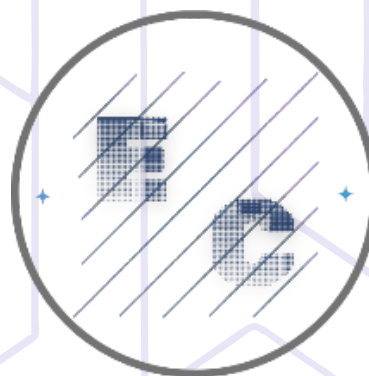
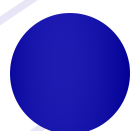
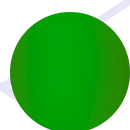




# REVISTA MULTIDISCIPLINAR EPISTEMOLOGÍA DE LAS CIENCIAS



**Universidad de Sonsonate**



**Volumen 3, Número 3**



**Julio - Septiembre, 2026**  
**Edición Especial**

CROSSREF PREFIX DOI: 10.71112

ISSN: 3061-7812, [www.omniscens.com](http://www.omniscens.com)

Revista Multidisciplinar Epistemología de las Ciencias

Volumen 3, Número 3  
Julio-Septiembre 2026  
Edición Especial

Publicación trimestral  
Hecho en México

La Revista Multidisciplinar Epistemología de las Ciencias acepta publicaciones de cualquier área del conocimiento, promoviendo una plataforma inclusiva para la discusión y análisis de los fundamentos epistemológicos en diversas disciplinas. La revista invita a investigadores y profesionales de campos como las ciencias naturales, sociales, humanísticas, tecnológicas y de la salud, entre otros, a contribuir con artículos originales, revisiones, estudios de caso y ensayos teóricos. Con su enfoque multidisciplinario, busca fomentar el diálogo y la reflexión sobre las metodologías, teorías y prácticas que sustentan el avance del conocimiento científico en todas las áreas.

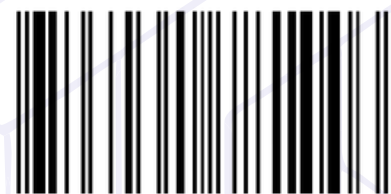
Contacto principal: [admin@omniscens.com](mailto:admin@omniscens.com)

Las opiniones expresadas por los autores no necesariamente reflejan la postura del editor de la publicación

Se autoriza la reproducción total o parcial del contenido de la publicación sin previa autorización de la Revista Multidisciplinar Epistemología de las Ciencias siempre y cuando se cite la fuente completa y su dirección electrónica.



Copyright © 2026: Los autores

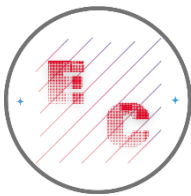


9773061781003

---

### Cintillo legal

Revista Multidisciplinar Epistemología de las Ciencias Vol. 3, Núm. 3, julio-septiembre 2026, edición especial, es una publicación trimestral editada por el Dr. Moises Ake Uc, C. 51 #221 x 16B , Las Brisas, Mérida, Yucatán, México, C.P. 97144 , Tel. 9993556027, Web: <https://www.omniscens.com>, [admin@omniscens.com](mailto:admin@omniscens.com), Editor responsable: Dr. Moises Ake Uc. Reserva de Derechos al Uso Exclusivo No. 04-2024-121717181700-102, ISSN: 3061-7812, ambos otorgados por el Instituto Nacional del Derecho de Autor (INDAUTOR). Responsable de la última actualización de este número, Dr. Moises Ake Uc, fecha de última modificación, 1 julio 2026.



**Revista Multidisciplinar Epistemología de las Ciencias**

**Vol. 3, Núm. 3, 2026, julio-septiembre, Edición Especial**

**DOI:** <https://doi.org/10.71112/5vy5w189>

**APLICACIÓN DE MODELOS DE INTELIGENCIA ARTIFICIAL EN LA  
INTERPRETACIÓN DE RADIOGRAFÍAS DE TÓRAX PARA EL DIAGNÓSTICO DE  
TUBERCULOSIS PULMONAR**

**APLICACIÓN O ARTIFICIAL INTELIGENCIE MODELOS IN CHEST X-RAY  
INTERPRETATION FOR PULMONARY TUBERCULOSIS DIAGNOSIS**

**Nelson Magdiel López Díaz**

**Edgardo Josué Ramos Rivas**

**Jacqueline Stefany Ramos Torres**

**El Salvador**

## **Aplicación de modelos de inteligencia artificial en la interpretación de radiografías de tórax para el diagnóstico de tuberculosis pulmonar**

### **Aplicación o artificial Inteligencie modelos in mhest x-ray Interpretation for pulmonary tuberculosis diagnosis**

Nelson Magdiel López Díaz<sup>a,\*</sup>

[nel070293@hmail.com](mailto:nel070293@hmail.com)

<https://orcid.org/0009-0007-5012-2160>

Edgardo Josué Ramos Rivas<sup>a</sup>

[joramedins@gmail.com](mailto:joramedins@gmail.com)

<https://orcid.org/0000-0002-1820-1674>

Jacqueline Stefany Ramos Torres<sup>b</sup>

[ramostorres964@gmail.com](mailto:ramostorres964@gmail.com)

<https://orcid.org/0009-0004-3563-7324>

\*Autor de correspondencia: [nel070293@hmail.com](mailto:nel070293@hmail.com), <sup>a</sup>Instituto Nacional de Salud, <sup>b</sup>Universidad de El Salvador, El Salvador.

## **RESUMEN**

La tuberculosis pulmonar (TBP) continúa siendo una causa relevante de morbilidad y mortalidad a nivel mundial. La inteligencia artificial (IA), especialmente mediante técnicas de aprendizaje profundo, se ha posicionado como una herramienta prometedora para el diagnóstico automatizado mediante radiografías de tórax. El objetivo de esta revisión de alcance fue mapear y caracterizar la evidencia científica sobre modelos de IA aplicados al diagnóstico de TBP. Se realizó una búsqueda sistemática en seis bases de datos internacionales siguiendo la metodología del Instituto Joanna Briggs y las directrices PRISMA-ScR. Se incluyeron veinte estudios publicados entre 2020 y 2025, predominando arquitecturas CNN, híbridas con transformers y modelos de segmentación contextual. Los algoritmos

alcanzaron exactitudes entre 96 % y 99.9 %. La IA mostró un desempeño superior al juicio humano en entornos controlados, aunque persisten desafíos en estandarización, validación clínica y representatividad geográfica.

**Palabras clave:** Tuberculosis pulmonar; Inteligencia artificial; Deep learning; Machine learning; Revisión de alcance.

## ABSTRACT

Pulmonary tuberculosis (PTB) remains a major cause of infectious morbidity and mortality worldwide. Artificial intelligence (AI), particularly through deep learning techniques, has emerged as a promising tool for automated diagnosis using chest X-ray imaging. This scoping review aimed to map and characterize the available scientific evidence on AI models applied to PTB diagnosis. A systematic search was conducted across six international databases following the Joanna Briggs Institute methodology and PRISMA-ScR guidelines. Twenty studies published between 2020 and 2025 were included, predominantly employing convolutional neural networks (CNN), hybrid architectures with transformers, and contextual segmentation models. Reported accuracies ranged from 96% to 99.9%. AI demonstrated diagnostic performance superior to human interpretation under controlled conditions, although challenges remain regarding standardization, clinical validation, and geographic representativeness.

