

Paciente de 35 años que acude a la consulta externa como parte de una evaluación médica anual de su trabajo. Niega asma y tabaquismo.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta.

Respuestas correctas: normal, radiografía de tórax normal, radiografía de pulmones normal, radiografía no patológica, sin hallazgos patológicos, sin hallazgos significativos

¿Cuál es el sexo del paciente? Hombre ( ) Mujer ( )

Respuesta

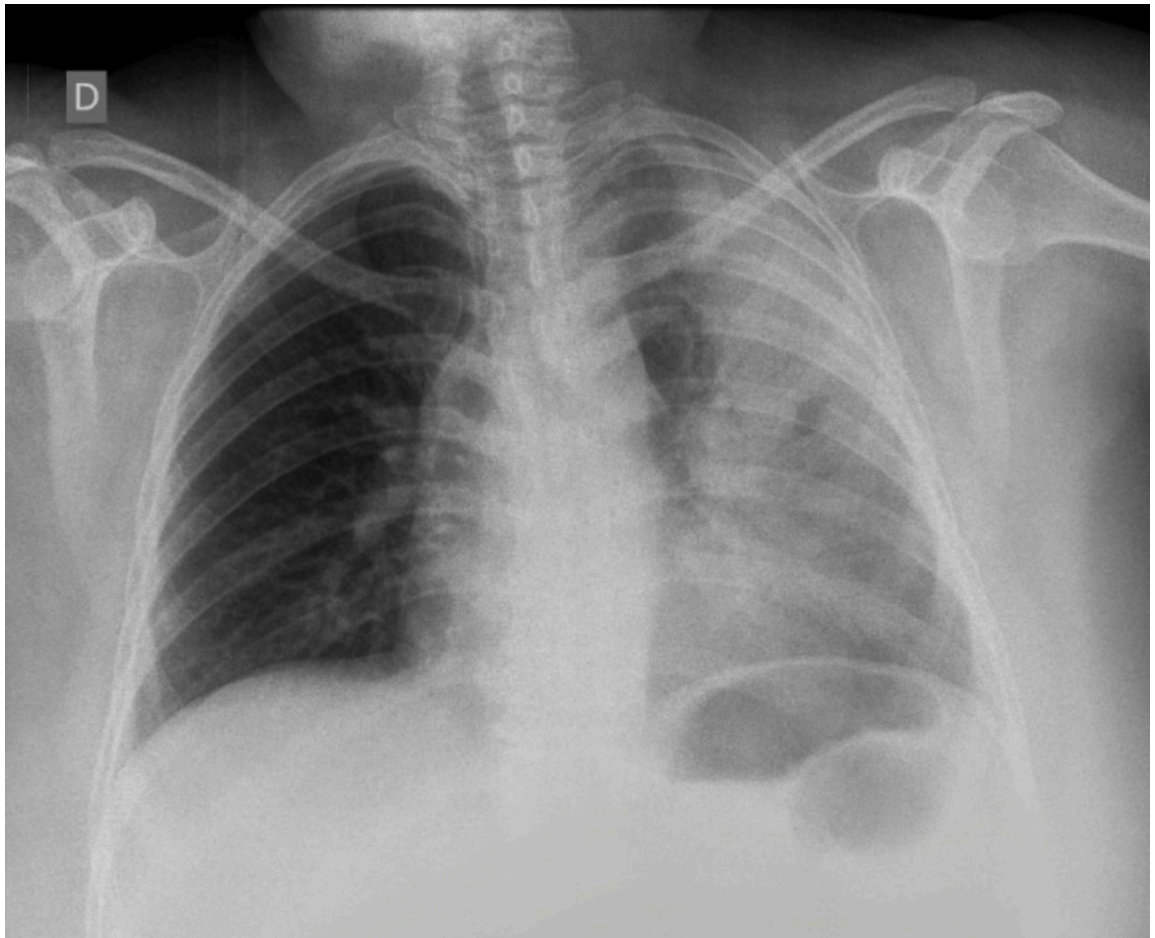
correcta:

Mujer

## Pregunta 2

Mujer de 55 años es llevada a la emergencia porque 3 días antes inicia con fiebre, tos productiva y disnea. El día de hoy se agrega alteración de la conciencia. Tiene antecedente de diabetes mellitus tipo 2 no controlada, niega hospitalizaciones y cirugías recientes.

Se obtiene la siguiente radiografía de tórax:



