

Propuesta de software para la predicción del fracaso terapéutico en tuberculosis pulmonar

Software proposal for predicting therapeutic failure in pulmonary tuberculosis

Martín del Río-Sánchez^{1*} <https://orcid.org/0009-0003-5500-4914>

Luis Eugenio Valdés García² <https://orcid.org/0009-0002-6862-8713>

Enrique Beldarrain Chaple³ <https://orcid.org/0009-0003-5500-4914>

Omarys Loyola-Cabrera¹ <https://orcid.org/0000-0002-9228-7816>

¹ Centro Provincial de Higiene, Epidemiología y Microbiología. Ciego de Ávila, Cuba.

² Instituto Finlay de Vacunas. La Habana, Cuba.

³ Centro Nacional de Información de Ciencias Médicas. La Habana, Cuba.

* Autor para la correspondencia: delriosanchezmartin@gmail.com

RESUMEN

Introducción: el fracaso terapéutico en la tuberculosis pulmonar representa un desenlace adverso crítico que compromete la salud del paciente y favorece la transmisión sostenida y el desarrollo de resistencias. La aplicación manual de

modelos predictivos no es factible en la práctica clínica diaria, por lo que se requiere de herramientas automatizadas.

Objetivo: evaluar la percepción de un panel de expertos sobre la necesidad, utilidad y viabilidad de implementar un software predictivo del riesgo de fracaso terapéutico en tuberculosis pulmonar.

Métodos: se realizó un estudio observacional, descriptivo y transversal durante marzo de 2026 en Ciego de Ávila, Cuba. Se aplicó una encuesta autoadministrada validada a 20 especialistas y se empleó el método Delphi para alcanzar consenso sobre los factores de riesgo y las funcionalidades del software.

Resultados: el 100 % de los expertos estuvo de acuerdo o totalmente de acuerdo con la necesidad y utilidad del software. Las principales barreras identificadas fueron: complejidad del cálculo (90 %), falta de tiempo en la consulta (85 %) y alto riesgo de errores (80 %). Se priorizaron 15 factores de riesgo y se diseñó una propuesta con seis módulos funcionales, que alcanzó un consenso de pertinencia superior al 85 %.

Conclusiones: los expertos reconocen unánimemente la necesidad de un software predictivo para tuberculosis pulmonar. La propuesta estructurada es viable y pertinente, aunque se requieren estudios de validación clínica prospectiva antes de su implementación generalizada en el Sistema Nacional de Salud cubano.

Palabras clave: Algoritmos; Predicción; Programas Informáticos; Técnica Delphi; Tuberculosis; Valor Predictivo de las Pruebas

ABSTRACT

Introduction: treatment failure in pulmonary tuberculosis represents a critical adverse outcome that compromises patient health and promotes sustained transmission and the development of resistance. The manual application of predictive models is not feasible in daily clinical practice, thus requiring automated tools.

Objective: to evaluate the perceptions of a panel of experts regarding the need for, usefulness of, and feasibility of implementing predictive software to assess the risk of treatment failure in pulmonary tuberculosis.

Methods: An observational, descriptive, and cross-sectional study was conducted in March 2026 in Ciego de Ávila, Cuba. A validated self-administered survey was administered to 20 specialists, and the Delphi method was used to reach a consensus on risk factors and software functionalities.

Results: 100 % of the experts agreed or strongly agreed with the need for and usefulness of the software. The main barriers identified were: complexity of the calculation (90 %), lack of time during consultations (85 %), and high risk of errors (80 %). Fifteen risk factors were prioritized, and a proposal with six functional modules was designed, achieving a consensus of relevance exceeding 85 %.

Conclusions: the experts unanimously recognize the need for predictive software for pulmonary tuberculosis. The structured proposal is viable and relevant, although prospective clinical validation studies are required before its widespread implementation in the Cuban National Health System.

Keywords: Algorithms; Prediction; Software; Delphi Technique; Tuberculosis; Predictive Value of Tests

Recibido: 11/08/2026

Aceptado: 03/09/2026

Introducción

La tuberculosis (TB) es una enfermedad transmisible y curable, pero con un fuerte componente social, que constituye tanto causa como consecuencia de la pobreza y afecta a personas de cualquier edad, sexo o raza. Causada por el bacilo de Koch, en 2021 la Organización Mundial de la Salud (OMS) reportó 10,6 millones de enfermos. Se estima que una cuarta parte de la población mundial está infectada, y entre el 5 y el 10 % desarrollará la enfermedad, que afecta principalmente a los pulmones (80-85 %). En América Latina y el Caribe, según la Organización Panamericana de la Salud (OPS), el 89 % de los casos se concentran en 13 países, con más de la mitad en Brasil, Perú y México.^(1,2,3)

La tuberculosis pulmonar (TBP) sigue siendo un desafío prioritario para la salud pública.⁽⁴⁾ A pesar de los esfuerzos, el fracaso terapéutico, definido como la persistencia de baciloscopías positivas tras el cuarto mes de tratamiento o la recaída bacteriológica representa un desenlace adverso crítico. Este evento no solo compromete la salud del paciente, sino que favorece la transmisión sostenida y el desarrollo de resistencias, obstaculizando el objetivo de eliminación de la enfermedad.⁽⁵⁾

