



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

“ADHERENCIA AL ENFOQUE GRADE DE LAS GUÍAS DE
PRÁCTICA CLÍNICA PERUANAS 2018-2022: UNA
REVISIÓN METODOLÓGICA”

TESIS PARA OPTAR EL GRADO DE MAESTRO EN
EPIDEMIOLOGÍA CLÍNICA

VÍCTOR ANDRES VELASQUEZ RIMACHI

LIMA - PERÚ

2024

Asesor

Mg. Carlos Antonio Vicente Salazar Ordóñez

JURADO DE TESIS

MG. César Antonio Loza Munarriz

PRESIDENTE

MG. Luis Alexander Orrego Ferreyros

VOCAL

MG. Manuel Tomas Castillo Portilla

SECRETARIO

DEDICATORIA.

A mi madre, Sonia Rimachi Góngora; a mi padre, Julián Velasquez Ortiz;
y a mi hermana, Nayeli Velasquez Rimachi, por su incondicional apoyo y soporte
durante esta etapa gratificante.

AGRADECIMIENTOS.

A mis colegas y asesor: Carlos Alva-Díaz, Isabel Pinedo-Torres, Gandy
Dolores-Maldonado, Karen Huamán-Sánchez, Nair Javier-Murillo, André
Lapeyre-Rivera y Carlos Salazar-Ordoñez, por su invaluable apoyo en el
desarrollo de este estudio.

FUENTE DE FINANCIAMIENTO.

Trabajo de investigación Autofinanciado.

ADHERENCIA AL ENFOQUE GRADE DE LAS GUÍAS DE PRÁCTICA CLÍNICA PERUANAS 2018-2022: UNA REVISIÓN METODOLÓGICA

INFORME DE ORIGINALIDAD



FUENTES PRIMARIAS

1	formacion.guiasalud.es Fuente de Internet	1%
2	ietsi.essalud.gob.pe Fuente de Internet	1%
3	Submitted to ITESM: Instituto Tecnológico y de Estudios Superiores de Monterrey Trabajo del estudiante	1%
4	Submitted to Universidad de Málaga - Tii Trabajo del estudiante	1%
5	duict.upch.edu.pe Fuente de Internet	1%
6	ddd.uab.cat Fuente de Internet	1%
7	Barriocanal, Ana Maria, Universitat Autònoma de Barcelona. Departament de Farmacologia, de Terapèutica i de Toxicologia. "Estudi observacional retrospectiu per avaluar la	<1%

TABLA DE CONTENIDO

INTRODUCCIÓN.....	1
OBJETIVOS.....	3
OBJETIVO GENERAL	3
OBJETIVOS ESPECÍFICOS	3
METODOLOGÍA.....	4
DISEÑO DE ESTUDIO	4
CRITERIOS DE ELEGIBILIDAD	4
ESTRATEGIA DE BÚSQUEDA	5
SELECCIÓN DE ESTUDIOS	5
EXTRACCIÓN DE DATOS.....	6
ANÁLISIS ESTADÍSTICO	7
RESULTADOS	8
SELECCIÓN DE GUÍAS.....	8
CARACTERÍSTICAS DE LAS GUÍAS.....	8
RIGOR DEL DESARROLLO Y ADHERENCIA A GRADE.....	9
DEFINICIÓN DE LA CERTEZA DE LA EVIDENCIA (Q1-Q3).....	9
FORMULACIÓN DE LA PREGUNTA DE SALUD (Q4).....	10
SELECCIÓN Y CLASIFICACIÓN DE LA IMPORTANCIA DE LOS	

RESULTADOS (Q5-Q6).....	10
EVALUACIÓN DE LA CERTEZA DE LA EVIDENCIA (Q7)	10
DETERMINACIÓN DE LA CERTEZA DE LA EVIDENCIA (Q8-Q9).....	11
RESUMEN DE HALLAZGOS (Q10-Q11)	11
ESTABLECIMIENTO DE RECOMENDACIONES BASADAS EN EVIDENCIA (Q12-Q15)	12
CALIDAD METODOLÓGICA	12
DISCUSIÓN	13
CONCLUSIONES.....	16
RECOMENDACIONES	17
REFERENCIAS BIBLIOGRÁFICAS	18
ANEXO	24

RESUMEN

Objetivo: Evaluar la adherencia al enfoque GRADE de las guías de práctica clínica (GPC) denominadas *de novo*, sobre intervenciones terapéuticas y diagnósticas, desarrolladas por las instituciones públicas gubernamentales peruanas y publicadas entre 2018 y 2022.