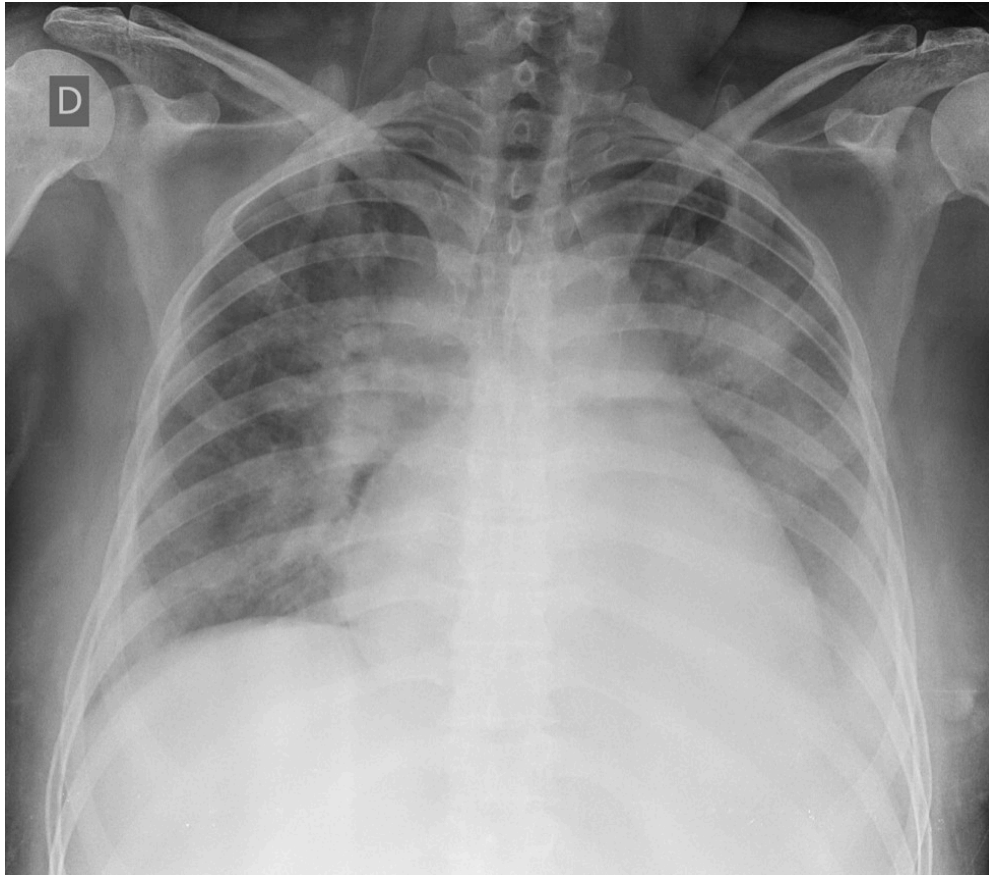
¿Cuál es el diagnóstico principal? Escriba usted su respuesta.

Respuestas correctas: neumonía, opacidades alveolares, consolidado, infiltrados alveolares, neumonía adquirida en la comunidad, neumonía severa, neumonía bacteriana

### Pregunta 3

Mujer de 50 años que acude a la emergencia por 3 días de disnea progresiva y tos nocturna. Tiene antecedentes de hipertensión arterial, diabetes mellitus tipo 2 y dislipidemia sin tratamiento regular.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico principal? Escriba usted su respuesta.

Respuestas correctas: Opacidades alveolares bilaterales, Edema agudo de pulmón, edema pulmonar cardiogénico, opacidades hilio fúgales, infiltrados alveolares bilaterales, infiltrados hilio fúgales, edema en alas de mariposa, opacidades en alas de mariposa, infiltrados en alas de mariposa, congestión pulmonar, insuficiencia cardiaca congestiva, insuficiencia cardiaca descompensada, (con o sin cardiomegalia)

#### Pregunta 4

Varón de 60 años que acude a consulta externa por 6 meses de disnea a esfuerzos moderados y tos seca. No tiene antecedentes de enfermedades previas, exposiciones laborales ni recreacionales, consumo de fármacos ni hábitos nocivos.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: fibrosis pulmonar, fibrosis pulmonar idiopática, enfermedad pulmonar intersticial, enfermedad pulmonar intersticial difusa, opacidades intersticiales, infiltrados intersticiales, opacidades reticulares difusas, neumonía intersticial usual

### Pregunta 5

Paciente mujer de 19 años que acude a la emergencia por disnea de inicio súbito. Tiene antecedente de asma leve para la cual se medica con salbutamol inhalado condicional a síntomas.

Se obtiene la siguiente radiografía de tórax:



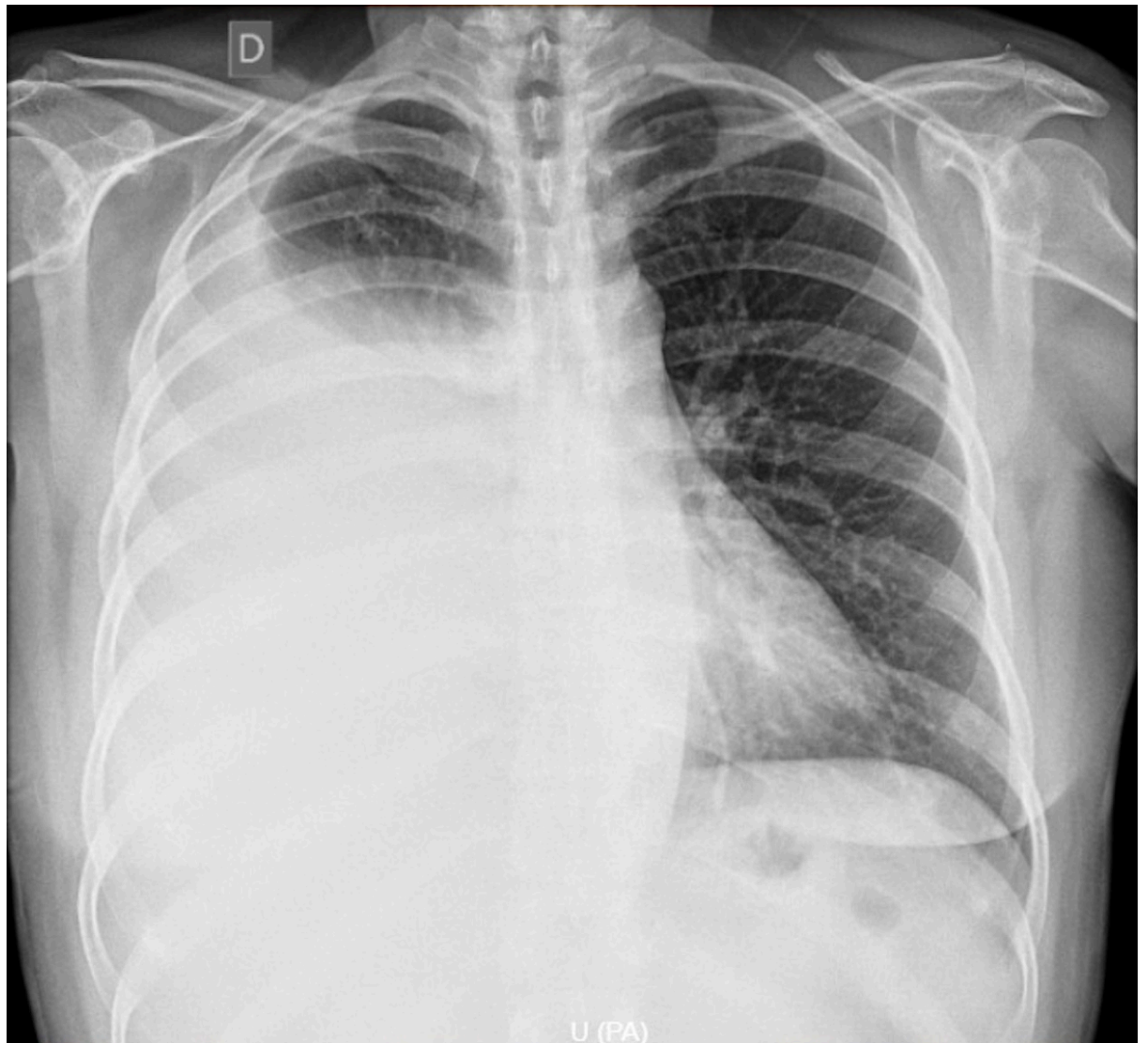
¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Neumotórax, neumotórax espontáneo, neumotórax primario, neumotórax espontáneo primario, aire en espacio pleural, (con o sin colapso pulmonar)

### Pregunta 6

Mujer de 30 años que acude a la emergencia por 3 meses de fiebre, disnea, dolor torácico y baja de peso.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Efusión pleural, efusión pleural derecho, derrame pleural, derrame pleural derecho, derrame pleural masivo, derrame pleural severo, líquido en espacio pleural

**Pregunta 7**

Varón de 85 años que acude a la emergencia por disnea, silbido de pecho, tos y dificultad para expectorar.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Atelectasia masiva, pulmón colapsado, atelectasia, colapso pulmonar, opacidad pulmonar que tracciona mediastino y pinza espacios intercostales

### Pregunta 8

Mujer de 35 años que es traída a la emergencia por 5 días de dolor torácico y disnea. Tiene antecedente de lupus eritematoso sistémico con tratamiento irregular.

Se obtiene la siguiente radiografía de tórax:



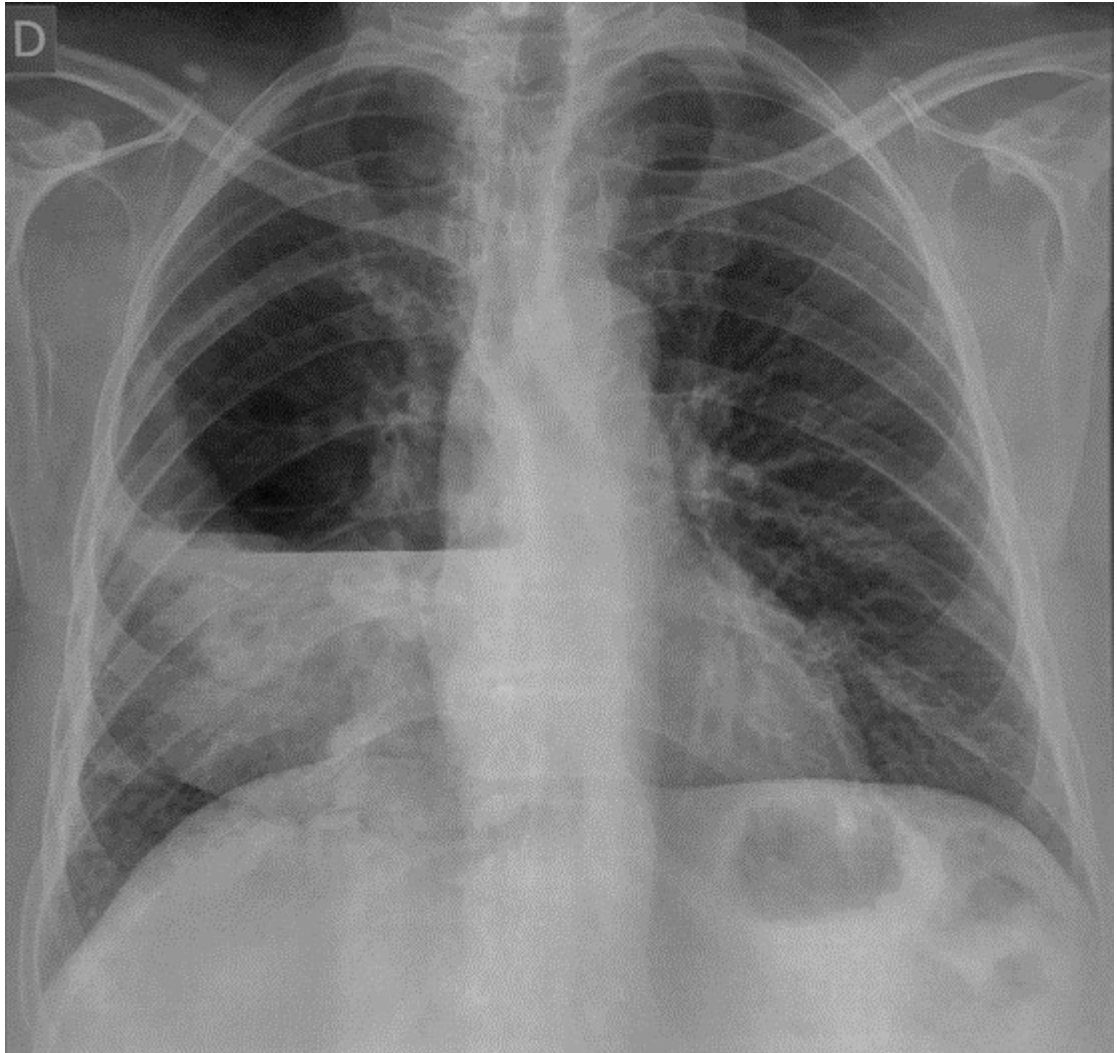
¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Derrame pericárdico, efusión pericárdica, líquido pericárdico, cardiomegalia, aumento de tamaño de silueta cardíaca, silueta cardíaca en botellón, silueta cardíaca en botellón

### Pregunta 9

Mujer de 45 años que acude a la consulta externa por 3 semanas de sensación de alza térmica, tos con expectoración purulenta abundante y astenia. Tiene antecedente de epilepsia con último episodio convulsivo hace 1 mes.

Se obtiene la siguiente radiografía de tórax:



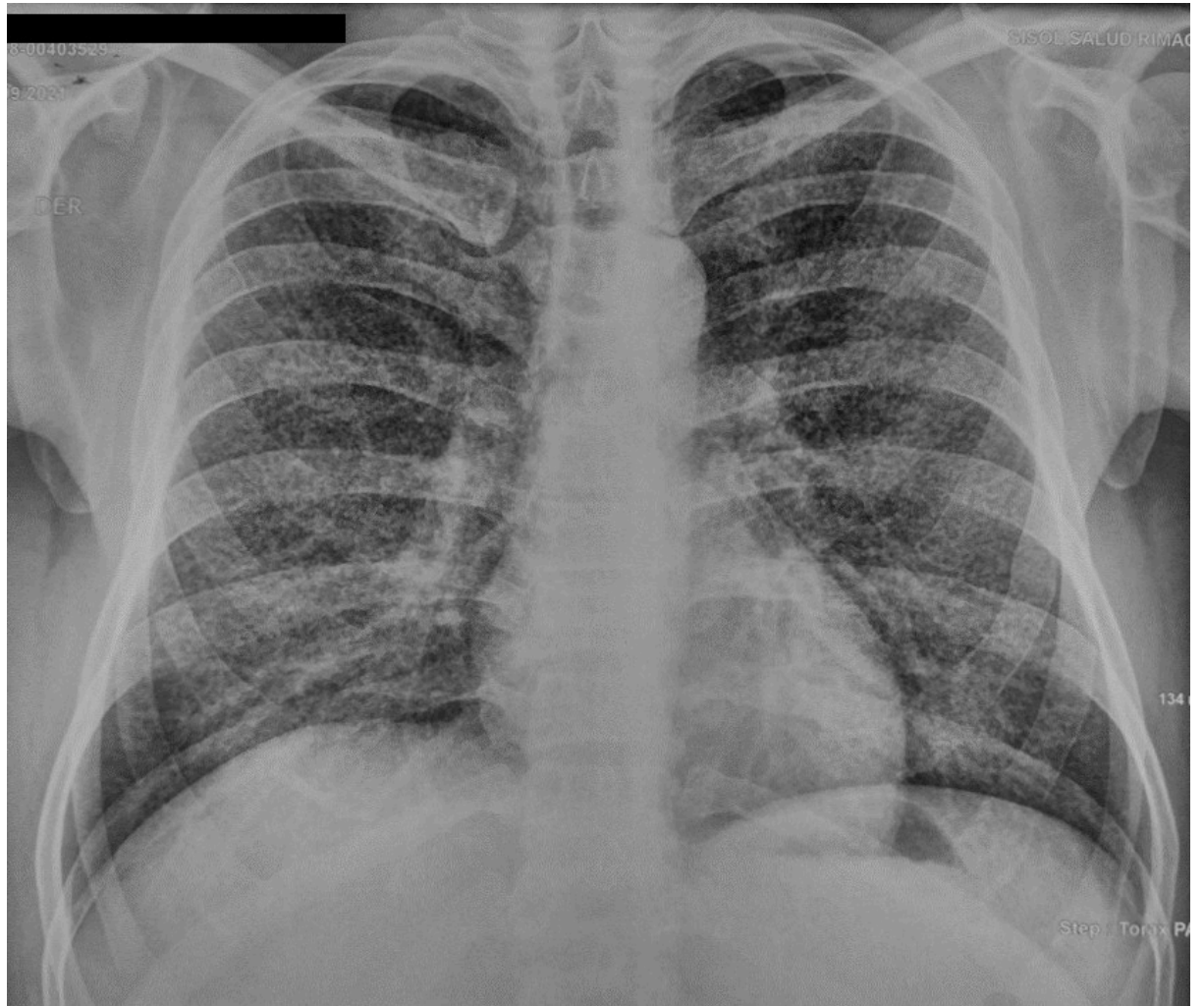
¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Absceso pulmonar, cavidad con nivel hidroaéreo, quiste complicado