Desde el punto de vista de la regulación sanitaria, los softwares utilizados con fines médicos han adquirido un protagonismo creciente en los últimos años, lo que ha motivado el desarrollo de marcos normativos específicos para garantizar su seguridad, eficacia y calidad. En este contexto, el Centro para el

Control Estatal de Medicamentos, Equipos y Dispositivos Médicos (CECMED), mediante la Resolución No. 45 de 2025, aprobó la regulación E 130-25, que establece los requisitos para la inscripción de fabricantes y el registro sanitario del software de uso médico.⁽⁶⁾ Aunque la herramienta predictiva propuesta carece de propósito diagnóstico estricto, su función de apoyo clínico mediante el cálculo automatizado del riesgo de fracaso terapéutico la sitúa entre los softwares médicos que deben cumplir requisitos regulatorios para su implementación en el Sistema Nacional de Salud.

El personal sanitario necesita herramientas integradas de forma fluida en su labor diaria con pacientes con tuberculosis; un software que calcule al instante la probabilidad de adherencia terapéutica tendría gran valor, pues reduciría los errores por subestimación del riesgo. Por ello, se propone un software predictivo que no solo predice el fracaso, sino que también estratifica el riesgo en tiempo real para apoyar la toma de decisiones clínicas. El desarrollo de un software predictivo para el fracaso terapéutico en tuberculosis pulmonar no es un componente opcional, sino el eslabón indispensable para trasladar la evidencia epidemiológica y el consenso de expertos a la práctica clínica cotidiana; en este contexto, la presente investigación tuvo como objetivo evaluar la percepción de un panel de expertos sobre la necesidad, utilidad y viabilidad de implementar dicho software.

Métodos

Se realizó un estudio observacional, descriptivo y de corte transversal, de carácter exploratorio, durante el mes de marzo de 2026 en la provincia de Ciego de Ávila, Cuba. Los criterios de selección de los expertos fueron: experiencia mínima de cinco años en tuberculosis pulmonar; autoría o coautoría de al

menos dos publicaciones científicas sobre la temática; y posesión de categoría docente o grado científico (Máster, Especialista de Segundo Grado o Doctor en Ciencias).

Se invitó inicialmente a 28 especialistas de la provincia, de los cuales 20 cumplieron los criterios, aceptaron participar y otorgaron su consentimiento informado, conformando así la población de estudio (tasa de respuesta del 71,4 %). Se diseñó una encuesta autoadministrada en formato digital, estructurada en tres secciones: la primera recogía datos generales del especialista; la segunda exploraba la percepción del modelo predictivo y su aplicabilidad; y la tercera indagaba sobre la necesidad y las expectativas respecto a una herramienta informática. Todos los ítems de las secciones segunda y tercera se evaluaron mediante una escala Likert de 5 puntos (1: Desacuerdo total; 5: Acuerdo total).

El instrumento fue sometido a un proceso de validación de contenido por un panel de siete expertos independientes —metodólogos, neumólogos y epidemiólogos— que no participaron en el estudio Delphi posterior. Cada juez calificó los ítems en términos de claridad, coherencia y relevancia en una escala de 1 (no cumple) a 4 (cumple en su totalidad). Se calculó el Índice de Validez de Contenido por ítem (IVC-i) como la proporción de jueces que asignaron una puntuación ≥ 3 ; se retuvieron los ítems con $\text{IVC-i} \geq 0.86$, y el IVC total del cuestionario fue de 0.94.

Adicionalmente, se realizó una prueba piloto con 15 médicos con características similares a las de la población de estudio, quienes no formaron parte de dicha población. En esta fase se evaluaron la comprensibilidad (mediante preguntas abiertas), el tiempo de llenado (promedio de 8 minutos) y la consistencia interna de las escalas sumativas mediante el coeficiente alfa de

Cronbach, que alcanzó un valor de 0,82. A raíz de estos resultados se ajustó la redacción de dos ítems para mejorar su claridad.

La versión definitiva de la encuesta se administró de forma remota mediante un formulario digital autoadministrado, enviado por correo electrónico y por WhatsApp a cada uno de los expertos. El acceso al formulario se limitó exclusivamente a los participantes del estudio, con el fin de preservar el anonimato de las respuestas. Se aplicó el método de Delphi modificado en tres rondas para alcanzar un consenso sobre la elaboración de la propuesta de software.

Posteriormente, se aplicó el método Delphi modificado en tres rondas con los 20 expertos que conformaron la población de estudio. En la primera ronda, los expertos evaluaron 23 factores de riesgo y propusieron funcionalidades mediante preguntas abiertas. Las respuestas se sintetizaron y, en la segunda ronda, cada experto recibió retroalimentación personalizada (su respuesta previa frente al promedio del grupo) para reevaluar los 15 factores con mediana ≥ 4 y asignarles ponderaciones. En la tercera ronda se validó la propuesta del software. El anonimato se preservó en todo momento. El consenso se definió como un acuerdo de $>80\%$; se alcanzó un 85% para la propuesta final. La variabilidad inter-ronda se redujo en un 42% , lo que evidencia convergencia.