Metodología: Se realizó una revisión metodológica sobre las GPC peruanas con alcance nacional denominadas *de novo* que declararon el uso de la metodología GRADE. Se describió el porcentaje de adherencia al enfoque GRADE incluyendo la evaluación metodológica con el instrumento AGREE II – Dominio 3. Se investigó la relación entre el grado de adherencia al enfoque GRADE y el rigor metodológico evaluado por el AGREE II – Dominio 3, presentando resultados dentro de un intervalo de confianza del 95% y considerando un valor de $p < 0,05$ como significativo, utilizando el software STATA 18.

Resultados: Se revisaron 746 registros de bases de datos y 226 documentos de sitios web peruanos, seleccionando 45 guías de GPC que declararon usar la metodología GRADE. La mayoría de las GPC fueron publicadas entre 2021 y 2022, y el 64,4% estaban dirigidas a mayores de 18 años. Oncología fue la especialidad más común (15,5%). El número mediano de preguntas PICO fue 7. El 51,1% de las guías incluyó tablas de resumen de hallazgos (SoF) y el 53,3% desarrolló un marco para la toma de decisiones (EtD) completo. La adherencia al enfoque GRADE fue alta (>80%) en la mayoría de las guías. Sin embargo, aunque el 77,8% siguió las definiciones propuestas por GRADE, varias GPC carecieron de definiciones explícitas de certeza de la evidencia. Todas las guías estructuraron sus preguntas siguiendo el marco PICO, y el 95,6% evaluó la certeza de la evidencia de manera

rigurosa. El 84,4% de las GPC presentó un marco EtD, y el 80% justificó la fuerza y dirección de las recomendaciones. La calidad metodológica evaluada con AGREE II – Dominio 3 mostró una mediana del 75%, indicando un rigor metodológico adecuado.

Conclusión: Los hallazgos muestran una adecuada adherencia al enfoque GRADE y una adecuada rigurosidad metodológica en las GPC peruanas, lo que contribuye a mejorar y estandarizar las recomendaciones en salud.

Palabras clave: Enfoque GRADE; Métodos; Guía de Práctica Clínica; Perú
(Fuente DeCS Bireme).

ABSTRACT

Objective: To evaluate the adherence to the GRADE approach in called de novo clinical practice guidelines (CPGs) for therapeutic and diagnostic interventions developed by Peruvian governmental public institutions and published between 2018 and 2022.

Methods: A methodological review was conducted on called de novo Peruvian national CPGs that declared the use of the GRADE methodology. The adherence percentage to the GRADE approach was described, including methodological evaluation using the AGREE II – Domain 3 instrument. The relationship between the degree of adherence to the GRADE approach and the methodological rigor assessed by AGREE II – Domain 3 was investigated, presenting results within a 95% confidence interval and considering a p-value <0.05 as significant, using STATA 18 software.

Results: We reviewed 746 database records and 226 documents from Peruvian websites, ultimately selecting 45 CPGs that declared the use of the GRADE methodology. Most CPGs were published between 2021 and 2022, with 64.4% targeting adults over 18 years old. Oncology was the most common specialty (15.5%). The median number of PICO questions was 7. A total of 51.1% of the guidelines included Summary of Findings (SoF) tables, and 53.3% developed a complete Evidence to Decision (EtD) framework. Adherence to the GRADE approach was high (>80%) in most guidelines. However, while 77.8% followed the GRADE definitions, several CPGs lacked explicit definitions of evidence certainty. All guidelines structured their questions using the PICO framework, and 95.6% rigorously evaluated the certainty of evidence. Furthermore, 84.4% of the CPGs

presented an EtD framework, and 80% provided transparent justification for the strength and direction of their recommendations. Methodological quality assessed using AGREE II – Domain 3, showed a median score of 75%, indicating adequate methodological rigor.

Conclusion: The findings indicate appropriate adherence to the GRADE approach and sufficient methodological rigor in Peruvian clinical practice guidelines, contributing to the improvement and standardization of health recommendations.

Keywords: GRADE Approach; Methods; Practice Guideline; Peru. (Fuente: MeSH NLM).

I. INTRODUCCIÓN

Las guías de práctica clínica (GPC) emiten recomendaciones para orientar la toma de decisiones en el cuidado de la salud^{1,2} y tienen como objetivo estandarizar las ofertas de atención al paciente para mejorar los resultados de salud en la población^{3,4}. El desarrollo de las GPC es un proceso complejo que requiere la incorporación de la mejor evidencia científica disponible en una circunstancia clínica específica para formular recomendaciones robustas⁵⁻⁷. Sin embargo, la mejora continua de este proceso requiere métodos adecuados para interpretar la evidencia identificada y la toma de decisiones subsiguiente en el contexto de una GPC⁸⁻¹².