**Keywords:** Pulmonary tuberculosis; Artificial intelligence; Deep Learning; Machine Learning; Scoping review.

Recibido: 16 febrero 2026 | Aceptado: 30 junio 2026 | Publicado: 1 julio 2026

## INTRODUCCIÓN

La tuberculosis pulmonar (TBP) continúa siendo una de las principales causas de morbilidad y mortalidad infecciosa a escala global. En 2023, la Organización Mundial de la Salud (OMS) estimó 10.8 millones de casos y alrededor de 1.25 millones de muertes, consolidándose como la infección más letal después de la fase aguda de la pandemia por COVID-19 y evidenciando persistentes brechas en el diagnóstico oportuno y el acceso al tratamiento, especialmente en países de ingresos bajos y medios (Global Tuberculosis Reporta 2024, s. f.; Tuberculosis (TB), s. f.). Esta carga epidemiológica ha impulsado la búsqueda de herramientas tecnológicas que mejoren la detección temprana, reduzcan errores diagnósticos y optimicen los recursos disponibles en los sistemas de salud.

En este escenario, la inteligencia artificial (IA), particularmente mediante el aprendizaje profundo aplicado a radiografías de tórax (CXR), ha emergido como una alternativa prometedora para el tamizaje y diagnóstico de la tuberculosis en contextos con escasez de radiólogos y alta demanda de estudios. Desde 2021, la OMS ha recomendado el uso de software de detección asistida por computadora (CAD) como apoyo para la interpretación de CXR en programas de cribado. Estas guías, actualizadas con cada nueva versión tecnológica, destacan que la combinación CXR + CAD es clave para reducir la brecha de casos no detectados, siempre que la implementación sea calibrada y validada en los contextos locales (Use o Computer-Aided Detección Software for Tuberculosis Screening, s. f.; WHO Aproes Six Software Productos for Computer-Aided Detección o TB on Chest X-Ray, s. f.).

El desarrollo de la IA en medicina ha dejado de ser un concepto teórico para convertirse en una herramienta aplicada de alto impacto. En campos como la radiología, dermatología, oftalmología y patología digital, los modelos de aprendizaje automático y profundo han demostrado capacidad para reconocer patrones complejos con precisión comparable o incluso superior a la de los especialistas humanos. Esta revolución tecnológica

está transformando los paradigmas del diagnóstico clínico, favoreciendo la detección temprana de enfermedades, la reducción del tiempo diagnóstico y la consolidación de una medicina más personalizada y basada en datos.

Entre 2020 y 2025, revisiones sistemáticas y metaanálisis han reportado que los sistemas CAD basados en IA alcanzan desempeños diagnósticos elevados, con áreas bajo la curva (AUC) frecuentemente superiores a 0.85 y resultados equiparables al juicio de expertos. Evaluaciones comparativas de soluciones comerciales como CAD4TB, ox y Lunit INSIGHT han mostrado rendimientos similares e incluso beneficios adicionales cuando el algoritmo se combina con la lectura humana. Sin embargo, la literatura enfatiza la necesidad de validaciones externas y calibraciones adaptadas a variables demográficas y epidemiológicas locales, como edad, sexo y prevalencia de la enfermedad (Han et al., 2025; Reagan et al., 2024; Sung et al., 2025).

Pese a estos avances, la investigación enfrenta desafíos metodológicos y de implementación. Entre ellos destacan la dependencia de conjuntos de datos públicos (Montgomery, Shenzhen, TBX11K) que no siempre representan la diversidad clínica real, la heterogeneidad en los protocolos de entrenamiento y validación, las variaciones en las métricas de desempeño (exactitud, sensibilidad, especificidad, AUC) y los cambios de dominio relacionados con equipos o condiciones de imagen que afectan la generalización de los modelos. A ello se suman desafíos éticos y operativos vinculados a la gobernanza de datos, la transparencia algorítmica y la equidad de acceso (Calibrating Computer-Aided Detección (CAD) for TB, s. f.; Index o public/Tuberculosis-Chest-X-ray-Datasets/, s. f.; TBX11K - Dataset Ninja, s. f.; Vbookshelf - TB X-Rays with Bounding Boxes, s. f.; Jaeger et al., 2014; Santosh et al., 2022).

El incremento global en la demanda de estudios de imagen, la limitada disponibilidad de personal especializado y la necesidad de mejorar la eficiencia diagnóstica consolidan el rol de la IA como apoyo clínico viable. No obstante, la heterogeneidad de los modelos, la ausencia

de estándares universales de validación y las limitaciones en la representatividad de datos siguen siendo obstáculos para su implementación rutinaria en sistemas de salud de bajos recursos.

Actualmente, no existen revisiones de alcance que sinteticen de manera integral la evidencia disponible sobre los modelos de inteligencia artificial aplicados al diagnóstico de tuberculosis en imágenes clínicas. Una búsqueda preliminar en MEDLINE y en Open Science Framework no identificó revisiones sistemáticas ni de alcance recientes o en curso sobre el tema. Ante este vacío, se plantea la necesidad de una revisión estructurada que compile la evidencia existente, evalúe los avances tecnológicos y determine las brechas de conocimiento que orienten futuras investigaciones, guías clínicas y políticas públicas.

El presente estudio tiene como objetivo mapear y caracterizar la evidencia científica disponible sobre el uso de modelos de inteligencia artificial aplicados al diagnóstico de tuberculosis pulmonar mediante imágenes médicas, describiendo los tipos de algoritmos (redes neuronales convolucionales, híbridos y de segmentación), las bases de datos empleadas, las métricas de desempeño y los desafíos metodológicos y clínicos asociados a su implementación.

## METODOLOGÍA

La presente investigación se desarrolló como una revisión de alcance (*scoping review*) siguiendo la metodología del Instituto Joanna Briggs (JBI, 2020) para la síntesis de evidencia científica, y se reporta conforme a las directrices PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews). Este diseño permitió mapear y describir la evidencia disponible sobre modelos de inteligencia artificial (IA) aplicados al diagnóstico de tuberculosis pulmonar (TBP), identificar vacíos de conocimiento y orientar futuras líneas de investigación (Tricco et al., 2018).

El protocolo fue registrado en la plataforma Open Science Framework (OSF) bajo su respectivo número identificador, y todas las etapas (búsqueda, selección y extracción de datos) fueron realizadas de forma independiente por dos revisores, resolviéndose las discrepancias mediante consenso o la intervención de un tercer evaluador.

Los criterios de inclusión se definieron de acuerdo con los componentes PCC (Participantes, Concepto y Contexto) (James Cook University Library, 2021). Se incluyeron estudios primarios que evaluaron a personas de cualquier edad o condición clínica sometidas a procedimientos diagnósticos para TBP. Se consideraron investigaciones que describieron, desarrollaron o validaron modelos de IA aplicados al diagnóstico o detección de tuberculosis mediante análisis de imágenes médicas (radiografía de tórax, tomografía computarizada o microscopía digital). Se excluyeron estudios en animales, revisiones narrativas, editoriales y trabajos sin texto completo disponible.