### Pregunta 10

Varón de 28 años que acude a la emergencia por 3 semanas de fiebre, tos, disnea y pérdida de peso. Tiene antecedente de infección con VIH sin tratamiento.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

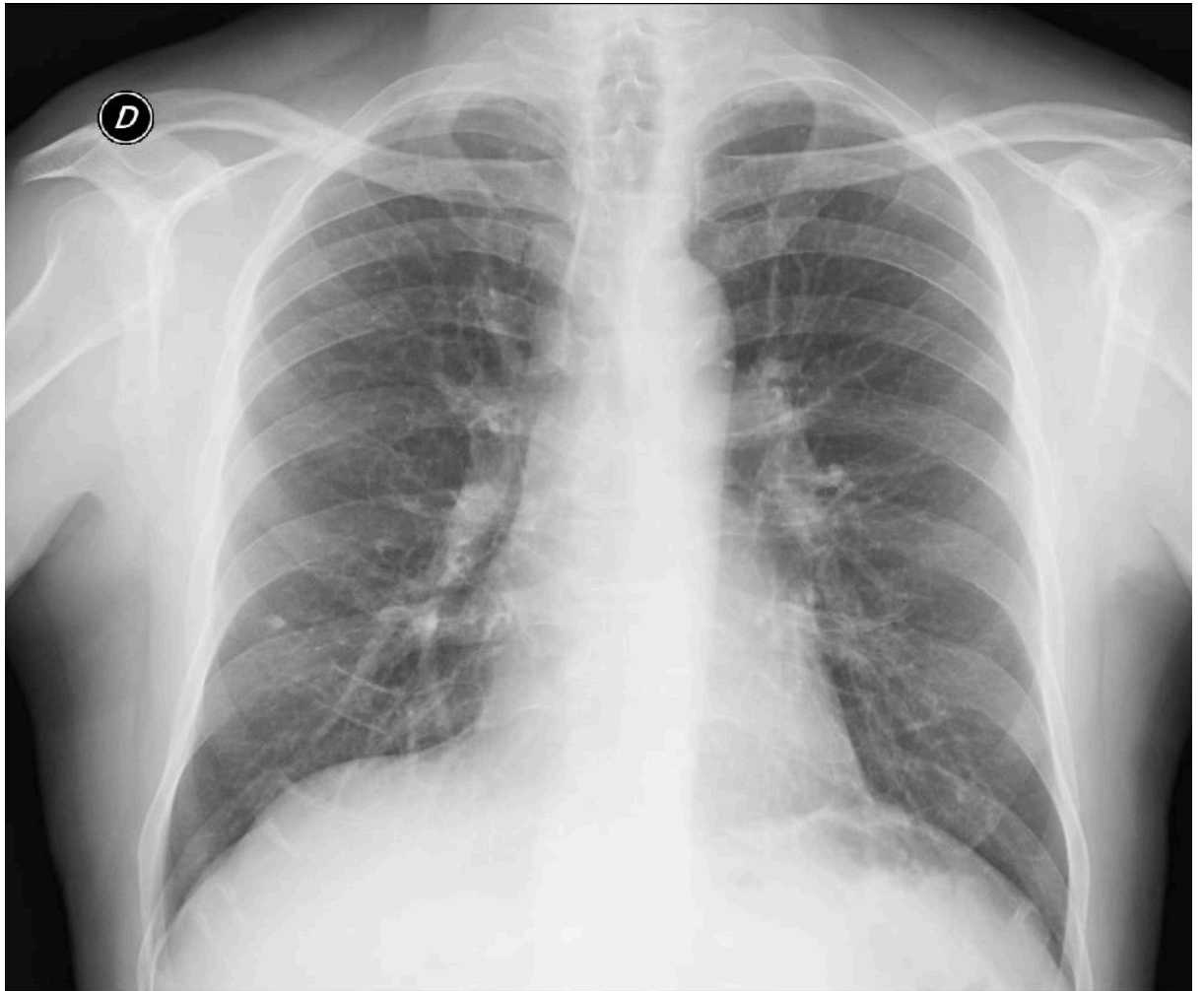
Respuestas correctas: milia, TBC miliar, patrón miliar, micronódulos bilaterales, patrón intersticial micronodular bilateral, opacidades micronodulares, micosis pulmonar, histoplasmosis, paracoccidioidomicosis

**Pregunta 11**

Varón de 40 años que acude a consulta externa para un examen médico general.

