

Evaluación crítica de la calidad de la Guía OMS Tuberculosis 2025 mediante el instrumento AGREE II

Jaime Angamarca-Iguago[1], Juan Marcos Parise-Vasco[2], Jaen Cagua-Ordoñez[1], Claudia Reytor-González[1]
Claudia Reytor-Gonzales Daniel Simancas-Racines[2]

1. Escuela de Medicina. Pontificia Universidad Católica del Ecuador, Santo Domingo, 230203, Ecuador.
2. Facultad de Salud y Bienestar. Pontificia Universidad Católica del Ecuador, Quito, 170143, Ecuador.

DOI: <https://doi.org/10.16921/pfr.v10i2.388>

PRÁCTICA FAMILIAR RURAL | Vol.10 | No.3 | Noviembre 2025 | Recibido: 12/10/2025 | Aprobado: 25/11/2025

Cómo citar este artículo

Angamarca-Iguago J, Parise-Vasco JM, Cagua-Ordoñez J, Reytor-González C, Simancas-Racines D. Evaluación crítica de la calidad de la Guía OMS Tuberculosis 2025 mediante el instrumento AGREE II. PFR [Internet]. Noviembre 2025; 10(3). Disponible en: <https://practicafamiliarrural.org/index.php/pfr/article/view/388>

Compartir en:



Resumen

Objetivo: Evaluar la calidad metodológica de la Guía Consolidada de la OMS sobre tuberculosis (Módulo 4: Tratamiento y Cuidado, 2025) mediante el instrumento AGREE II. **Métodos:** Se realizó un estudio descriptivo de evaluación metodológica transversal. Tres evaluadores independientes (JWAI, JMPV, JCCO) con experiencia certificada en el instrumento AGREE II calificaron los 23 ítems en una escala Likert de 7 puntos. Se calcularon puntuaciones escaladas de dominio según la fórmula oficial de AGREE II, estadística descriptiva (media, desviación estándar, rango) y confiabilidad interevaluador mediante el coeficiente kappa de Fleiss. **Resultados:** La Guía de Práctica Clínica (GPC) demostró una alta calidad metodológica. Las puntuaciones medias escaladas fueron: Dominio 1: Alcance y Propósito, 100.0% (DE=0.0); Dominio 2: Implicación de Implicados, 88.89% (DE=6.43); Dominio 3: Rigor en el Desarrollo, 95.14% (DE=4.17); Dominio 4: Claridad de Presentación, 100.0% (DE=0.0); Dominio 5: Aplicabilidad, 80.56% (DE=6.29); Dominio 6: Independencia Editorial, 91.67% (DE=14.43). La confiabilidad interevaluador global fue buena (Kappa de Fleiss $\kappa=0.65$). Se identificaron discrepancias significativas en el Ítem 14 (Procedimiento de actualización, $\kappa=0.35$) y en el Ítem 22 (Influencia del financiador, $\kappa=0.18$, clasificado como "pobre"). La puntuación global de calidad fue 7/7, unánime, con una recomendación mayoritaria de uso sin modificaciones (2/3 de los evaluadores). **Conclusiones:** La GPC OMS TB 2025 representa un estándar de oro para el desarrollo de guías de práctica clínica internacionales. Su rigor metodológico (95.14%), transparencia editorial (91.67%) y claridad de presentación (100.0%) la hacen confiable para la práctica global. Las áreas de mejora identificadas se focalizan en la integración de herramientas de implementabilidad (Dominio 5: 80.56%) y la clarificación de procedimientos prospectivos de actualización. Las discrepancias observadas en Independencia Editorial (Dominio 6) revelan ambigüedad interpretativa en el instrumento AGREE II sobre cómo evaluar la independencia del financiador cuando la documentación es implícita más que explícita.

Palabras clave: guía de práctica clínica; tuberculosis; AGREE II; calidad metodológica; confiabilidad interevaluador; medicina basada en evidencia; Organización Mundial de la Salud

Critical quality assessment of the WHO Tuberculosis Guidelines 2025 using the AGREE II instrument