El enfoque *Grading of Recommendations Assessment, Development, and Evaluation* (GRADE) permite establecer la certeza de la evidencia y formular recomendaciones a través de la evaluación del riesgo-beneficio de diferentes intervenciones¹. GRADE utiliza un proceso sistemático que permite: i) plantear la pregunta de salud, ii) seleccionar y priorizar los resultados clínicamente relevantes, iii) evaluar la certeza de la evidencia científica y resumir los hallazgos (tabla SoF), iv) establecer la certeza de la evidencia, y v) determinar la fuerza y dirección de las recomendaciones de salud^{13,14}. Por lo tanto, el uso de este enfoque es fundamental para guiar la toma de decisiones en salud y contribuir a la implementación de tratamientos efectivos y seguros para los pacientes.

En Australia, Barker et al.¹⁵ informaron que los desarrolladores de GPC están utilizando el enfoque GRADE con inconsistencias entre la priorización de resultados relevantes (27%; n=4/15), resumen de los hallazgos (53%; n=8/15), y el proceso de EtD (20%; n=3/15). Mientras que en América Latina y el Caribe,

Cabrera et al.¹⁶ informaron que solo nueve países habían incorporado el enfoque GRADE en sus regulaciones para el desarrollo de GPC hasta 2019. En Perú, el uso del enfoque GRADE ha sido regulado para GPC de Novo desde 2015¹⁷, siendo promovido por varias instituciones gubernamentales como el “Instituto de Evaluación de Tecnologías en Salud e Investigación” (IETSI) del Seguro Social de Salud del Perú (EsSalud) y la “Unidad de Análisis y Generación de Evidencias en Salud Pública” (UNAGESP) del Instituto Nacional de Salud^{5,17}.

Sin embargo, el uso del enfoque GRADE en Perú es reciente y no está claro si el desarrollo de sus GPC se lleva a cabo según este enfoque, lo cual afectaría su generalizabilidad, reproducibilidad y comparaciones futuras con otras GPC. El propósito de esta revisión metodológica es evaluar la adherencia al enfoque GRADE de las GPC de Novo, sobre intervenciones terapéuticas y diagnósticas, desarrolladas por las instituciones públicas gubernamentales peruanas y publicadas entre 2018 y 2022.

II. OBJETIVOS

OBJETIVO GENERAL

- Evaluar la adherencia al enfoque GRADE de las GPC de Novo, sobre intervenciones terapéuticas y diagnósticas, desarrolladas por las instituciones públicas gubernamentales peruanas y publicadas entre 2018 y 2022.

OBJETIVOS ESPECÍFICOS

- Describir las características de las GPC peruanas *de novo*, incluyendo la especialidad, la enfermedad de referencia, el año de realización, la organización desarrolladora, la normativa empleada y la declaración de financiamiento y conflicto de intereses.
- Evaluar la adherencia de las GPC peruanas *de novo* al enfoque GRADE, utilizando una lista de verificación adaptada al español peruano (GRADE ADHERENCE CHECK-LIST) (ver anexo 01).
- Evaluar la calidad metodológica del desarrollo de las GPC peruanas *de novo* que han declarado el uso del enfoque GRADE como metodología, utilizando el instrumento AGREE-II (dominio 3).

III. METODOLOGÍA

DISEÑO DE ESTUDIO

Se llevó a cabo una revisión metodológica definida como un tipo de investigación secundaria sistemática que se enfoca en compilar y resumir las prácticas metodológicas más avanzadas en un campo o tema relevante para examinar cualquier aspecto metodológico relacionado con el diseño, la ejecución y la revisión de estudios de investigación, así como con la síntesis de evidencia^{18,19}. Empleamos procesos sistemáticos de búsqueda, selección, extracción y análisis de información tomando como referencia las pautas de la colaboración Cochrane⁹, el "*Preferred Reporting Items for Systematic Reviews and Meta-Analysis Extension for Scoping Reviews*" (PRISMA-ScR)^{20,21} y el capítulo 11 de la "*JBI for Evidence Synthesis guideline for scoping review*"²².

CRITERIOS DE ELEGIBILIDAD

Se incluyeron GPC denominadas *de novo* que hayan sido desarrolladas por entidades gubernamentales o institutos de salud en Perú con alcance nacional, destinadas a proporcionar recomendaciones basadas en evidencia para intervenciones terapéuticas y diagnósticas. Una GPC *de novo* es aquella que se desarrolla íntegramente a partir de una revisión sistemática de la evidencia, sin recurrir a la adaptación o adopción directa de guías preexistentes. Solo se incluyeron las GPC que declararon utilizar el enfoque GRADE en su desarrollo. Se consideró la versión más reciente de una GPC si existían múltiples versiones, y el texto completo debía estar disponible en inglés o español. Se excluyeron las GPC que no

estaban completamente disponibles para revisión, aquellas que carecían de metodología detallada para la evaluación del nivel de evidencia, y las GPC basadas únicamente en consenso de expertos.