Tiene antecedente de contacto con familiar con tuberculosis hace 30 años.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: Nódulo pulmonar solitario, nódulo, granuloma, hamartoma

**Pregunta 12**

Mujer de 20 años de edad que acude a consulta externa por 2 años de episodios repetidos de tos productiva y disnea.

Se obtiene la siguiente radiografía de tórax:



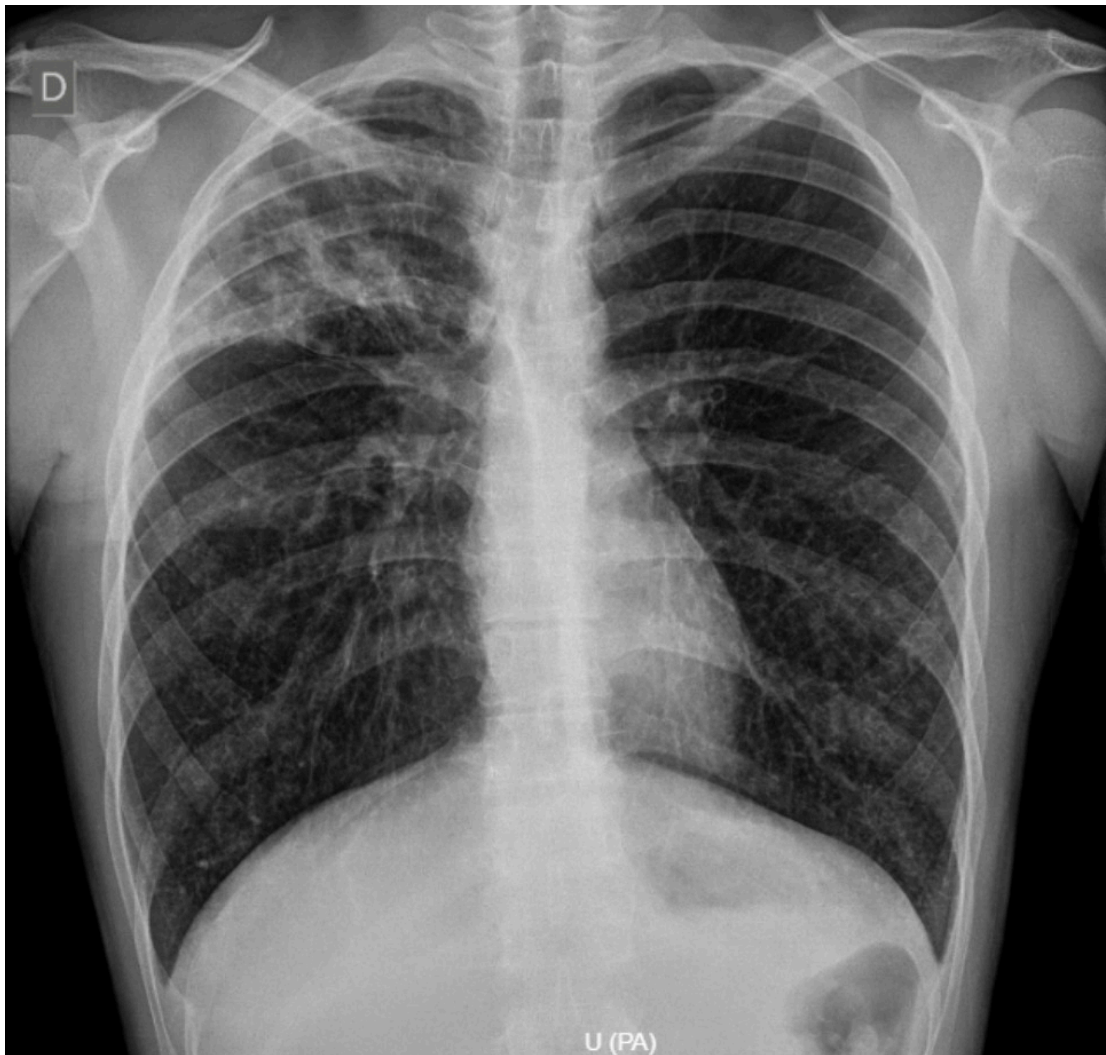
¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: bronquiectasias, dilataciones bronquiales

### Pregunta 13

Varón de 18 años que acude a consulta externa por 3 semanas de tos productiva y baja de peso. No tiene antecedentes de importancia.