Abstract

Objective: To assess the methodological quality of the WHO Consolidated Guidelines on Tuberculosis (Module 4: Treatment and Care, 2025) using the AGREE II instrument. **Methods:** A descriptive cross-sectional methodological appraisal was conducted. Three independent raters (JWAI, JMPV, JCCO), certified experts in the AGREE II instrument, scored all 23 items on a 7-point Likert scale. Scaled domain scores were calculated according to the official AGREE II formula, along with descriptive statistics (mean, standard deviation, range) and inter-rater reliability using Fleiss' kappa coefficient. **Results:** The clinical practice guideline (CPG) demonstrated high methodological quality. Mean scaled scores were: Domain 1: Scope and Purpose, 100.0% (SD=0.0); Domain 2: Stakeholder Involvement, 88.89% (SD=6.43); Domain 3: Rigor of Development, 95.14% (SD=4.17); Domain 4: Clarity of Presentation, 100.0% (SD=0.0); Domain 5: Applicability, 80.56% (SD=6.29); Domain 6: Editorial Independence, 91.67% (SD=14.43). Overall inter-rater reliability was good (Fleiss' kappa $\kappa=0.65$). Significant discrepancies were observed in Item 14 (Update procedure, $\kappa=0.35$) and Item 22 (Funder influence, $\kappa=0.18$, rated as "poor"). The overall quality rating was 7/7, unanimous, with a majority recommending use without modifications (2/3 raters). **Conclusions:** The WHO TB 2025 CPG represents a gold standard for international clinical practice guideline development. Its methodological rigor (95.14%), editorial transparency (91.67%), and clarity of presentation (100.0%) make it highly reliable for global practice. Improvement areas include integrating implementation tools (Domain 5: 80.56%) and clarifying prospective update procedures. Discrepancies in Editorial Independence (Domain 6) highlight interpretive ambiguity in AGREE II regarding how to assess funder independence when documentation is implicit rather than explicit.

Keywords: clinical practice guideline; tuberculosis; AGREE II; methodological quality; inter-rater reliability; evidence-based medicine; World Health Organization

Introducción

La tuberculosis (TB) persiste como uno de los retos sanitarios más significativos a nivel mundial. Según el último informe de la Organización Mundial de la Salud (OMS), en 2024 se registraron aproximadamente 10,8 millones de casos nuevos de TB y 1,25 millones de defunciones atribuibles a la enfermedad (1). Dentro de este panorama epidemiológico complejo, la tuberculosis resistente a medicamentos (TB-MDR) y la tuberculosis extensamente resistente (TB-XDR) representan un desafío creciente y de magnitud crítica. La OMS estima que anualmente surgen aproximadamente 400.000 nuevos casos de TB resistente a isoniazida y rifampicina (1,2), entidades clínicas caracterizadas por mayor fracaso terapéutico, regímenes de tratamiento prolongados, mayor toxicidad, costos sustancialmente elevados y mortalidad incrementada.

Ante esta compleja epidemia, la estandarización del tratamiento basada en evidencia científica sólida constituye el pilar fundamental para el control y la eliminación de la TB a escala global. Las Guías de Práctica Clínica (GPC) de alta calidad metodológica son instrumentos esenciales en esta estrategia (3,4). Las recomendaciones emitidas por la OMS, particularmente, funcionan como documentos de referencia internacional que informan y alinean las políticas de los Programas Nacionales de Tuberculosis en más de 190 países (2). El propósito fundamental de estas guías es optimizar los resultados clínicos en pacientes, garantizando adherencia a regímenes terapéuticos efectivos y facilitando la adopción de nuevas intervenciones farmacológicas (5).

Sin embargo, la autoridad y credibilidad de una GPC no deben derivarse únicamente de la reputación institucional de su organismo emisor, sino fundamentalmente de la solidez y la transparencia de su proceso de desarrollo metodológico. La confianza que merece una recomendación depende críticamente de que el proceso de desarrollo sea sistemático, riguroso, transparente y libre de sesgos que pudieran comprometer la validez de sus conclusiones (6). Evaluaciones históricas de la calidad metodológica de GPC sobre TB han identificado deficiencias significativas en múltiples dominios metodológicos, cuestionando la fiabilidad y aplicabilidad práctica de estas herramientas (3).

Para cuantificar objetivamente la calidad del proceso de desarrollo y la exhaustividad del reporte de guías clínicas, se ha desarrollado, validado y adoptado internacionalmente el instrumento AGREE II (Appraisal of Guidelines for Research & Evaluation II). AGREE II constituye actualmente el estándar de oro global para la evaluación sistemática de GPC, proporcionando un marco metodológico estructurado que comprende 23 ítems organizados en 6 dominios: 1. Alcance y Propósito, 2. Implicación de Agentes Clave, 3. Rigor en el Desarrollo, 4. Claridad de la Presentación, 5. Aplicabilidad, y 6. Independencia Editorial (6). La aplicación rigurosa de AGREE II permite identificar fortalezas y limitaciones metodológicas

que informan tanto sobre la confiabilidad de las recomendaciones como sobre las oportunidades de mejora en futuros procesos de desarrollo.