El alcance geográfico fue global, incorporando estudios de diferentes contextos asistenciales (hospitales, programas de cribado, laboratorios o entornos penitenciarios) y fuentes de datos tanto públicas (Shenzhen, Montgomery, TBX11K, RSNA) como locales. No se establecieron restricciones de idioma ni de fecha de publicación.

La búsqueda sistemática se realizó en seis bases de datos principales MEDLINE (PubMed), Embase, Web of Science, IEEE Xplore, ACM Digital Library y Biblioteca Virtual en Salud (BVS), complementada con búsquedas en Epistemonikos y Google Scholar. Se emplearon descriptores MeSH y DeCS relacionados con “Tuberculosis”, “Artificial Intelligence”, “Machine Learning”, “Deep Learning” y “Diagnosis”. Los registros identificados se gestionaron mediante Zotero y Rayyan, donde se eliminaron duplicados y se aplicó un proceso de cribado doble ciego por título, resumen y texto completo. Se documentaron los motivos de exclusión en cada etapa y se presentó el flujo de selección según el diagrama PRISMA-ScR. La extracción de datos se efectuó con una plantilla estructurada en Microsoft Excel, que incluyó variables

bibliográficas, características metodológicas y métricas de desempeño (exactitud, sensibilidad, especificidad, exactitud). La evaluación de calidad metodológica se realizó de manera descriptiva utilizando las listas de verificación JBI correspondientes al tipo de estudio, sin excluir artículos por su puntaje.

Finalmente, la síntesis de resultados se efectuó mediante análisis descriptivo y narrativo, apoyado en gráficos, tablas y mapas elaborados en RStudio 4.4.1. Los modelos de IA se agruparon por tipo de arquitectura (CNN, híbridos, segmentación contextual y ensambles), generando una visión panorámica del desempeño, las tendencias tecnológicas y los principales desafíos para su aplicación clínica.

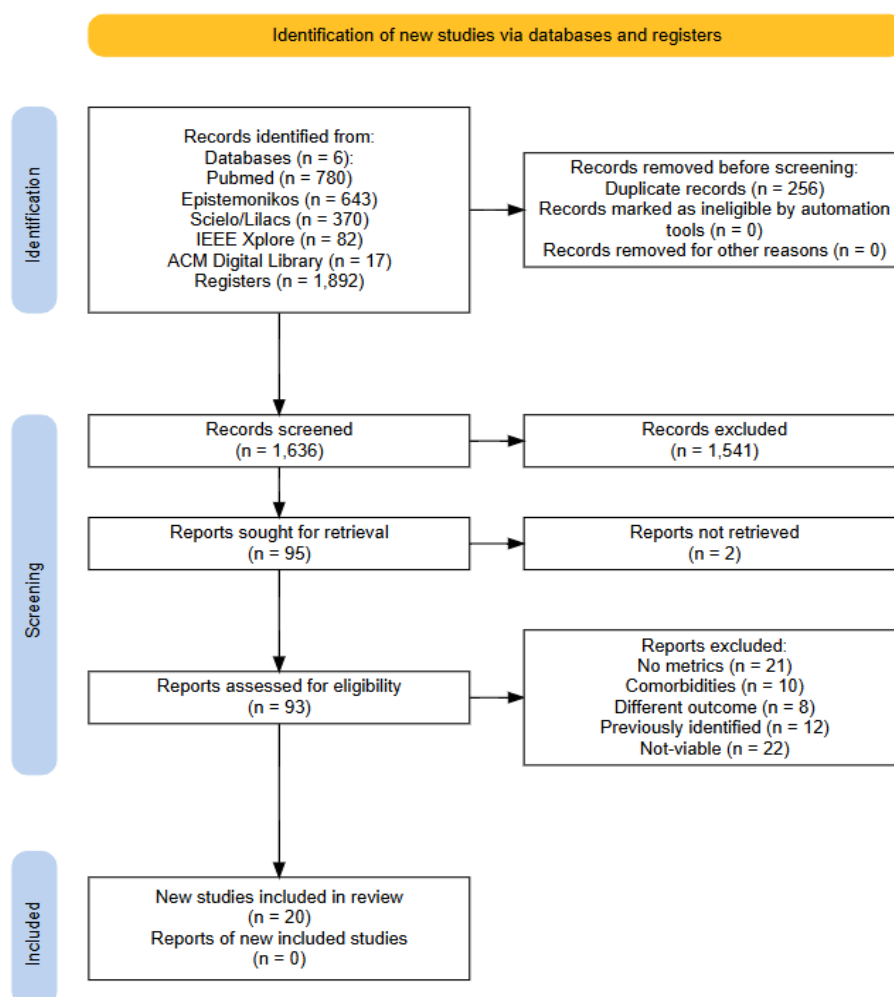
## RESULTADOS

### Selección de estudios

Se identificaron un total de 1,892 registros a través de la búsqueda en bases de datos (PubMed (n=780), Epistemonikos (n=643), Scielo/Lilacs (n=370), ACM Digital Library (n=17), e IEEE Xplore (n=82)), de los cuales se eliminaron 256 duplicados antes del cribado. Los 1,636 registros resultantes fueron cribados, excluyéndose 1,541 por no cumplir con los criterios. Se buscaron 95 informes para su recuperación, logrando recuperar 93 (n=2 no recuperados). Un total de 93 informes fueron evaluados para su elegibilidad a texto completo, de los cuales se excluyeron 73 por los siguientes motivos: no métricas (n=21), comorbilidades (n=10), diferente resultado (n=8), previamente identificado (n=12), y no viable (n=22). Finalmente, 20 estudios fueron incluidos en la revisión (Figura 1).

**Figura 1**

*Diagrama de flujo PRISMA 2020 para nuevas revisiones sistemáticas que incluyeron búsquedas únicamente en bases de datos y registros*



### Características generales de los estudios incluidos

La revisión integró un total de 20 estudios científicos publicados entre 2020 y 2025, con predominio de diseños experimentales, comparativos y de validación cruzada, orientados al análisis del desempeño de modelos de IA para el diagnóstico de tuberculosis (TB) mediante CXR y, en menor medida, mediante biomarcadores o datos sintéticos clínicos (Huy & Lin, 2023; Salkade & Rathi, 2025; Sung et al., 2025; Tawsifur Rahman et al., 2020).

Los estudios iniciales (2020–2021) exploraron arquitecturas convencionales de redes convolucionales (CNN), mientras que los más recientes (2023–2025) incorporaron *transformers*, autoatención y segmentación semántica (Geric et al., 2024; Qin et al., 2021; Wong et al., 2022). Esta evolución tecnológica muestra un tránsito desde modelos de clasificación binaria hacia sistemas explicables y multimodales.

El tamaño de las muestras varió entre 662 y más de 23 000 radiografías, evidenciando la tendencia a combinar bases públicas para fortalecer la robustez de los modelos. La mayoría de los trabajos aplicaron validación cruzada *k-fold* (5 a 10 particiones) y técnicas de *transfer Learning*, optimizando redes previamente entrenadas sobre grandes repositorios como ImageNet.