Para representar visualmente los conceptos más relevantes identificados durante el consenso, se elaboró una nube de palabras a partir de las expresiones textuales aportadas por los expertos en las rondas cualitativas del método Delphi. El proceso consistió en la extracción de términos significativos, la exclusión de palabras vacías y la ponderación por frecuencia de aparición. La representación gráfica se generó con la herramienta en línea [WordArt.com](https://wordart.com). Las palabras de mayor relevancia temática fueron destacadas con mayor tamaño y un color diferenciado, con fines ilustrativos y de síntesis visual.

El motor de cálculo del software se basó en un modelo predictivo de fracaso terapéutico desarrollado previamente mediante regresión logística multivariable, elaborado según la guía TRIPOD. Dicho modelo se derivó de una cohorte retrospectiva de pacientes con tuberculosis pulmonar en la provincia de Ciego de Ávila ($n = 420$), obtuvo un área bajo la curva ROC de 0,81 (IC95 %: 0,76 – 0,85) y fue sometido a validación externa en una cohorte de Camagüey mediante la recalibración del intercepto conforme a la prevalencia local.

A partir de las expectativas y necesidades expresadas por los expertos se definieron las funcionalidades mínimas de la herramienta: automatización del cálculo del riesgo, estratificación en categorías (bajo, medio y alto), generación de recomendaciones clínicas personalizadas, elaboración de reportes imprimibles, monitoreo anónimo de casos y configuración de umbrales personalizados. La propuesta quedó estructurada en seis módulos funcionales (entrada de datos, cálculo del riesgo, recomendaciones, reportes, monitoreo y configuración), atendiendo además a consideraciones de diseño centrado en el usuario, privacidad y ética, validación en un entorno real y capacitación. La propuesta final, evaluada de manera integral, fue sometida a revisión y validación por el mismo panel de expertos, alcanzándose un consenso superior al 85 % en cuanto a su pertinencia y viabilidad globales; dicho porcentaje corresponde al acuerdo sobre la propuesta en su conjunto y no a una valoración independiente de cada módulo funcional.

Dado que el software propuesto maneja información sensible de salud, se definieron requisitos preliminares de seguridad en consonancia con la Resolución No. 45/2025 del CECMED (Regulación E 130-25) y con la normativa cubana de protección de datos personales;⁽⁶⁾ estos requisitos incluyen: (a) almacenamiento local de los datos en los dispositivos clínicos, sin transmisión automática hacia servidores externos, excepto para actualizaciones anonimizadas del modelo; (b) cifrado AES-256 de la base de datos local y de las

comunicaciones con repositorios centrales; (c) control de acceso mediante autenticación por usuario y contraseña, con roles diferenciados (clínico, administrador); (d) registro de auditoría no modificable que consigne quién, cuándo y qué datos introdujo o consultó; (e) anonimización irreversible de los casos exportados para monitoreo o reentrenamiento del modelo, con eliminación de cualquier identificador directo (nombre, carné de identidad, dirección) y con seudonimización mediante clave separada; y (f) función de eliminación segura de datos al finalizar el tratamiento o cuando el paciente lo solicite. Por su parte, el análisis de los datos se realizó con el software SPSS versión 27.0, mediante un análisis descriptivo de las variables, con el cálculo de frecuencias absolutas y relativas.

El protocolo de investigación fue aprobado por el Comité de Ética de la Investigación del Centro Provincial de Higiene, Epidemiología y Microbiología de Ciego de Ávila. En todo momento se garantizó la confidencialidad de la información, conforme a la Declaración de Helsinki y las normas nacionales vigentes.

Resultados

La Tabla 1 presenta la caracterización de los expertos participantes. Predominó la profesión médica, con 18 de los 20 integrantes, mientras que solo 2 eran licenciados en Tecnologías de la Salud. En cuanto a la especialidad, Neumología aportó 6 expertos, seguida de Medicina Interna (5), Medicina Familiar (4) y Epidemiología (3). Respecto a la categoría docente, 18 panelistas poseían alguna categoría: 7 eran profesores auxiliares, 6 titulares y 5 asistentes; únicamente 2 carecían de ella. En cuanto al grado científico, más de la mitad tenía título de Máster (11) o de Especialista de Segundo Grado (11), y 7 habían

alcanzado el grado de Doctor en Ciencias. En cuanto al cargo, 12 combinaban funciones asistenciales y docentes, 4 se dedicaban exclusivamente a la asistencia y 4 ocupaban cargos directivos. Esta composición refleja un panel heterogéneo en cuanto a especialidades, niveles de formación académica y perfiles laborales, lo que aporta una perspectiva amplia y complementaria sobre el problema estudiado.

Tabla 1- Características generales de los expertos participantes

Categoría		No.	%
Profesión	Médico Especialista	18	90
	Lic. en Tecnologías de la Salud	2	10
Especialidad	Neumología	6	33,3
	Medicina Interna	5	27,8
	Medicina Familiar	4	22,2
	Epidemiología	3	16,7
Categoría docente	Profesor Titular	6	30
	Profesor Auxiliar	7	35
	Profesor Asistente	5	25
	Sin categoría docente	2	10
Grado científico	Doctor en Ciencias	7	35
	Máster	11	55
	Especialista de 2do grado	11	55
Cargo actual	Asistencial únicamente	4	20
	Asistencial y docente	12	60
	Directivo	4	20

La figura 1 muestra la percepción de los expertos sobre la necesidad, la utilidad y la aplicabilidad de un software predictivo sobre la tuberculosis, se observa que 5 del total de los expertos está de acuerdo y 15 de ellos totalmente de acuerdo. Ninguno de los participantes manifestó estar neutral, en desacuerdo o totalmente en desacuerdo.