La GPC Consolidada de la OMS para tuberculosis 2025 (Módulo 4: Tratamiento y Cuidado) representa la respuesta institucional más actualizada y exhaustiva al desafío global de TB-MDR y TB-XDR (7). Esta guía consolida, actualiza y jerarquiza las recomendaciones que previamente se encontraban dispersas en múltiples documentos de actualización rápida (2010, 2017, 2020, 2022). Dado que esta GPC servirá como documento de referencia para millones de pacientes y guiará decisiones políticas en decenas de países, una evaluación rigurosa e independiente de su calidad metodológica no representa un mero ejercicio académico, sino una verificación necesaria y responsable de que la guía que dirige la práctica clínica y política global en TB se fundamenta en el más alto estándar de rigor, evidencia y transparencia. Por consiguiente, el objetivo de este estudio se centró en evaluar críticamente la calidad metodológica de la Guía Consolidada de la OMS sobre tuberculosis (Módulo 4) aplicando el instrumento AGREE II mediante evaluación independiente de tres investigadores, identificando fortalezas, limitaciones y oportunidades de mejora, así como analizando la confiabilidad interevaluador en la aplicación de este instrumento.

Métodos

Diseño del Estudio

Se realizó un estudio descriptivo transversal aplicando el instrumento AGREE II, de manera independiente por tres evaluadores expertos en medicina, enfermedades infecciosas, salud pública, epidemiología clínica, cuantificando la calidad metodológica global, analizando la variabilidad interevaluador e identificando fortalezas y debilidades.

Selección del Informe/Guía Evaluada

Se seleccionó para evaluación el documento oficial titulado "WHO consolidated guidelines on tuberculosis. Module 4: Treatment and care" (Organización Mundial de la Salud, Ginebra, 2025; ISBN 978-92-4-010724-3; versión electrónica) (7). La justificación para su selección fue: (i) constituye la publicación más reciente, consolidada y actualizada de la OMS respecto al tratamiento de TB drogo-sensible y drogo-resistente; (ii) su alcance global y autoridad institucional la convierten en un documento prioritario para evaluación metodológica rigurosa; (iii) consolida y jerarquiza recomendaciones previamente dispersas en múltiples actualizaciones rápidas de 2010 a 2025, requiriendo evaluación integral de calidad metodológica.

Herramienta de Evaluación: Instrumento AGREE II

Se utilizó la versión oficial traducida al español del instrumento AGREE II, cuya validez y confiabilidad han sido establecidas en múltiples estudios (2,6,8). El instrumento AGREE II está diseñado específicamente para evaluar tanto el proceso de desarrollo de una guía clínica como la calidad y la exhaustividad de su reporte. El instrumento comprende:

- **23 ítems** distribuidos en 6 dominios temáticos
- **Escala de puntuación Likert de 7 puntos** para cada ítem: 1 (Totalmente en Desacuerdo), 2, 3, 4, 5, 6, 7 (Totalmente de Acuerdo)
- **Dominios evaluados:**
 - Alcance y Propósito (3 ítems): Objetivos, preguntas de salud y población objetivo
 - Implicación de Agentes Clave (3 ítems): Composición del grupo desarrollador, inclusión de pacientes y definición de usuarios objetivo
 - Rigor en el Desarrollo (8 ítems): Búsqueda sistemática, criterios de selección de evidencia, certeza de evidencia, formulación de recomendaciones, consideración de beneficios/daños, vinculación evidencia-recomendación, revisión externa y actualización
 - Claridad de la Presentación (3 ítems): Especificidad de recomendaciones, presentación de opciones de manejo e identificabilidad de recomendaciones clave
 - Aplicabilidad (4 ítems): Barreras y facilitadores, herramientas de implementación, implicaciones de recursos y criterios de monitoreo/auditoría
 - Independencia Editorial (2 ítems): Independencia del financiador y manejo de conflictos de interés

Proceso de Evaluación e Independencia de los Evaluadores

Tres evaluadores (JWAI, JMPV, JCCO), autores del presente manuscrito, con experiencia previa certificada en la evaluación de guías clínicas y en la aplicación del instrumento AGREE II, realizaron la evaluación de forma completamente independiente. Cada evaluador completó el instrumento AGREE II sin conocer las puntuaciones de los demás evaluadores, con el propósito explícito de medir la variabilidad y no de alcanzar un consenso en una única puntuación. Esta aproximación metodológica permitió cuantificar la confiabilidad interevaluador de la aplicación de AGREE II e identificar ítems de interpretación equívoca o problemática.