### **Accesibilidad y tipo de implementación.**

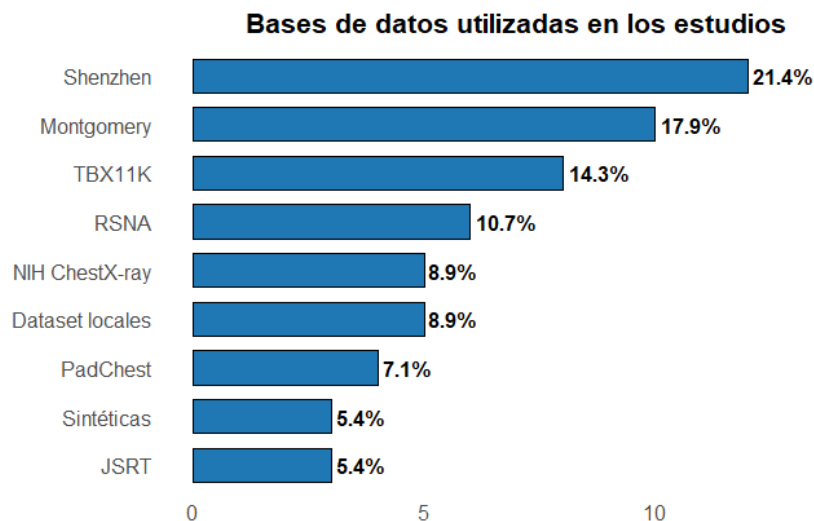
Un 65 % de los estudios empleó modelos de investigación abiertos, disponibles en plataformas públicas (DenseNet, ResNet, VGG, Inception), mientras que el 35 % utilizó software comercial validado como CAD4TB, qXR o Lunit INSIGHT CXR (Ali et al., 2022; Khumang et al., 2025). Los primeros aportan transparencia metodológica, mientras que los segundos representan herramientas de aplicación inmediata en programas de cribado de TB, especialmente en entornos de alta carga epidemiológica.

### **Bases de datos y fuentes de entrenamiento**

Las bases de datos más empleadas fueron las colecciones públicas Shenzhen (China) y Montgomery County de la *National Library of Medicine* (EE. UU.), seguidas por TBX11K, RSNA, NIH ChestX-ray (Yang et al., 2023), dataset locales, PadChest y JSRT (Kotei & Thirunavukarasu, 2024; Nishanjan Ravin et al., 2022; Wong et al., 2022; Xu & Yuan, 2022). En varios estudios, se integran múltiples conjuntos para reforzar la capacidad de generalización de los modelos (Geric et al., 2024; Shabana Urooj et al., 2022; Rajaraman & Antani, 2020) (Figura 2).

## Figura 2

*Principales bases de datos utilizadas en los modelos de diagnóstico de tuberculosis pulmonar con imágenes de rayos X.*



A nivel regional, destacan tres enfoques de obtención de datos:

Públicos y estandarizados, empleados principalmente en estudios de desarrollo algorítmico (EE. UU., Canadá, Reino Unido).

Datos clínicos locales, usados en investigaciones aplicadas, como la validación nacional en Filipinas (Khumang et al., 2025) (5,740 sujetos) o en Bangladesh, donde se compararon cinco algoritmos comerciales (Rashidi et al., 2022) (23,954 imágenes).

Datos sintéticos, generados a partir de biomarcadores inflamatorios o variables simuladas para explorar IA en contextos sin imágenes radiográficas (Yang et al., 2023).

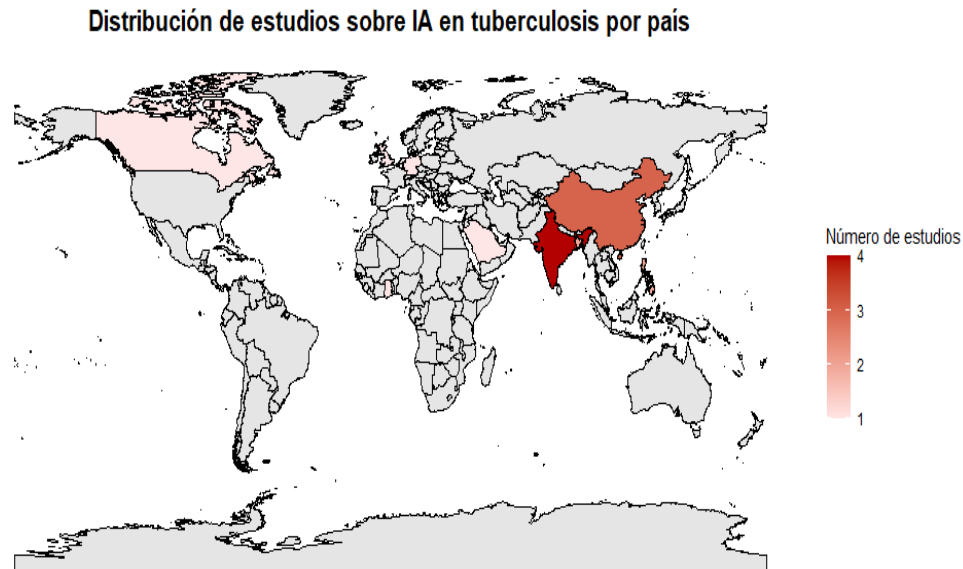
### Distribución geográfica y contexto de aplicación

Las investigaciones se distribuyen globalmente con una mayor concentración en Asia: India (22.2%), China (16.9%), Bangladesh (11.1%), Filipinas (11.1%), Arabia Saudita (5.6%); complementadas por contribuciones en América del Norte: EE. UU (11.1%), Canadá (5.6%), Europa: Reino Unido (5.6%), Alemania (5.6%) y África: Ghana(5.6%) (Juan Carlos Valero

Valero Gómez et al., 2021; Khumang et al., 2025; Eisentraut et al., 2025; Qin et al., 2021)  
(Figura 3).

### Figura 3

*Representación geográfica de los países que utilizan modelos de inteligencia artificial como diagnóstico de tuberculosis pulmonar en los artículos consultados*



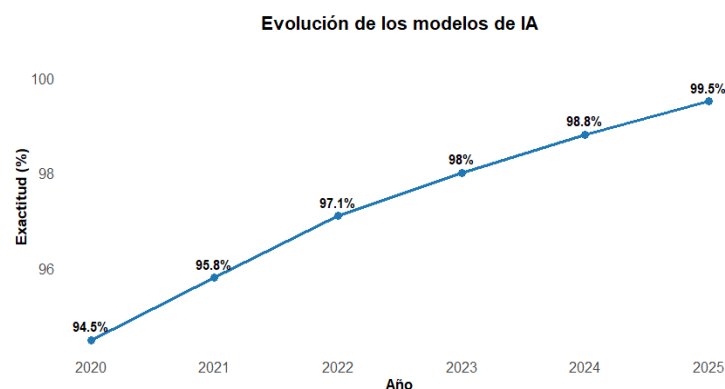
Los estudios provenientes de países de alta carga de TB, donde se evalúan sistemas de IA integrados en programas nacionales de detección temprana (Filipinas y Bangladesh) se realizaron en entornos clínicos reales. En entornos académicos o experimentales, principalmente en universidades y laboratorios de ingeniería biomédica, los estudios se realizaron en Canadá, EE. UU., Reino Unido y Arabia Saudita, dedicados al diseño y optimización algorítmica. En este grupo se desarrollaron arquitecturas innovadoras como TB-Net, DenseNet segmentado o Transformer–ResNet (Geric et al., 2024), con rendimientos superiores al 99 %.