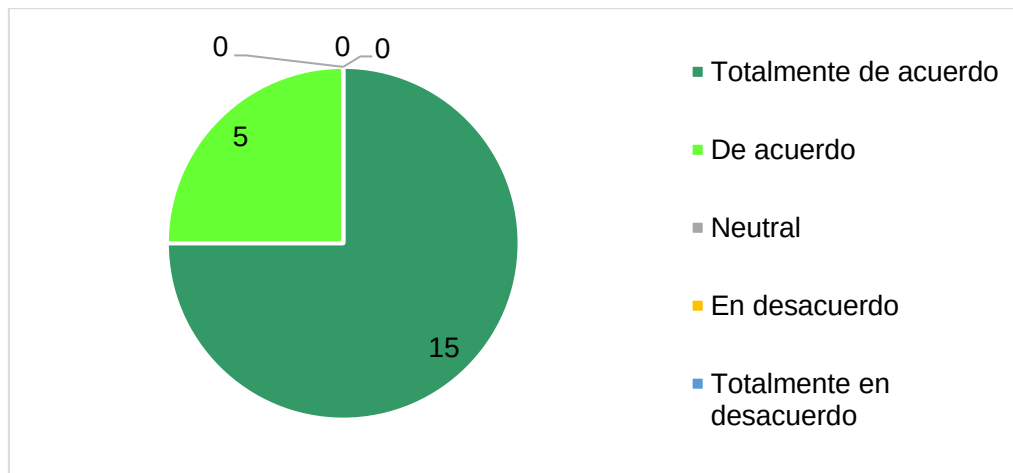


Fig. 1- Percepción de los expertos sobre la necesidad, la utilidad y aplicabilidad de un software predictivo sobre la tuberculosis pulmonar.

Fuente: Elaboración propia.

La figura 2 muestra las barreras identificadas por los expertos en la implementación del software. Del total de los expertos, 18 identificaron como barrera la complejidad del cálculo, 17 seleccionaron la falta de tiempo en la consulta, 16 el alto riesgo de errores, 15 la baja estandarización entre clínicos, 14 la dificultad para recordar todos los factores y 8 la resistencia al cambio.

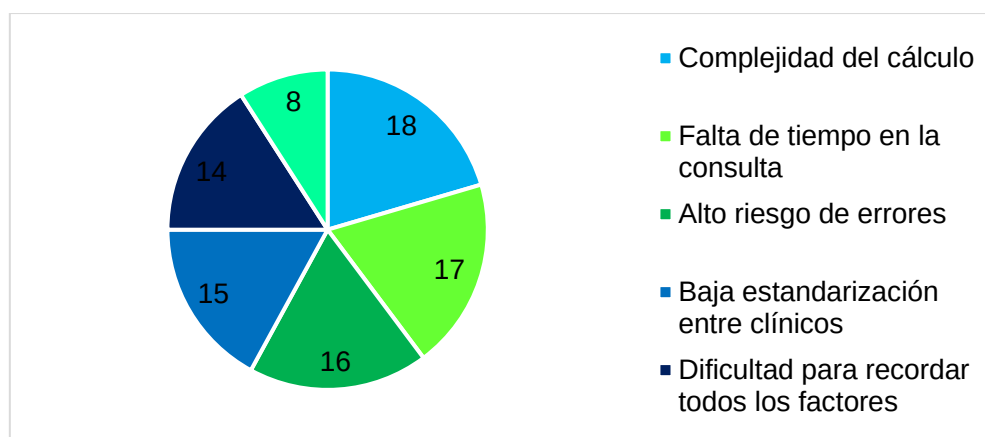


Fig. 2- Barreras para la implementación del software.

Fuente: Elaboración propia.

El análisis cualitativo de las percepciones de los expertos, derivado del consenso alcanzado mediante el método Delphi modificado, reveló un acuerdo unánime en torno a la utilidad potencial del modelo predictivo para identificar pacientes con alto riesgo de fracaso terapéutico. No obstante, los expertos coincidieron en señalar que la aplicación manual del modelo no es factible en la práctica clínica diaria, principalmente debido a la complejidad que representa integrar las 15 variables priorizadas, cada una con sus respectivas ponderaciones e interacciones.

Esta percepción se vio reforzada por el reconocimiento generalizado de que la falta de una herramienta automatizada constituye una barrera crítica para la implementación generalizada del modelo. Los especialistas manifestaron, de forma consistente, que un software específico no solo automatizaría el cálculo del riesgo, sino que también estandarizaría los criterios de evaluación entre diferentes observadores y reduciría significativamente los errores en la estimación del pronóstico.

Además, según los expertos, la herramienta no debe limitarse a calcular una probabilidad. Debe estratificar el riesgo en tiempo real y generar recomendaciones personalizadas de manejo. Los especialistas también aceptaron ampliamente una versión offline y solicitaron capacitación institucional. Esto evidencia su preocupación por garantizar la equidad en el acceso a la herramienta. Asimismo, buscan su efectiva adopción en todos los entornos asistenciales, incluidos aquellos con limitaciones de conectividad (Ver Figura 3).