Se obtiene la siguiente radiografía de tórax:



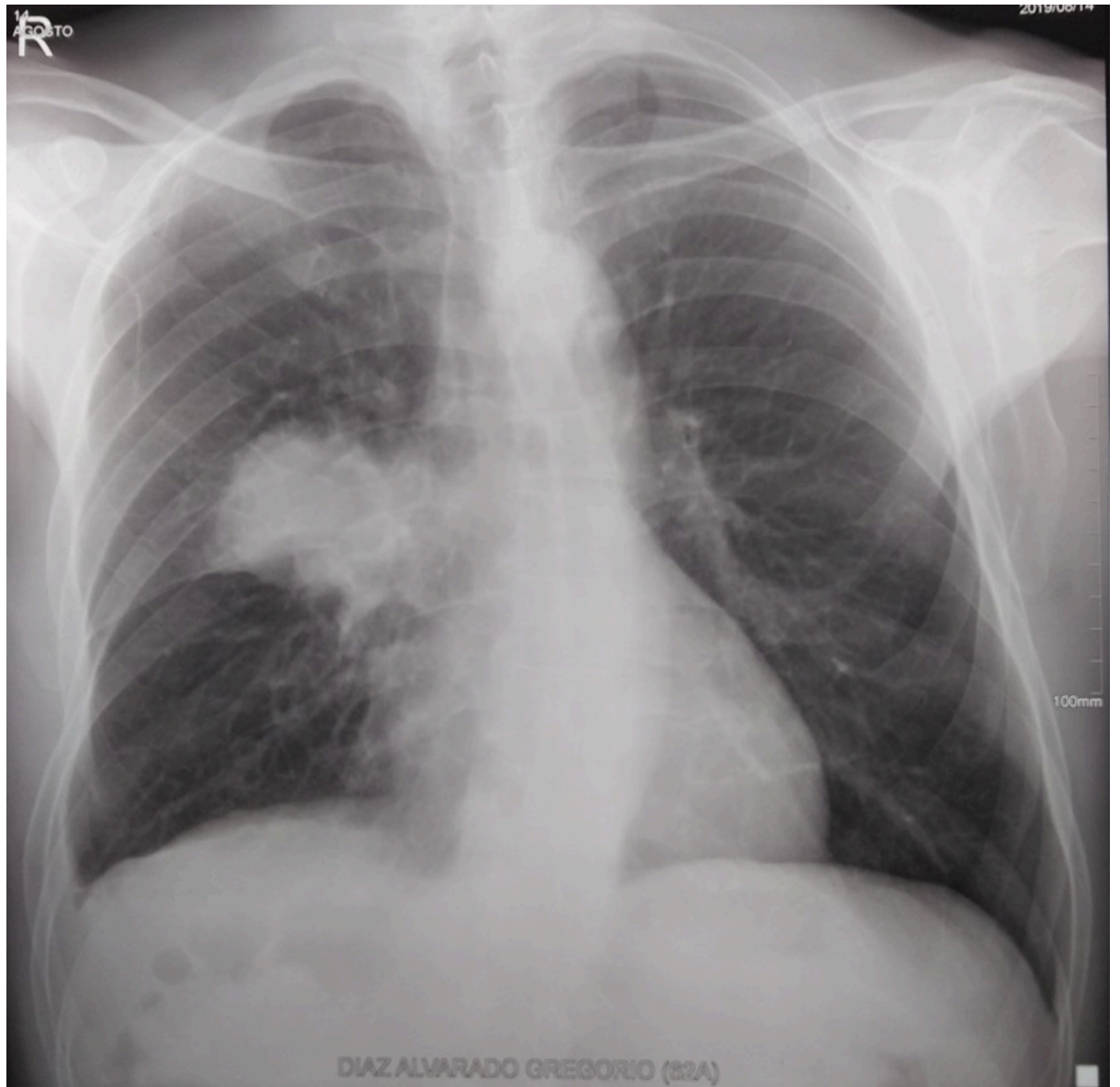
¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: tuberculosis, tuberculosis pulmonar, tuberculosis postprimaria, tuberculosis broncógena, tuberculosis por reactivación, cavernas tuberculosas, cavernas, opacidades alveolo-intersticiales cavitarios, infiltrados alveolo-intersticiales cavitarios, lesión cavitaria

#### Pregunta 14

Varón de 55 años que acude a consulta externa por 8 meses de tos, esputo hemoptoico y baja de peso.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: masa pulmonar, masa pulmonar parahiliar, cáncer de pulmón, neoplasia maligna pulmonar, neoplasia pulmonar, tumor pulmonar.

### Pregunta 15

Varón de 48 años es traído a la emergencia por trauma torácico producto de accidente vehicular.

Se obtiene la siguiente radiografía de tórax:



¿Cuál es el diagnóstico? Escriba usted su respuesta

Respuestas correctas: hidroneumotórax, aire y líquido en el espacio pleural, hemoneumotórax (con o sin traumático)

### **Anexo 3.** Invitación a los expertos validadores

Estimado validador:

Somos un grupo de investigadores de la Universidad Peruana Cayetano Heredia y nos complace invitarlo a participar como experto en un estudio que tiene como objetivo la validación de una herramienta que mide el nivel de interpretación de una radiografía de tórax por estudiantes de medicina.

Esta herramienta consiste en una serie de 15 radiografías de tórax acompañadas de una viñeta clínica (incluyendo 1 normal) en la cual el participante debe escribir su diagnóstico/hallazgo más probable; las radiografías son extraídas del archivo del hospital y contienen diagnósticos que un estudiante de medicina debería ser capaz de diagnosticar. Las radiografías a evaluar son: normal, neumonía, edema pulmonar cardiogénico, fibrosis pulmonar, neumotórax, derrame pleural, atelectasia masiva, taponamiento cardiaco, absceso pulmonar, TBC miliar, nódulo pulmonar solitario, bronquiectasias, TBC secundaria, masa pulmonar, hidroneumotórax.