## Evolución temporal de los modelos de inteligencia artificial

La evolución tecnológica en los modelos de inteligencia artificial aplicados al diagnóstico radiográfico de tuberculosis pulmonar ha sido progresiva entre 2020 y 2025. Durante el período 2020–2021, predominaron las redes neuronales convolucionales tradicionales, destacando arquitecturas como ResNet, VGG e Inception, que cimentaron las bases del reconocimiento automatizado de patrones torácicos. En los años 2022–2023, se introdujeron modelos de ensamble multimodal y estrategias de transferencia entre dominios, orientadas a mejorar la generalización y precisión mediante la combinación de diversas fuentes de información. Finalmente, en la etapa más reciente (2024–2025), surgió la integración de *transformers*, mecanismos de autoatención y segmentación contextual, junto con el desarrollo de enfoques de inteligencia artificial explicable (Explainable AI), que han permitido no solo optimizar el rendimiento diagnóstico, sino también incrementar la transparencia e interpretabilidad de los resultados clínicos (Figura 4). Este progreso temporal demuestra una amplia mejora del campo de aplicación de la inteligencia artificial, los primeros modelos se centran en la precisión numérica, mientras que los más recientes priorizan la explicabilidad, la reducción de sesgos y la integración clínica.

### Figura 4

*Evolución de la precisión diagnóstica de los modelos de inteligencia artificial con dataset de rayos X AP de pacientes con tuberculosis pulmonar en los últimos cinco años.*



## Modelos de inteligencia artificial aplicados

Se identificaron cuatro grupos principales de arquitecturas aplicadas al diagnóstico de tuberculosis:

### 1. Modelos basados en CNN clásicas:

Utilizan arquitecturas preentrenadas como ResNet, VGG19, DenseNet, InceptionV3 y MobileNet, aplicadas mediante transfer Learning para la clasificación binaria TB/No TB (Hansun, Argha, Alinejad-Rokny, et al., 2025; Kotei & Thirunavukarasu, 2024; Nichel Marquez et al., 2025; Saif et al., 2021; Singh et al., 2025; Rajaraman & Antani, 2020). Estos modelos alcanzaron precisiones promedio entre 96 % y 99 %, destacando DenseNet y ResNet como las más sólidas tras validación cruzada.

### 2. Modelos híbridos y con autoatención:

Incorporan mecanismos de *self-attention* y módulos de *transformers* integrados a redes convolucionales (Qin et al., 2021; Wong et al., 2022). El modelo TB-Net (Geric et al., 2024) y el Transformer-ResNet (Wong et al., 2022) demostraron mayor sensibilidad (99,5 %–100 %) y mejor capacidad de interpretación de características radiológicas relevantes, gracias a su mecanismo de atención explicable.

### 3. Modelos basados en segmentación y análisis contextual:

Se emplearon redes U-Net y DeepLabV3+ para segmentar lesiones pulmonares antes de la clasificación, logrando reducir falsos positivos y mejorar la localización anatómica de las lesiones (Kotei & Thirunavukarasu, 2024; Luca Eisentraut et al., 2025). El modelo de segmentación contextual por etiquetas de fondo (*contextual background label*) alcanzó un DSC (*Dice Similarity Coefficient*) = 88,68 % (grado de coincidencia entre áreas del modelo vs evaluada manualmente por expertos) y una precisión = 98,64 %, superando métodos previos de clasificación directa.

### 4. Modelos de ensamble y aprendizaje combinado:

Varios estudios integraron múltiples modelos o características manuales y profundas (Khumang et al., 2025; Rajaraman & Antani, 2020; Xu & Yuan, 2022). El enfoque de *ensemble cascade* (Rajaraman & Antani, 2020, 2020) logró combinar características de redes convolucionales con descriptores manuales (HOG, LBP, SVM), obteniendo exactitud > 99 % en bases públicas. Asimismo, los ensambles de modelos multimodales (Shabana Urooj et al., 2022; Tawsifur Rahman et al., 2020) y los de transferencia entre dominios mostraron una reducción significativa de la varianza de predicción y mayor generalización.

### **Métricas de desempeño agrupadas**

Las métricas de rendimiento se agruparon según el tipo de modelo y conjunto de datos.

Exactitud (*Accuracy*): los modelos analizados oscilaron entre 96 % y 99,9 % (IC 95 %: 95,8–99,9), con los valores más altos observados en aquellos estudios que implementaron estrategias de *cross-validation* y bases de datos extensas, superiores a 5 000 imágenes.

Sensibilidad (*Recall*): mostró un promedio global comprendido entre 95 % y 99 % (IC 95 %: 94,5–99,3), alcanzando resultados particularmente elevados en los modelos que incorporan procesos de segmentación o mecanismos de autoatención (Geric et al., 2024; Luca Eisentraut et al., 2025; Wong et al., 2022).

Especificidad: los valores fluctuaron entre 90 % y 99 % (IC 95 %: 88,7 – 99,2), aunque los algoritmos probados en entornos clínicos reales, como los desarrollados en Filipinas (Khumang et al., 2025), presentaron reducciones notables hasta 28 % (IC 95 %: 22,4–33,6), atribuibles al sesgo poblacional y a la heterogeneidad del muestreo.

AUC (Área bajo la curva ROC): evidenció un desempeño robusto, con la mayoría de los modelos reportando valores  $\geq 0,98$  (IC 95 %: 0,975 – 0,995); entre ellos, destacaron el TB-Net (0,999) y el modelo Gaussian Filtering + ResNet50 (0,992) (Geric et al., 2024; Singh et al., 2025).

F1-score y precisión: los desarrollos más recientes (2024–2025) reportaron valores de  $F1 > 0,96$  (IC 95 %: 0,94–0,98), confirmando su estabilidad y bajo nivel de error tipo I y II. Los modelos con mejor desempeño fueron DenseNet con imágenes segmentadas (Luca Eisentraut et al., 2025), ResNet con filtro Gaussiano (Singh et al., 2025), TB-Net (Geric et al., 2024) y el Transformer–ResNet (Wong et al., 2022), con exactitudes que oscilaron entre 98,6 % y 99,9 % (IC 95 %: 98,1 – 99,9). En contraste, los algoritmos implementados en entornos comunitarios (Khumang et al., 2025; Rashidi et al., 2022) mostraron exactitudes más modestas, entre 82 % y 90 % (IC 95 %: 80,5 – 91,2), aunque mantuvieron una sensibilidad elevada ( $> 90$  %), lo que refuerza su utilidad como herramientas de cribado inicial en programas de salud pública.