Fig. 3- Nube de palabras.

A partir del consenso alcanzado y de las necesidades identificadas en la encuesta de percepción, se elaboró una propuesta de software predictivo que integra las funcionalidades mínimas consideradas imprescindibles para su aplicabilidad clínica.

Tabla 2- Propuesta de funcionalidades mínimas del software

Módulo	Funcionalidad
Entrada de datos	Formulario con los 15 factores priorizados (listas desplegadas, escalas, valores numéricos). Ayudas contextuales para cada factor (definición de “bajo IMC”, “cavitaciones”, etc.).
Cálculo del riesgo	Motor que aplica el modelo final (regresión logística, árboles o conjunto) y devuelve: (a) probabilidad numérica (0-100 %), (b) categoría de riesgo (bajo/medio/alto), (c) interpretación clínica en lenguaje sencillo.
Recomendaciones	Acciones sugeridas según el nivel de riesgo y los factores modificables detectados (ej. “Alto riesgo por adherencia deficiente y bajo apoyo social: activar consejería y DOT familiar”).
Reportes	Generación de informe imprimible para la historia clínica del paciente, con fecha, factores introducidos y predicción.
Monitoreo	Registro anónimo de casos (factores + desenlace real) para futuras actualizaciones del modelo.
Configuración	Opción de uso de umbrales personalizados (si el centro prefiere mayor sensibilidad o especificidad).

Seguridad	Almacenamiento local con cifrado AES-256, control de acceso por usuario y rol, registro de auditoría no modificable, anonimización irreversible de exportaciones y eliminación segura de datos al final del tratamiento o por solicitud del paciente.
------------------	---

Consideraciones para el desarrollo

- **Diseño centrado en el usuario:** involucrar a clínicos, neumólogos y epidemiólogos desde la fase de prototipado para asegurar interfaces intuitivas y una carga de trabajo mínima. Debe considerarse la diversidad de perfiles profesionales y de niveles de alfabetización digital, evitando sesgos de diseño que favorezcan a usuarios con mayor experiencia tecnológica.
- **Privacidad, ética y autonomía:** los datos deben anonimizarse y cumplir con las normativas locales vigentes. El paciente debe ser informado sobre el uso de esta herramienta predictiva y debe otorgar su consentimiento. Asimismo, debe garantizarse el principio de autonomía profesional. Deben observarse, además, los principios de beneficencia, no maleficencia y justicia, asegurando que la herramienta no introduzca discriminación por edad, sexo, procedencia o condición social.
- **Accesibilidad y equidad:** el software debe funcionar sin conexión para garantizar su uso en entornos con conectividad limitada o nula. Debe ser de acceso gratuito para todas las instituciones del Sistema Nacional de Salud, independientemente de su nivel de atención o de su ubicación geográfica. Se recomienda su disponibilidad en múltiples plataformas y la inclusión de ajustes de accesibilidad que faciliten su uso para profesionales con diferentes capacidades visuales o cognitivas. Asimismo, debe evitarse cualquier sesgo algorítmico derivado de datos históricos que puedan subrepresentar a poblaciones vulnerables.
- **Validación en el entorno real:** antes de la implementación generalizada, el software debe ser evaluado en un estudio piloto multicéntrico que incluya, al menos, dos o tres centros de salud con características

diferentes. Este piloto debe evaluar no solo la usabilidad, la aceptación y el rendimiento diagnóstico, sino también la presencia de sesgos de implementación, la equidad en el acceso y la percepción de utilidad por parte de usuarios con distintos niveles de experiencia. Los resultados deben retroalimentar el rediseño de la herramienta antes de su escalamiento.

- Capacitación integral: desarrollar tutoriales breves, videos y material de ayuda integrado en el propio software. Además, se recomienda complementar estos recursos con talleres presenciales y semipresenciales dirigidos a los profesionales que utilizarán la herramienta en su práctica diaria.
- Evaluación de sesgos y monitoreo continuo: durante el desarrollo y la implementación, debe establecerse un plan de monitoreo que permita detectar posibles sesgos en la predicción, tanto a nivel del modelo como en su aplicación clínica. Este plan debe incluir auditorías periódicas del rendimiento del software en distintos subgrupos poblacionales y la revisión sistemática de los desenlaces reales frente a las predicciones generadas.

El propio modelo predictivo deberá reentrenarse periódicamente con nuevos datos epidemiológicos locales para mantener su validez y precisión a lo largo del tiempo. Para ello, el software debe facilitar la actualización centralizada, de modo que las mejoras o ajustes se implementen desde un repositorio común sin necesidad de reinstalar la aplicación en cada ordenador.

Asimismo, resulta imprescindible incorporar un sistema de monitorización del rendimiento que permita registrar los desenlaces reales (éxito o fracaso terapéutico) y compararlos sistemáticamente con las predicciones generadas, lo que permite detectar oportunamente la denominada "deriva conceptual", es decir, aquella situación en la que el modelo pierde precisión debido a cambios

en las características de la población atendida o en los esquemas de tratamiento empleados. Finalmente, la herramienta debe contemplar mecanismos de retroalimentación de los usuarios, mediante los cuales los clínicos puedan reportar problemas, dificultades técnicas o sugerencias de mejora directamente desde la interfaz del software, asegurando así un proceso continuo de optimización y adaptación a las necesidades reales del entorno asistencial.