Análisis de Datos

Las puntuaciones (1-7) de los tres revisores para los 23 ítems se tabularon en una matriz de datos en Microsoft Excel. Para cada dominio, se calculó la puntuación escalada utilizando la fórmula oficial AGREE II (6):

$$\text{Puntuación Escalada (\%)} = [(\text{Suma Obtenida} - \text{Mínimo Posible}) / (\text{Máximo Posible} - \text{Mínimo Posible})] \times 100$$

Donde:

- Puntuación obtenida = suma de puntuaciones (1-7) de todos los ítems del dominio para los 3 evaluadores
- Puntuación mínima posible = número de ítems del dominio $\times$ 3 evaluadores $\times$ 1
- Puntuación máxima posible = número de ítems del dominio $\times$ 3 evaluadores $\times$ 7

Las puntuaciones individuales de cada evaluador se registraron y se calcularon la media aritmética, la desviación estándar (DE) y el rango (mínimo-máximo) para cada dominio. Se calculó el coeficiente Kappa de Fleiss (κ) para medir la concordancia inter-evaluador más allá del azar, tanto a nivel global como a nivel de ítems individuales. La interpretación de valores Kappa se realizó según estándares establecidos: $\kappa < 0.20$ = Pobre; 0.20-0.40 = Débil; 0.41-0.60 = Moderado; 0.61-0.80 = Bueno; 0.81-1.00 = Muy Bueno.

RESULTADOS

Descripción del documento evaluado:

La GPC evaluada (7) es un documento exhaustivo que consolida todas las recomendaciones vigentes de la OMS sobre manejo de TB drogo-sensible (TB-DS) y TB drogo-resistente (TB-DR/TB-MDR/TB-XDR). Las características generales del documento se presentan en la Tabla 1.

Tabla 1. Características Generales de la GPC OMS TB 2025 (Módulo 4) Evaluada

Componente	Detalle
Autor Institucional	World Health Organization (OMS), Departamento de Stop TB, Ginebra
Título	WHO consolidated guidelines on tuberculosis. Module 4: Treatment and care
Año de Publicación	2025
ISBN	978-92-4-010724-3 (versión electrónica)
Alcance	Consolidación de recomendaciones nuevas, actualizadas y validadas para tratamiento de TB-DS y TB-DR en adultos y niños
Objetivo Principal	Proporcionar recomendaciones basadas en evidencia sobre tratamiento, cuidado y seguimiento de TB, así como resumir cambios respecto a guías anteriores (2010-2022)
Usuarios Objetivo	Formuladores de políticas en ministerios de salud, gestores de Programas Nacionales de TB, profesionales de salud (médicos, enfermeras, educadores), organizaciones gubernamentales y no gubernamentales
Financiamiento	Unitaid (agencia internacional no comercial de salud pública)
Documentación Complementaria	Manual operativo separado (WHO Operational Handbook TB Treatment and Care) contiene herramientas de implementación, algoritmos y listas de verificación