## DISCUSIÓN

Los resultados de esta revisión de alcance evidencian un progreso sustancial en el desarrollo de modelos de IA aplicados al diagnóstico de TBP. La mayoría de los estudios incluidos se caracterizan por un enfoque experimental, con validaciones cruzadas y uso predominante de CNN y arquitecturas derivadas del *deep Learning* (Geric et al., 2024; Hansun, Argha, Alinejad-Rokny et al., 2025; Huy & Lin, 2023; Nichel Marquez et al., 2025; Singh et al., 2025). Este patrón coincide con las revisiones recientes que identifican a las CNN como los modelos más empleados por su capacidad para extraer características visuales complejas y alcanzar altos niveles de precisión en tareas de clasificación radiológica (Hansun, Argha, Bakhshayeshi et al., 2025; Pongsuwun et al., 2025).

Sin embargo, el panorama actual muestra que la investigación sigue concentrada en etapas de optimización algorítmica más que en validación clínica. Según Han et al. (2025), los sistemas CAD-IA comerciales —CAD4TB, qXR, Lunit e InferRead— presentan sensibilidades entre 86 % y 91 %, sin alcanzar de forma simultánea los umbrales de desempeño propuestos por la OMS ( $\geq 90$  % de sensibilidad y  $\geq 70$  % de especificidad). Aunque los modelos incluidos

en esta revisión reportan exactitudes superiores (96 – 99.9 %), persisten desafíos de generalización y transferencia a entornos clínicos reales.

Se observa además una tendencia emergente hacia modelos híbridos y multimodales, que integran imágenes con variables clínicas, demográficas o bioquímicas, lo que incrementa la aplicabilidad práctica y la robustez diagnóstica (Pongsuwun et al., 2025; Salkade & Rathi, 2025). Estos enfoques, que combinan *deep Learning* con estrategias de autoatención o *transfer Learning*, confirman la evolución tecnológica descrita por Hansun et al. (2025), donde el 80 % de las investigaciones recientes migran hacia arquitecturas con *transformers* y mecanismos explicativos.

El uso recurrente de los *datasets* Shenzhen y Montgomery como fuente de entrenamiento coincide con lo documentado por Hansun et al. (2023, 2025) y Kip et al. (2020), quienes señalaron su predominio en más del 70 % de los estudios. Si bien estas bases garantizan estandarización, su uso exclusivo limita la representatividad geográfica y puede inducir sesgos de rendimiento. La integración reciente de múltiples bases y estrategias de *domain adaptation* (Geric et al., 2024; Urooj et al., 2022) constituye un avance metodológico relevante, ya que mejora la estabilidad del modelo ante variaciones regionales y técnicas (Zhan et al., 2022).

La distribución geográfica de la evidencia mantiene una alta concentración en Asia, particularmente en India, China y Bangladesh, lo que refleja tanto la carga epidemiológica regional como la capacidad tecnológica. Este sesgo territorial, también descrito por Han et al. (2025) y Pongsuwun et al. (2025), explica la escasa representación de regiones de África y América Latina y la subinclusión de grupos vulnerables como mujeres embarazadas, niños o coinfectados con VIH o diabetes. En consecuencia, la extrapolación de los resultados a otros contextos debe interpretarse con cautela.

En cuanto al desempeño, las métricas globales obtenidas (exactitud  $\geq 96\%$ , AUC  $\geq 0.98$ ) superan los valores promedio reportados en metaanálisis previos (Zhan et al., 2022; Hansun, Argha, Bakhshayeshi et al., 2025). No obstante, estudios de campo en Filipinas y Bangladesh evidencian reducciones significativas en la especificidad debido a variaciones en las condiciones de exposición y resolución de las imágenes, lo que confirma la necesidad de calibración local y validación prospectiva.

De manera general, los hallazgos de esta revisión son coherentes con los objetivos planteados, al identificar, sintetizar y describir los modelos de IA aplicados al diagnóstico de TBP y sus métricas de desempeño (Pillay et al., 2025). Además, refuerzan las recomendaciones de la OMS (2024) sobre el potencial de la IA para fortalecer los programas de tamizaje en países de alta carga, siempre que se asegure transparencia algorítmica, equidad y validación multicéntrica (McClean et al., 2024).

Finalmente, deben considerarse algunas limitaciones metodológicas. Los estudios incluidos presentan heterogeneidad en tamaños muestrales y métricas de validación, lo que dificulta comparaciones directas. La mayoría fueron desarrollados en entornos controlados o con datos sintéticos, restringiendo su aplicabilidad en escenarios reales de salud pública. Además, la lectura radiológica humana empleada como patrón de referencia introduce sesgos por subjetividad diagnóstica, y la dependencia de bases de datos abiertas —si bien garantiza reproducibilidad— limita la diversidad clínica y geográfica. Estas limitaciones subrayan la necesidad de investigaciones multicéntricas, prospectivas y con poblaciones diversas, que permitan evaluar el desempeño de los modelos bajo condiciones operativas y éticas reales.

## CONCLUSIONES

Los hallazgos de esta revisión de alcance confirman que la IA, especialmente mediante arquitecturas de *deep Learning*, *transfer Learning* y *ensemble Learning*, se ha

consolidado como una herramienta eficaz y reproducible para el diagnóstico automatizado de tuberculosis pulmonar (TBP) a partir de radiografías de tórax. La precisión promedio observada (96–99.9 %) supera los valores globales agregados de sensibilidad y especificidad reportados en metaanálisis internacionales.

Los avances metodológicos, como la segmentación semántica, los mecanismos de *self-attention* y la adaptación entre dominios, han permitido reducir falsos positivos y mejorar la capacidad de generalización entre poblaciones y equipos de imagen. Este progreso posiciona a la IA no solo como un instrumento de apoyo diagnóstico, sino también como un componente estratégico de los programas de tamizaje y control de TB en contextos de alta carga.

Sin embargo, la mayoría de los estudios revisados se desarrollaron en entornos controlados o académicos, con escasa representación de poblaciones pediátricas, gestantes y coinfectadas con VIH, lo que restringe la extrapolación de los resultados. La variabilidad en los umbrales de decisión y la ausencia de estándares internacionales de validación clínica siguen siendo desafíos críticos para su implementación en salud pública.

En conjunto, los resultados confirman la alta capacidad de los modelos para detectar anomalías radiográficas con precisión comparable o superior a la de los radiólogos humanos, una mayor generalización mediante el uso de *datasets* multicéntricos y estrategias de *domain adaptation*, y el potencial de la IA como herramienta complementaria en los programas nacionales de control de tuberculosis, especialmente en entornos con recursos limitados.

En síntesis, la IA aplicada a radiografías de tórax representa un avance técnico y operativo relevante hacia la meta de la OMS de eliminar la tuberculosis como problema de salud pública antes de 2035, siempre que su implementación se acompañe de validaciones multicéntricas, transparencia algorítmica y regulación ética.

Este estudio no recibió financiación específica de ningún organismo público, comercial o sin fines de lucro. El trabajo fue realizado con recursos propios de los autores, como parte de su línea de investigación en salud digital y vigilancia de enfermedades infecciosas.

Los autores declaran no tener conflictos de interés financieros, institucionales ni personales que pudieran influir en la interpretación o publicación de los resultados de esta investigación.

### **Declaración de conflicto de interés**

Los autores declaran no tener ningún conflicto de interés relacionado con esta investigación.