Discusión

La evidencia internacional señala que la complejidad y el volumen de variables constituyen barreras sustanciales para la adopción de modelos predictivos en la práctica clínica habitual.⁽⁷⁾ Además, la identificación de la complejidad del cálculo manual y la falta de tiempo como principales barreras resulta particularmente relevante. Estos factores son críticos en entornos de alta carga asistencial, donde cualquier herramienta que no se integre de forma fluida en el flujo de trabajo tiene pocas probabilidades de ser adoptada. Los expertos demandan un software que no solo automatice el cálculo, sino que también proporcione recomendaciones de manejo, lo que apunta a la necesidad de un sistema de apoyo a la decisión clínica, y no de una simple calculadora.

Es preciso señalar que, si bien el software predictivo propuesto en esta investigación no constituye propiamente un dispositivo médico en el sentido tradicional, pues no tiene como finalidad el diagnóstico de una enfermedad de forma directa, su desarrollo y futura implementación en el contexto sanitario cubano deben considerar el marco regulatorio establecido por la autoridad competente. La definición del Foro Internacional de Reguladores de Dispositivos Médicos (IMDRF), establece que el "software de uso médico"

incluye tanto el software como dispositivo médico (SaMD) como el software embebido en un dispositivo médico (SiMD).⁽⁶⁾

El SaMD se define como aquel software destinado a ser utilizado para uno o más fines médicos, sin ser parte de un dispositivo de hardware, y que puede proporcionar información para el diagnóstico, la detección, el seguimiento, el pronóstico, la predicción o la determinación del estado fisiológico.⁽⁶⁾ Aunque la herramienta aquí propuesta tiene un propósito predictivo y de apoyo a la toma de decisiones clínicas, se ha tenido en cuenta esta definición regulatoria durante su diseño, especialmente en lo relativo a los requisitos de verificación, validación clínica, gestión de riesgos y ciberseguridad exigidos por la citada resolución.

Los hallazgos del presente estudio, coinciden con experiencias internacionales recientes en el campo de la predicción clínica mediante modelos automatizados. En este sentido, Lai y cols. desarrollaron y validaron una herramienta simplificada de evaluación de fragilidad utilizando aprendizaje automático, logrando un rendimiento superior (AUC 0,963 en entrenamiento y 0,850 en validación externa) a partir de solo ocho parámetros clínicos fácilmente disponibles, lo que refuerza la viabilidad de herramientas parsimoniosas pero precisas para la práctica clínica habitual.⁽⁸⁾

Asimismo, Coppola y cols. demostraron que modelos como Support Vector Machine (AUC 0,97) y XGBoost (AUC 0,94) predicen con alta exactitud la necesidad de oxigenoterapia en niños con neumonía por *Mycoplasma pneumoniae* utilizando datos de ingreso de rutina, destacando la importancia de incorporar herramientas interpretables como SHAP para facilitar la adopción clínica.⁽⁹⁾

Por otra parte, Alzghaibi identificó que las barreras técnicas y de usabilidad, como la compatibilidad del sistema y la lentitud en el rendimiento, constituyen los obstáculos más significativos para la adopción de aplicaciones de salud

móvil por parte de los profesionales, hallazgo que guarda estrecha relación con las barreras percibidas por los expertos en el presente estudio respecto a la complejidad del cálculo manual y la falta de tiempo.⁽¹⁰⁾

Olawade y cols., en su revisión narrativa sobre modelos predictivos en servicios de urgencias, subrayan que la integración de modelos de aprendizaje automático en los flujos de trabajo clínicos enfrenta desafíos persistentes relacionados con la calidad de los datos, la interpretabilidad y la capacidad de generalización, lo que refuerza la necesidad de un desarrollo centrado en el usuario y de estudios multicéntricos de validación externa, tal como se propone en este trabajo.⁽¹¹⁾ En conjunto, estas experiencias respaldan la estrategia de desarrollar herramientas informáticas predictivas que, basadas en modelos rigurosos y validadas con la participación de los usuarios finales, puedan cerrar la brecha entre la evidencia epidemiológica y la toma de decisiones en escenarios de recursos limitados.

Olawade y cols., en una revisión reciente sobre aplicaciones de inteligencia artificial en medicina de emergencia, subrayan que el éxito de un software clínico no depende únicamente de su capacidad predictiva, sino también de su precisión operativa, entendida como la exactitud y la consistencia de los resultados generados en condiciones reales de uso. Los autores destacan que las herramientas con alto rendimiento en fases de desarrollo pueden mostrar una disminución de su precisión al implementarse en entornos clínicos heterogéneos, debido a variaciones en la calidad de los datos, la interoperabilidad con los sistemas existentes y la carga cognitiva del personal sanitario. En este sentido, enfatizan la necesidad de validar el software mediante estudios multicéntricos que evalúen no solo la exactitud de las predicciones, sino también la usabilidad, la reproducibilidad y la estabilidad del sistema a lo largo del tiempo.⁽¹²⁾