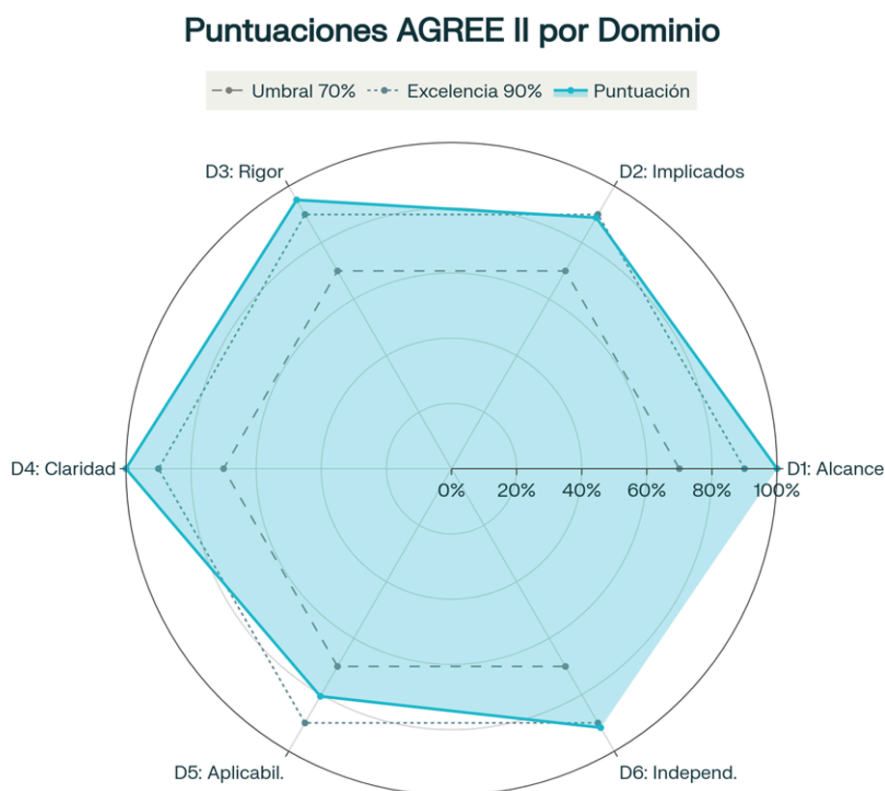
A continuación, se presentan las puntuaciones escaladas de cada dominio, junto con su desviación estándar.

Tabla 2. Puntuaciones Escaladas AGREE II por Dominio (N=3 Evaluadores Independientes)

Dominio	N Ítems	Suma Total (R1+R2+R3)	Mínimo Posible	Máximo Posible	Puntuación Escalada (%)	DE
D1: Alcance y Propósito	3	63	9	63	100,00	0,00
D2: Implicación de Implicados	3	57	9	63	88,89	6,43
D3: Rigor en el Desarrollo	8	161	24	168	95,14	4,17

D4: Claridad de Presentación	3	63	9	63	100,00	0.00
D5: Aplicabilidad	4	70	12	84	80,56	6,29
D6: Independencia Editorial	2	39	6	42	91,67	14,43

Nota: Puntuación promedio global consolidada de los 6 dominios: 92,71 %. DE = desviación estándar.



La Figura 1 presenta, mediante un gráfico de radar, las puntuaciones escaladas de cada dominio, lo que permite una apreciación integral del perfil de calidad metodológica de la GPC. El área rellena en azul representa el desempeño de la GPC en cada dominio. Las líneas de referencia horizontales al 70% y al 90% indican umbrales de calidad metodológica moderada y excelente, respectivamente.

La forma del radar demuestra que la GPC mantiene un desempeño consistente y alto en todos los dominios, excepto por una depresión notable en el Dominio 5 (Aplicabilidad: 80,56%), lo que refleja la limitación mencionada respecto a las herramientas de implementación. Los dominios 1 y 4 alcanzan la perfección (100%), mientras que el Dominio 3 (Rigor en el Desarrollo) alcanza 95,14%, el segundo nivel más alto, indicando excelencia en la metodología subyacente del desarrollo de la guía.

Análisis de Concordancia Inter-evaluador

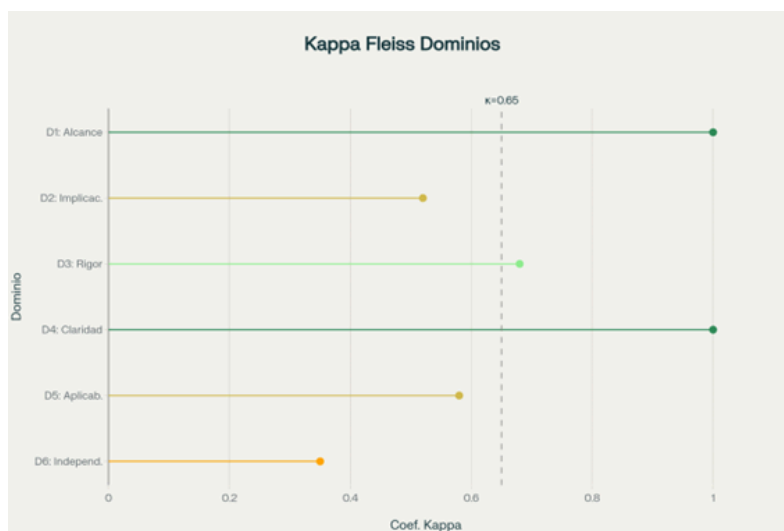
El análisis de concordancia reveló una confiabilidad interevaluador global buena ($\kappa=0.65$). La Tabla 3 detalla los ítems con variabilidad significativa.