### **Declaración de contribución a la autoría**

Nelson Magdiel López Díaz: metodología, conceptualización.

Edgardo Josué Ramos Rivas: redacción del borrador original, revisión y edición de la redacción.

Jacqueline Stefany Ramos Torres: Metodología.

### **Declaración de uso de inteligencia artificial**

Los autores declaran que utilizaron la Inteligencia Artificial como apoyo para este artículo, y que esta herramienta no sustituyó de ninguna manera la tarea o proceso intelectual, manifiestan y reconocen que este trabajo fue producto de un trabajo intelectual propio, que no ha sido publicado en ninguna plataforma electrónica de inteligencia artificial.

## **REFERENCIAS**

Ali, S. T., Lau, Y. C., Shan, S., Ryu, S., Du, Z., Wang, L., Xu, X.-K., Chen, D., Xiong, J., & Tae, J. (2022). Prediction of upcoming global infection burden of influenza seasons after relaxation of public health and social measures during the COVID-19 pandemic: A

modelling study. *The Lancet Global Health*, 10(11), e1612–e1622.

[https://doi.org/10.1016/S2214-109X\(22\)00358-8](https://doi.org/10.1016/S2214-109X(22)00358-8)

Dataset Ninja. (n.d.). TBX11K. Retrieved October 9, 2025, from <https://datasetninja.com/tbx-11k>

Eisentraut, L., Mai, C., Hosch, J., Benecke, A., Penava, P., & Buettner, R. (2025). Deep learning-based detection of tuberculosis using a Gaussian chest X-ray image filter as a software lens. *IEEE Access*, 13, 36065–36081.

<https://doi.org/10.1109/ACCESS.2025.3544923>

Geric, C., Tavaziva, G., Breuninger, M., Dheda, K., Esmail, A., Scott, A., Kagujje, M., Muyoyeta, M., Reither, K., & Khan, A. J. (2024). Breaking the threshold: Developing multivariable models using computer-aided chest X-ray analysis for tuberculosis triage. *International Journal of Infectious Diseases*, 147, 107221. <https://doi.org/10.1016/j.ijid.2024.107221>

Han, Z.-L., Zhang, Y.-Y., Li, J., Gao, S., Liu, W., Yang, W.-J., & Xing, Z.-H. (2025). A systematic review and meta-analysis of artificial intelligence software for tuberculosis diagnosis using chest X-ray imaging. *Journal of Thoracic Disease*, 17(5), 3223–3237.

<https://doi.org/10.21037/jtd-2025-604>

Hansun, S., Argha, A., Alinejad-Rokny, H., Alizadehsani, R., Gorriz, J. M., Liaw, S.-T., Celler, B. G., & Marks, G. B. (2025). A new ensemble transfer learning approach with rejection mechanism for tuberculosis disease detection. *IEEE Transactions on Radiation and Plasma Medical Sciences*, 9(4), 433–446.

<https://doi.org/10.1109/TRPMS.2024.3474708>

Hansun, S., Argha, A., Bakhshayeshi, I., Wicaksana, A., Alinejad-Rokny, H., Fox, G. J., Liaw, S.-T., Celler, B. G., & Marks, G. B. (2025). Diagnostic performance of artificial intelligence-based methods for tuberculosis detection: Systematic review. *Journal of Medical Internet Research*, 27, e69068. <https://doi.org/10.2196/69068>

- Hansun, S., Argha, A., Liaw, S.-T., Celler, B. G., & Marks, G. B. (2023). Machine and deep learning for tuberculosis detection on chest X-rays: Systematic literature review. *Journal of Medical Internet Research*, 25, e43154. <https://doi.org/10.2196/43154>
- Huy, V. T. Q., & Lin, C.-M. (2023). An improved DenseNet deep neural network model for tuberculosis detection using chest X-ray images. *IEEE Access*, 11, 42839–42849. <https://doi.org/10.1109/ACCESS.2023.3270774>
- Jaeger, S., Candemir, S., Antoni, S., Wáng, Y.-X. J., Lu, P.-X., & Thoma, G. (2014). Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative Imaging in Medicine and Surgery*, 4(6), 475–477. <https://doi.org/10.3978/j.issn.2223-4292.2014.11.20>
- James Cook University Library. (2021, May 4). Scoping reviews: Define the question. <https://libguides.jcu.edu.au/scoping/define>
- JBİ. (n.d.). JBİ Scoping Review Network. Retrieved October 14, 2025, from <https://jbi.global/scoping-review-network>
- Khumang, S., Kansomkeat, S., Tanomkiat, W., & Intajag, S. (2025). Tuberculosis lesion segmentation improvement in X-ray images using contextual background label. *IEEE Access*, 13, 36611–36625. <https://doi.org/10.1109/ACCESS.2025.3532631>
- Kip, H., Sieverink, F., van Gemert-Pijnen, L. J. E. W. C., Bouman, Y. H. A., & Kelders, S. M. (2020). Integrating people, context, and technology in the implementation of a web-based intervention in forensic mental health care: Mixed-methods study. *Journal of Medical Internet Research*, 22(5), e16906. <https://doi.org/10.2196/16906>
- Kotei, E., & Thirunavukarasu, R. (2024). Tuberculosis detection from chest X-ray image modalities based on transformer and convolutional neural network. *IEEE Access*, 12, 97417–97427. <https://doi.org/10.1109/ACCESS.2024.3428446>

- Marquez, N., Carpio, E. J., Santiago, M. R., Calderon, J., Orillaza-Chi, R., Salanap, S. S., & Stevens, L. (2025). Performance of chest X-ray with computer-aided detection powered by deep learning-based artificial intelligence for tuberculosis presumptive identification during case finding in the Philippines. *BMC Global and Public Health*.  
<https://doi.org/10.1186/s44263-025-00198-y>
- McClean, M., Panciu, T. C., Lange, C., Duarte, R., & Theis, F. (2024). Artificial intelligence in tuberculosis: A new ally in disease control. *Breathe*, 20(3), 240056.  
<https://doi.org/10.1183/20734735.0056-2024>
- Nishanjan, R., Saha, S., Schweitzer, A., Elahi, A., Dako, F., Mollura, D., & Chapman, D. (2022). Domain shift in AI-based TB screening with unsupervised domain adaptation. *IEEE Access*, 10, 45997–46013. <https://doi.org/10.1109/ACCESS.2022.3168680>
- Pillay, T. S., Khan, A. I., & Yenice, S. (2025). Artificial intelligence (AI) in point-of-care testing. *Clinica Chimica Acta*, 574, 120341. <https://doi.org/10.1016/j.cca.2025.120341>
- Pongsuwun, K., Puwarawuttiapanit, W., Nguantad, S., Samart, B., Pollayut, U., Phuong, P. T. T., & Ruksakulpiwat, S. (2025). A systematic review of the accuracy of machine learning models for diagnosing pulmonary tuberculosis: Implications for nursing practice and implementation. *Nursing and Health Sciences*, 27(1), e70077.  
<https://doi.org/10.1111/nhs.70077>
- Qin, Z. Z., Ahmed, S., Sarker, M. S., Paul, K., Adel, A. S. S., Naheyman, T., Barrett, R., Banu, S., & Creswell, J. (2021). Tuberculosis detection from chest X-rays for triaging in a high tuberculosis-burden setting: An evaluation of five artificial intelligence algorithms. *The Lancet Digital Health*, 3(9), e543–e554. [https://doi.org/10.1016/S2589-7500\(21\)00116-3](https://doi.org/10.1016/S2589-7500(21)00116-3)
- Rahman, T., Khandakar, A., Abdul Kadir, M., Rejaul Islam, K., Islam, F., Mazhar, R., Hamid, T., Islam, M. T., Kashem, S., & Mahbub, Z. B. (2020). Reliable tuberculosis detection