Por otra parte, la revisión sistemática de Zhang y cols. identificó que las redes neuronales convolucionales y las máquinas de vectores de soporte destacan en la predicción de la duración del tratamiento, mientras que Random Forest y las redes neuronales artificiales son prometedoras para anticipar reacciones adversas y desenlaces terapéuticos, lo que respalda la inclusión de múltiples módulos funcionales en el software propuesto.⁽¹³⁾

En consonancia con lo anterior, Subbaraman y cols. señalaron que las tecnologías digitales de adherencia (DATs), como los recordatorios por SMS, los pastilleros digitales y la observación directa por video, pueden facilitar la estratificación del riesgo y la provisión de cuidados diferenciados, siempre que se integren con estrategias clínicas que aborden las causas subyacentes de la falta de adherencia. No obstante, estos autores advierten que la evidencia sobre la efectividad clínica de estas tecnologías aún es limitada y que su implementación requiere considerar aspectos éticos como la privacidad, la confidencialidad y la autonomía del paciente. Lo anterior refuerza la necesidad de incorporar capas de seguridad en el diseño del software desde sus fases iniciales.⁽¹⁴⁾

Desde el punto de vista del diagnóstico y el monitoreo, el manual de algoritmos para el diagnóstico de tuberculosis de ORAS-CONHU subraya la importancia de contar con flujos de trabajo estandarizados y pruebas rápidas aprobadas por la OMS (como Xpert MTB/RIF) para la detección temprana de casos y de resistencias.⁽¹⁵⁾ La integración de estos principios diagnósticos en un software predictivo que, además, se alinee con el marco regulatorio cubano constituye un valor añadido de la presente propuesta. En conjunto, estos antecedentes respaldan la estrategia de desarrollar una herramienta informática que no solo automatice el cálculo del riesgo de fracaso terapéutico, sino que también genere recomendaciones clínicas personalizadas, permitiendo así trasladar la

evidencia epidemiológica a la práctica clínica cotidiana en escenarios de recursos limitados.

En opinión del equipo investigador, la herramienta propuesta no debe concebirse como un sustituto del juicio clínico, sino como un sistema de apoyo a la decisión que optimice el flujo de trabajo del personal sanitario y reduzca la carga cognitiva asociada al manejo de múltiples factores de riesgo y sus interacciones. Se enfatiza que el éxito de la implementación dependerá, en gran medida, de un diseño centrado en el usuario que involucre a los clínicos desde las fases tempranas del desarrollo, así como de una capacitación institucional sostenida que garantice la alfabetización digital de los profesionales.

Asimismo, los autores opinan que el software debería integrarse progresivamente en la historia clínica electrónica y que su uso debería promoverse mediante guías clínicas oficiales, sin carácter obligatorio, respetando siempre la autonomía del médico tratante. Además, es importante que la herramienta sea gratuita y funcione sin conexión, dadas las limitaciones de conectividad en el país.

El presente estudio presenta varias limitaciones que deben considerarse al interpretar sus resultados. El estudio se basó en la percepción de los expertos y no en la validación clínica prospectiva del software propuesto, por lo que no se dispone de datos objetivos sobre su efectividad, usabilidad real o impacto en los desenlaces terapéuticos. Por otra parte, el método Delphi, si bien útil para alcanzar consensos, está sujeto a sesgos inherentes, como el sesgo de deseabilidad social o la influencia de las opiniones dominantes. Por último, no se realizó un análisis de costo-efectividad de la implementación del software, un aspecto crucial para evaluar su viabilidad en un contexto de recursos limitados como el de Cuba.

A partir de los resultados y las limitaciones del presente estudio, se proponen varias líneas de investigación futuras. En primer lugar, se hace impostergable la

realización de un estudio piloto multicéntrico en al menos tres provincias cubanas para evaluar la usabilidad, la aceptabilidad y el rendimiento diagnóstico del software en condiciones reales de práctica clínica, con un diseño cuasiexperimental que compare los desenlaces antes y después de su implementación.

Conclusiones

Los expertos consultados reconocen unánimemente la necesidad y la utilidad de un software predictivo del fracaso terapéutico en la tuberculosis pulmonar, identificando la complejidad del cálculo manual y la falta de tiempo como las principales barreras para su implementación clínica. Mediante el método Delphi se priorizaron 15 factores de riesgo y se diseñó una propuesta de software con seis módulos funcionales, que alcanzó un consenso de pertinencia superior al 85 %. La herramienta propuesta es factible y pertinente para estandarizar la estratificación del riesgo y apoyar la toma de decisiones clínicas, aunque se requiere su validación prospectiva y estudios multicéntricos previos a su implementación generalizada en el Sistema Nacional de Salud cubano, de acuerdo con el marco regulatorio vigente del CECMED.