Tabla 3. Concordancia Inter-evaluador (Kappa de Fleiss) para Ítems con Alta Variabilidad

Ítem AGREE II	Puntuaciones (R1, R2, R3)	Rango	Kappa (κ)	Interpretación
Ítem 5 (Opiniones pacientes)	5, 6, 5	1	0.47	Moderado
Ítem 14 (Procedimiento actualización)	6, 4, 6	2	0.35	Débil
Ítem 19 (Herramientas implementación)	5, 5, 5	0	1.00	Perfecto
Ítem 21 (Criterios auditoría)	6, 5, 6	1	0.47	Moderado
Ítem 22 (Influencia financiador)	7, 4, 7	3	0.18	Pobre

Nota: 15 ítems (65%) obtuvieron concordancia perfecta ($\kappa=1.0$); 6 ítems (26%) obtuvieron concordancia alta/moderada; 2 ítems (9%) obtuvieron concordancia débil/pobre.

La Figura 2 presenta un gráfico de lollipop mostrando el coeficiente Kappa de Fleiss estratificado por dominio AGREE II. Los colores codifican el nivel de concordancia según estándares internacionales. Esta visualización identifica claramente el Dominio 6 (Independencia Editorial, $\kappa=0.35$) como el área de mayor discordancia inter-evaluador.



La discrepancia más notable se observa en el Ítem 22 (Influencia del financiador), donde dos evaluadores asignaron 7/7 y uno 4/7, lo que da lugar a $\kappa=0.18$ (Pobre). Este hallazgo es de importancia metodológica considerable: sugiere que el instrumento AGREE II presenta ambigüedad interpretativa en los ítems de Independencia Editorial, particularmente cuando se evalúan financiadores no comerciales y existe una política robusta de gestión de conflictos de interés (Ítem 23).

Evaluación Global

Hubo consenso unánime entre los tres evaluadores para otorgar la puntuación máxima de calidad global (7/7). Respecto a la recomendación de uso:

- **JWAI:** Sí (sin modificaciones)
- **JMPV:** Sí, con modificaciones menores (mejoras en ítems 5, 14, 21)
- **JCCO:** Sí (sin modificaciones)

Recomendación consolidada: Sí (unánime), con 66,7 % de recomendación sin modificaciones.

Discusión

Este estudio evaluó la GPC de la OMS sobre TB 2025 mediante una metodología robusta de triple evaluación independiente. Los hallazgos confirman que esta GPC es un documento de calidad metodológica excepcional, estableciendo un estándar de oro para el desarrollo de guías a nivel global, en particular en 'Rigor en el Desarrollo' (Media: 95,14 %).

Análisis de fortalezas metodológicas

Las puntuaciones perfectas y unánimes en el Dominio 1 (Alcance y Propósito) y el Dominio 4 (Claridad de Presentación) demuestran una maestría en la formulación de GPC (9,10). Los objetivos de la guía son explícitos, las preguntas de salud están meticulosamente detalladas en formato PICO (Población, Intervención, Comparador, Desenlaces), y las poblaciones objetivo están claramente definidas con criterios de inclusión y exclusión (7). Del mismo modo, las recomendaciones son inequívocas, específicas y fácilmente identificables mediante tablas resumen y un formato visual claro.

El Dominio 3 (Rigor en el Desarrollo, 95.14%) fue la fortaleza más significativa. La GPC demuestra una aplicación ejemplar del marco GRADE (Grading of Recommendations Assessment, Development, and Evaluation) (11), métodos sistemáticos para la búsqueda de evidencia (incluyendo revisiones sistemáticas y análisis de datos de pacientes individuales) y un proceso de revisión externa transparente y bien documentado (7).

Estos hallazgos contrastan claramente con evaluaciones de GPC de TB más antiguas. El estudio de Gallardo et al. (2011) evaluó 36 GPC sobre TB publicadas entre 1998 y 2008, y encontró que apenas el 7% alcanzó puntuaciones AGREE II satisfactorias (>60% escalado) en el Dominio de Rigor en Desarrollo (12). En contraste, la GPC OMS 2025 alcanza 95,14 % en este dominio crítico, reflejando la adopción generalizada de estándares metodológicos modernos (6, 12, 14).