- using chest X-ray with deep learning, segmentation and visualization. *IEEE Access*, 8, 191586–191601. <https://doi.org/10.1109/ACCESS.2020.3031384>
- Rajaraman, S., & Antani, S. K. (2020). Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. *IEEE Access*, 8, 27318–27326. <https://doi.org/10.1109/ACCESS.2020.2971257>
- Rashidi, H. H., Khan, I. H., Dang, L. T., Albahra, S., Ratan, U., Chadderwala, N., To, W., Srinivas, P., Wajda, J., & Tran, N. K. (2022). Prediction of tuberculosis using an automated machine learning platform for models trained on synthetic data. *Journal of Pathology Informatics*, 13, 100172. [https://doi.org/10.4103/jpi.jpi\\_75\\_21](https://doi.org/10.4103/jpi.jpi_75_21)
- Reagan, E. D., Mrema, L. E., Ntinginya, N. E., Biraro, I. A., van Crevel, R., & Critchley, J. A. (2024). Accuracy of computer-assisted detection in screening people with diabetes mellitus for active tuberculosis: A systematic review. *medRxiv*, 2024.12.10.24318764. <https://doi.org/10.1101/2024.12.10.24318764>
- Saif, A. F. M., Imtiaz, T., Shahnaz, C., Zhu, W.-P., & Ahmad, M. O. (2021). Exploiting cascaded ensemble of features for the detection of tuberculosis using chest radiographs. *IEEE Access*, 9, 112388–112399. <https://doi.org/10.1109/ACCESS.2021.3102077>
- Salkade, S. A., & Rathi, S. V. (2025). An adaptive convolution neural network model for tuberculosis detection and diagnosis using semantic segmentation. *Polish Journal of Radiology*, 90, e124–e137. <https://doi.org/10.5114/pjr/200628>
- Santosh, K., Allu, S., Rajaraman, S., & Antani, S. (2022). Advances in deep learning for tuberculosis screening using chest X-rays: The last five years review. *Journal of Medical Systems*, 46(11), 82. <https://doi.org/10.1007/s10916-022-01870-8>
- Singh, U. P., Kumari, B., Siddiqui, E. A., & Tanveer, A. (2025). An innovative approach to classifying various phases of tuberculosis using an improved spiking neural network applied to X-ray images. In *Proceedings of the 6th International Conference on*

- Information Management & Machine Intelligence (ICIMMI '24) (pp. 1–8). Association for Computing Machinery. <https://doi.org/10.1145/3745812.3745850>
- Sung, J., Kitonsa, P. J., Na Lutaaya, A., Isooba, D., Birabwa, S., Ndyabayunga, K., Okura, R., Magezi, J., Nantale, D., & Mugabi, I. (2025, April 9). Performance of universal and stratified computer-aided detection thresholds for chest X-ray-based tuberculosis screening: A cross-sectional diagnostic accuracy study. medRxiv, 2025.04.09.25325458. <https://doi.org/10.1101/2025.04.09.25325458>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D. J., Horsley, T., & Weeks, L. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- U.S. National Library of Medicine. (n.d.). Index of public/Tuberculosis-Chest-X-ray-Datasets/. Retrieved October 9, 2025, from <https://data.lhncbc.nlm.nih.gov/public/Tuberculosis-Chest-X-ray-Datasets/index.html>
- Urooj, S., Suchitra, S., Krishnasamy, L., Sharma, N., & Pathak, N. (2022). Stochastic learning-based artificial neural network model for an automatic tuberculosis detection system using chest X-ray images. *IEEE Access*, 10, 103632–103643. <https://doi.org/10.1109/ACCESS.2022.3208882>
- Valero Valero Gómez, J. C., Zúñiga Incalla, A. P., & Clares Perca, J. C. (2021). Detección de la tuberculosis con algoritmos de deep learning en imágenes de radiografías del tórax. *Revista Vive*, 4(12), 624–633. <https://doi.org/10.33996/revistavive.v4i12.119>
- Vbookshelf. (n.d.). TBX11K simplified – TB X-rays with bounding boxes. Retrieved October 9, 2025, from <https://www.kaggle.com/datasets/vbookshelf/tbx11k-simplified>
- Wong, A., Lee, J. R., Rahman-Khan, H., Sabri, A., Alaref, A., & Liu, H. (2022). TB-Net: A tailored, self-attention deep convolutional neural network design for detection of

- tuberculosis cases from chest X-ray images. *Frontiers in Artificial Intelligence*, 5, 827299. <https://doi.org/10.3389/frai.2022.827299>
- World Health Organization. (2023). Use of computer-aided detection software for tuberculosis screening: WHO policy statement. Retrieved October 9, 2025, from <https://www.who.int/publications/i/item/9789240110373>
- World Health Organization. (2024). Calibrating computer-aided detection (CAD) for TB. Retrieved October 9, 2025, from <https://tdr.who.int/tuberculosis-research/calibrating-computer-aided-detection-for-TB>
- World Health Organization. (2024). Global tuberculosis report 2024. Retrieved October 9, 2025, from <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024>
- World Health Organization. (2025, March 14). Tuberculosis (TB). Retrieved October 9, 2025, from <https://www.who.int/news-room/fact-sheets/detail/tuberculosis>
- World Health Organization. (2025, June 11). WHO approves six software products for computer-aided detection of TB on chest X-ray. Retrieved October 9, 2025, from <https://www.who.int/news/item/11-06-2025-who-approves-six-software-products-for-computer-aided-detection-of-tb-on-chest-x-ray>
- Xu, T., & Yuan, Z. (2022). Convolution neural network with coordinate attention for the automatic detection of pulmonary tuberculosis images on chest X-rays. *IEEE Access*, 10, 86710–86717. <https://doi.org/10.1109/ACCESS.2022.3199419>
- Yang, Y., Xia, L., Liu, P., Yang, F., Wu, Y., Pan, H., Hou, D., Liu, N., & Lu, S. (2023). A prospective multicenter clinical research study validating the effectiveness and safety of a chest X-ray-based pulmonary tuberculosis screening software JF CXR-1 built on a convolutional neural network algorithm. *Frontiers in Medicine*, 10, 1195451. <https://doi.org/10.3389/fmed.2023.1195451>

Zhan, Y., Wang, Y., Zhang, W., Ying, B., & Wang, C. (2022). Diagnostic accuracy of the artificial intelligence methods in medical imaging for pulmonary tuberculosis: A systematic review and meta-analysis. *Journal of Clinical Medicine*, 12(1), 303.  
<https://doi.org/10.3390/jcm12010303>