Referencias bibliográficas

1. Organización Mundial de la Salud (OMS). Tuberculosis. Washington DC: OMS [Internet]. 2022 [citado 20/01/2026]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/tuberculosis>
2. Organización Panamericana de la Salud (OPS). La Tuberculosis en las Américas. Informe Regional 2021. Washington DC: OPS/OMS [Internet]. 2022 [citado 20/01/2026]. Disponible en:

https://iris.paho.org/bitstream/handle/10665.2/57084/9789275326497_spa.pdf?sequence=11&isAllowed=y

3. Secretaría de Salud 129. México avanza en la reducción de incidencia y mortalidad por tuberculosis. México: Gobierno de México [Internet]. 2021 [citado 20/01/2026]. Disponible en: <https://www.gob.mx/salud/prensa/129-mexico-avanza-en-la-reduccion-de-incidencia-y-mortalidad-por-tuberculosis>
4. Zhao CY, Song C, Lu YB, Huang AC, Zeng CM, Liu RH, Li WW, Xie ZH, Zhu QD. A New Era in Tuberculosis Prevention and Treatment: Breakthroughs in Drug Development and Future Prospects. Infect Dis Ther. [Online]. 2025;14(10):2205-2219. DOI: <http://doi.org/10.1007/s40121-025-01221-3>
5. Oña Cerna HY, Guanica Sarmiento NB, Miranda Velazquez E. Factores asociados al fracaso del tratamiento antifímico de pacientes con tuberculosis pulmonar. Rev Ciencia Latina [Internet]. 2025; 9 (1). DOI: https://doi.org/10.37811/cl_rcm.v9i1.16073
6. Centro para el Control Estatal de Medicamentos, Equipos y Dispositivos Médicos (CECMED). Resolución No. 45/2025. Regulación E 130-25 Requisitos para la Inscripción de Fabricantes y el Registro Sanitario del Software de Uso Médico. Ámbito Regulador [Internet]. 2025 [citado 20/01/2026]. Disponible en: <https://www.cecmec.med.cu/reglamentacion/ambito-regulador/indice?page=34%2C1>
7. Cowley LE, Farewell DM, Maguire S, Kemp AM. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. Diagnostic and Prognostic Research [Online]. 2019;3:16. DOI: <https://doi.org/10.1186/s41512-019-0060-y>
8. Lai J, Cheng C, Liang T, Tang L, Guo X, Liu X. Development and multi-cohort validation of a machine learning-based simplified frailty assessment tool for

clinical risk prediction. J Transl Med. [Online]. 2025 [cited 20/01/2026];23:921. Available from: <https://link.springer.com/article/10.1186/s12967-025-06728-4>

9. Coppola C, Andrew JJ, Ruggieri M, La Spina M, La Bianca MR, Leonardi S, *et al.* Machine learning prediction of oxygen therapy in pediatric Mycoplasma pneumoniae pneumonias. Front Digit Health. [Online]. 2026 [cited 20/01/2026]; 8:1755878. Available from: <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2026.1755878/full>

10. Alzghaibi H. Healthcare practitioners' perceptions of mHealth application barriers: challenges to adoption and strategies for enhancing digital health integration. Healthcare. [Online]. 2025 [cited 20/01/2026];13(5):494. Available from: <https://www.mdpi.com/resolver?pii=healthcare13050494>

11. Olawade DB, Da'Costa A, Origbo JE, Osonuga A, Egbon E, Teke J, *et al.* Smart emergency care: a narrative review of predictive machine learning models. Ann Transl Med. [Online]. 2025 [cited 20/01/2026];13(5):60. Available from: <https://atm.amegroups.org/article/view/144876/html>

12. Olawade DB, Wada OZ, Ige AO, Egbewole BI, Olojo A, Oladapo BI. Artificial intelligence in emergency medicine: A review of current applications, challenges, and future directions. J Clin Med. [Internet]. 2024;13(16):4746. DOI: <https://doi.org/10.3390/jcm13164746>

13. Zhang F, Zhang F, Li L, Pang Y. Clinical utilization of artificial intelligence in predicting therapeutic efficacy in pulmonary tuberculosis. J Infect Public Health. [Online]. 2024;17(4):633-40. DOI: <https://doi.org/10.1016/j.jiph.2024.02.012>

14. Subbaraman R, de Mondesert L, Musiimenta A, Pai M, Mayer KH, Thomas BE, *et al.* Digital adherence technologies for the management of tuberculosis therapy: mapping the landscape and research priorities. BMJ Glob Health.

[Online]. 2018 [cited 20/01/2026];3(5):e001018. Available from:

<https://pmc.ncbi.nlm.nih.gov/articles/PMC6195152/pdf/bmjgh-2018-001018.pdf>

15. Organismo Andino de Salud – Convenio Hipólito Unanue (ORAS-CONHU).

Manual de algoritmos para el diagnóstico de tuberculosis. [Internet] Lima:

ORAS-CONHU; 2018. [citado 20/01/2026]. Disponible en:

<https://www.paho.org/sites/default/files/2019-cde-manual-algoritmos-dx-tb-comisca.pdf>

Declaración de conflictos de interés

Los autores declaran no tener ningún conflicto de intereses, financiero o personal, que pudiera haber influido en el desarrollo, los resultados o la interpretación del presente trabajo.

Contribución de autoría

Conceptualización, metodología, investigación, curación de datos, escritura (borrador original), visualización: Martín del Río-Sánchez.

Metodología, investigación, análisis formal, escritura – revisión y edición, supervisión: Luis Eugenio Valdés García.

Conceptualización, Curación de datos, Investigación, Metodología, Recursos: Enrique Beldarrain Chaple.

Validación, análisis formal, recursos, escritura, revisión y edición, supervisión y administración del proyecto: Omarys Loyola-Cabrera.

Todos los autores leyeron y aprobaron la versión final del manuscrito.