Análisis de debilidades y áreas de mejora

1. El Desafío de la Aplicabilidad (Dominio 5: 80,56%)

El Dominio 5 (Aplicabilidad) recibió la puntuación más baja. Este es un hallazgo común en la literatura de evaluación de GPC, donde la aplicabilidad es frecuentemente el dominio más débil (9). Las puntuaciones más bajas se centraron en el Ítem 19 (Consejos y/o herramientas para la implementación, promedio de 5,0/7) y en el Ítem 21 (Criterios de monitoreo y/o auditoría, promedio de 5,67/7).

El análisis de las justificaciones de los evaluadores revela la fuente de esta puntuación: todos los evaluadores notaron que la GPC **remite** explícitamente al usuario a un "operational handbook" (manual operativo) separado para estas herramientas, algoritmos y listas de verificación. Esto plantea una cuestión metodológica crucial: ¿Debe una GPC ser un documento autocontenido para puntuar alto en AGREE II, o es la delegación en un manual complementario una estrategia válida?

El instrumento AGREE II evalúa *la GPC* (el documento en mano), y este carece de estas herramientas de implementación integradas. Por lo tanto, una puntuación más baja en estos ítems es metodológicamente correcta según los estándares AGREE II (6). La implicación práctica es que esta separación deliberada de la OMS (GPC vs. Manual) puede ser eficiente en términos organizacionales, pero genera una "brecha de aplicabilidad" en el punto de atención, lo que reduce la utilidad de la GPC como herramienta independiente para el gerente de programa o el clínico que necesita implementar las recomendaciones.

2. La Ambigüedad de la Independencia Editorial (Dominio 6: 91,67%, DE=14.43)

El Dominio 6 (Independencia Editorial) presentó la **mayor discrepancia** entre evaluadores (DE=14.43) y el acuerdo más bajo ($\kappa=0.18$ en el Ítem 22). Este hallazgo no revela necesariamente una falla en la GPC de la OMS, sino una **ambigüedad crítica en el instrumento AGREE II** (4,5).

Todos los evaluadores otorgaron una puntuación perfecta (7/7) al Ítem 23 (Conflictos de interés), elogiando la transparencia "ejemplar" de la GPC al detallar y gestionar activamente los conflictos de interés de los miembros del GDG, incluyendo la recusación de expertos en votaciones específicas (7).

Toda la discrepancia provino del Ítem 22 (Influencia del financiador). Dos evaluadores (JWAI, JCCO) otorgaron 7/7, adoptando una interpretación "holística": razonaron que la financiación proveniente de agencias no comerciales (Unitaid) y la robusta política de gestión de conflictos de interés (Ítem 23) eran salvaguardas suficientes. En contraste, un evaluador (JMPV) otorgó 4/7, adoptando una interpretación "literal": argumentó que la GPC carecía de una declaración explícita y afirmativa (p. ej., "los financiadores no tuvieron influencia en el contenido"), un requisito estricto para el cumplimiento total del ítem según su interpretación de AGREE II.

Esta discrepancia (que se traduce en puntuaciones de dominio del 100% frente al 75% en el mismo documento) demuestra que el Ítem 22 es altamente susceptible a la interpretación del evaluador. Este hallazgo metodológico es una contribución clave de nuestro estudio y subraya la necesidad de que The AGREE Research Trust clarifique la guía de puntuación para los ítems 22 y 23, y su interdependencia (13).

Limitaciones y Fortalezas

Las limitaciones de este estudio incluyen el número de evaluadores (tres), aunque es estándar y ampliamente aceptado en la evaluación AGREE II. En relación con la metodología sugerida por el instrumento, no se realizó una reunión de consenso (por diseño) para calcular una única puntuación "corregida"; en su lugar, el enfoque de este estudio fue el análisis de la variabilidad, que se considera una fortaleza metodológica. Finalmente, el instrumento AGREE II evalúa la calidad del *reporte* de la guía; se asume que esto es un proxy de la calidad subyacente del proceso de desarrollo.

Conclusión

Síntesis de Calidad

La GPC Consolidada de la OMS sobre tuberculosis (Módulo 4, 2025) (7) demuestra una calidad metodológica general excepcional con puntuación promedio de **92,71%** a través de los seis dominios. Sus fortalezas sobresalientes y unánimes en Rigor en el Desarrollo (95,14 %), Alcance (100 %) y Claridad (100 %) la posicionan como un documento ejemplar y de la más alta confianza para guiar la política global de TB.