Su contribución como experto consistirá en evaluar cada ítem de la herramienta (1 ítem = radiografía + viñeta clínica) en las siguientes 5 categorías, poniéndole una nota del 1 al 4 en el formato de validación:

1. Calidad técnica de la radiografía
2. Hallazgos radiográficos consistentes con respuestas
3. Si es que la viñeta clínica induce a la respuesta
4. Redacción de la viñeta clínica
5. Si es que está de acuerdo con el listado de respuestas posibles

Además, si es que desea hacer algún comentario o quisiera que cambiemos la viñeta clínica, radiografía o respuestas posibles, puede incluirlo en la sección de observaciones en el formato de validación.

### Anexo 3. Formato de validación

ITEM	Criterios a evaluar										Observaciones
	Calidad Técnica de la radiografía aceptable (coloque un número del 1 al 4 siendo 1: muy mala calidad, 2: mala calidad, 3: calidad aceptable, 4: buena calidad)	Observaciones	Radiografía coherente con respuesta (coloque un número del 1 al 4 siendo 1: muy poco coherente, 2: poco coherente, 3: coherente, 4: muy coherente)	Observaciones	Vineña clínica induce a la respuesta (coloque un número del 1 al 4 siendo 1: respuesta inducida fácilmente por la vineña, 2: respuesta podría ser inducida por la vineña, 3: respuesta difícilmente inducida por la vineña, 4: respuesta no se induce por la vineña)	Observaciones	Redacción clara de la pregunta y vineña clínica (coloque un número del 1 al 4 siendo 1: redacción muy mala, 2: redacción mala, 3: redacción aceptable, 4: redacción buena)	Observaciones	Esta de acuerdo con las alternativas de respuestas correctas: SI O NO	Observaciones	
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											
11											
12											
13											
14											
15											
Validador experto											
Validado por											
Fecha											
Firma Y Sello											

## Anexo 4. Consentimiento informado



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

VICERRECTORADO DE INVESTIGACIÓN

Validación de una herramienta para la evaluación de los conocimientos y las destrezas para la interpretación de la radiografía de tórax

Investigadores: Blas Miguel Cornejo Esparza, Sebastian Raul Orbegoso Aliaga, Juan Felipe Vasquez Rioja  
**CONSENTIMIENTO INFORMADO**

### Propósito del estudio:

Se le invita a participar en un estudio para la validación de una herramienta la cual busca medir el nivel de interpretación de una radiografía de tórax por estudiantes de medicina de último año de la Universidad Peruana Cayetano Heredia. Este es un estudio realizado por investigadores de la Universidad Peruana Cayetano Heredia.

### Procedimientos:

Si usted decide participar del estudio, se realizará lo siguiente:

Deberá completar una encuesta anónima donde se le mostrarán 13 radiografías de tórax de distintas patologías pulmonares, pleurales y cardíacas para las cuales deberá precisar un diagnóstico en la casilla de respuesta libre. Luego de haber escrito su respuesta deberá indicar el nivel de seguridad que tiene usted en la misma mediante una escala del 1 al 4, 1 siendo muy inseguro y 4 siendo muy seguro. Se le pide que utilice únicamente sus conocimientos y no apuntes u otros medios de información tales como internet, libros y artículos científicos. El cuestionario tomará un tiempo aproximado de 45 minutos por vía de preguntas de respuesta abierta en un documento Google Forms.

### Riesgos:

No existe ningún riesgo para los participantes del estudio.

### Beneficios:

No existen beneficios para los participantes del estudio.

### Costos y compensaciones:

No es necesario pagar por participar del estudio ni tampoco recibirá compensación económica por hacerlo.

### Confidencialidad:

Se le garantiza que la información brindada en el cuestionario es confidencial y solo el grupo de investigadores tendrá acceso a sus respuestas mas no a su identidad.

### Derechos del participante:

Si decide participar del estudio puede retirarse en cualquier momento sin consecuencias por hacerlo. Ante cualquier duda comunicarse con el equipo de investigadores del estudio o contacte al investigador principal del estudio Blas Miguel Cornejo Esparza al número [REDACTED] Si tiene preguntas sobre los aspectos éticos del estudio o siente que ha sido tratado injustamente se puede comunicar con el Dr. Manuel Raúl Pérez Martinot, presidente del Comité Institucional de Ética en Investigación de la Universidad Peruana Cayetano Heredia al teléfono 01-319000 anexo 201355 o al correo electrónico [orvei.ciei@oficias-upch.pe](mailto:orvei.ciei@oficias-upch.pe)

### Declaración y/o consentimiento

Acepto voluntariamente participar del estudio y entiendo las actividades en las cuales tomaré parte. Entiendo que si decido no participar me puedo retirar en cualquier momento.

Si (  ) No (  )

### Declaración del investigador

Declaro que el participante ha leído la descripción del proyecto, ha aclarado sus dudas sobre el estudio y ha decidido participar voluntariamente en él. Se le ha informado que los datos que provea se mantendrán anónimos y que los resultados del estudio serán utilizados para fines de investigación.

Versión 1.3 de fecha 31 de julio 2023