Síntesis de Mejoras

Las áreas para futura mejora metodológica se centran en el Dominio 5 (Aplicabilidad: 80,56 %). Se alienta a la OMS a considerar la integración de un conjunto básico de herramientas de implementación (algoritmos clave) y de criterios de auditoría programática (Ítems 19 y 21) directamente en el documento principal de la GPC, en lugar de delegarlos

exclusivamente en manuales operativos complementarios. Esto cerraría la brecha entre la recomendación y la implementación en el punto de atención.

Adicionalmente, la inclusión de un procedimiento prospectivo explícito de actualización (Ítem 14) fortalecería aún más la calidad metodológica, al especificar el cronograma y los criterios disparadores para revisiones futuras.

Llamado a la Acción

Este estudio genera dos llamados a la acción:

1. **A los Programas Nacionales de TB y Clínicos:** Se recomienda la adopción y adaptación de esta guía con alta confianza en su validez metodológica.
2. **A los Metodólogos y al Consorcio AGREE:** Se recomienda una revisión de la redacción y de la guía de puntuación del instrumento AGREE II, específicamente de los ítems 22 (Independencia del financiador) y 23 (Conflictos de interés). Se debe aclarar si un manejo robusto de los conflictos de interés (Ítem 23) puede compensar la falta de una declaración explícita de independencia del financiador (Ítem 22), con el fin de reducir la ambigüedad interpretativa y mejorar la confiabilidad interevaluador de esta herramienta crucial.

Referencias

1. Lv H, Wang L, Zhang X, Dang C, Liu F, Zhang X, et al. Further analysis of tuberculosis in eight high-burden countries based on the Global Burden of Disease Study 2021 data. *Infect Dis Poverty*. 2024 Sep 30;13(1):70.
2. Nahid P, Mase SR, Migliori GB, Sotgiu G, Bothamley GH, Brozek JL, et al. Treatment of Drug-Resistant Tuberculosis. An Official ATS/CDC/ERS/IDSA Clinical Practice Guideline. *Am J Respir Crit Care Med*. 2019 Nov 15;200(10):e93–142.
3. Dartois VA, Rubin EJ. Anti-tuberculosis treatment strategies and drug development: challenges and priorities. *Nat Rev Microbiol*. 2022 Nov
4. Nakkazi E. New TB drug could increase treatment success with shorter regime. *Nature Africa*. 2022 Dec 4;
5. Sarkies MN, Jones LK, Gidding SS, Watts GF. Improving clinical practice guidelines with implementation science. *Nat Rev Cardiol*. 2022 Jan 19;19(1):3–4.
6. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *Can Med Assoc J*. 2010 Dec 14;182(18):E839–42.
7. WHO consolidated guidelines on tuberculosis. Module 4: Treatment and care. Geneva; 2025.
8. Maes-Carballo M, García-García M, Martín-Díaz M, Estrada-López CR, Iglesias-Álvarez A, Filigrana-Valle CM, et al. A comprehensive systematic review of colorectal cancer screening clinical practices guidelines and consensus statements. *Br J Cancer*. 2023 Apr 6;128(6):946–57.
9. Hoffmann-Eßer W, Siering U, Neugebauer EAM, Brockhaus AC, Lampert U, Eikermann M. Guideline appraisal with AGREE II: Systematic review of the current evidence on how users handle the 2 overall assessments. *PLoS One*. 2017 Mar 30;12(3):e0174831.
10. Hoffmann-Eßer W, Siering U, Neugebauer EAM, Brockhaus AC, McGauran N, Eikermann M. Guideline appraisal with AGREE II: online survey of the potential influence of AGREE II items on overall assessment of guideline quality and recommendation for use. *BMC Health Serv Res*. 2018 Dec 27;18(1):143.
11. Brożek JL, Akl EA, Alonso-Coello P, Lang D, Jaeschke R, Williams JW, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines. *Allergy*. 2009 May;64(5):669–77.
12. Gallardo CR, Rigau D, Irfan A, Ferrer A, Caylà JA, Bonfill X, et al. Quality of tuberculosis guidelines: urgent need for improvement. *Int J Tuberc Lung Dis*. 2010;14(9):1191–1198.
13. Huang T, Chen Q, Wu G, Tang S. Interpretation of the World Health Organization consolidated guidelines on tuberculosis Module 4: treatment and care (2025 Edition). *Chinese Journal of Tuberculosis and Respiratory Diseases*. 2025 Aug 12;48(8):708–18.