ESTRATEGIA DE BÚSQUEDA

La estrategia de búsqueda tuvo como objetivo identificar estudios publicados en base de datos y repositorios y páginas web de institucionales gubernamentales. Se desarrollaron estrategias de búsqueda específicas para cada fuente de datos utilizando palabras clave relacionadas con "guías de práctica clínica," "Perú," "GRADE," etc. (ver anexo 02). La búsqueda incluyó estudios publicados en inglés o español en los últimos cinco años hasta el 31 de diciembre del 2022.

SELECCIÓN DE ESTUDIOS

Para los documentos publicados en bases de datos, se eliminaron duplicados en EndNote X9, seguido de una selección de duplicados (ALR y NJM) basada en títulos y resúmenes, y luego una selección de texto completo. Se registraron y resolvieron por consenso o por un tercer revisor (VVR) las razones de exclusión y los desacuerdos.

Para los documentos obtenidos de entidades gubernamentales o institutos de salud en Perú, se listaron los documentos de texto completo disponibles para descarga con el objetivo de contrastarlos finalmente con los documentos identificados en las bases de datos. Durante esta etapa, se confirmó que los documentos cumplían con los estándares de una GPC aprobada por el Instituto de Medicina (IOM)² y se utilizó la metodología GRADE para llevar a cabo una codificación preliminar. Finalmente,

se verificó si la GPC se denominaba como *de novo*, excluyendo aquellas adoptadas-adaptadas para la codificación final.

EXTRACCIÓN DE DATOS

Se extrajeron los datos utilizando un formulario en Excel. Las dudas se aclararon contactando a los desarrolladores de las guías. Los datos incluyeron detalles sobre especialidad, enfermedad de referencia, año de publicación, organización desarrolladora, preguntas que presentaron tabla de perfil de evidencia o resumen de hallazgos (SoF) y preguntas que desarrollaron EtD completa-parcial. Además, se evaluó la adherencia a GRADE utilizando una lista de verificación adaptada de Días et al.²⁴ y basada en el capítulo 8 del manual GRADE¹⁴ que contiene un total de 18 preguntas (ver anexo 01).

La elaboración de esta lista respondió a la necesidad de una herramienta para extraer datos relacionados con GRADE de las GPC. Un equipo de expertos metodológicos revisó la pertinencia y la comprensión de la traducción al español peruano de la lista de chequeo de Días et al.²⁴ mediante consenso por mayoría (ver anexo 03). Dado que esta lista no operativiza la valoración de los resultados de cada pregunta en sus reportes, el equipo de metodólogos desarrolló criterios inéditos basados en el manual GRADE para mejorar la transparencia y facilitar la evaluación. Con estos criterios dos revisores independientes evaluaron el rigor del desarrollo de las GPC que declararon el uso del enfoque GRADE utilizando la herramienta AGREE II – Dominio 3.

ANÁLISIS ESTADÍSTICO

Se realizó un análisis descriptivo de las características inherentes a las GPC. Este examen abarcó múltiples dominios: las especialidades médicas implicadas, las enfermedades o condiciones objetivo, los perfiles clínicos de las poblaciones destinatarias, los años de publicación, las entidades desarrolladoras y la extensión de las tablas SoF asociados con preguntas PICO.

Posteriormente, se evaluó la adherencia a la metodología GRADE determinando la proporción de preguntas clínicas que recibieron un “SI” en la lista de verificación de adherencia a GRADE. También se cuantificó el cumplimiento con AGREE II – Dominio 3, una métrica que evalúa el rigor del desarrollo. Además, se realizó un análisis basado en el número de preguntas PICO, categorizándolas en dos grupos usando la mediana del número de preguntas PICO como referencia. Se presentan los hallazgos dentro de un intervalo de confianza del 95%, interpretando un valor p inferior a 0,05 como estadísticamente significativo. Se ejecutaron todos los análisis estadísticos utilizando el software STATA 18.

IV. RESULTADOS

SELECCIÓN DE GUÍAS

Se identificaron un total de 746 registros de bases de datos y 226 documentos de sitios web de GPC y fuentes organizacionales peruanas. Después de la exclusión de 169 duplicados, se evaluaron 577 registros de bases de datos mediante una revisión de títulos y resúmenes. Esta revisión culminó con la evaluación de 30 registros para determinar su elegibilidad mediante una evaluación de texto completo. Paralelamente, la búsqueda en sitios web de GPC y organizaciones peruanas identificó 223 documentos para evaluación de texto completo, ya que 3 documentos no estaban disponibles para descarga. Tras la revisión de los textos completos, se incluyeron 45 GPC que declararon haber empleado una metodología *de novo* y empleaban el sistema GRADE. Los detalles de los 207 documentos excluidos se describen en el diagrama de flujo PRISMA (ver anexo 04).

CARACTERÍSTICAS DE LAS GUÍAS

La mayoría de las GPC fueron publicadas en 2021 (n=10), seguidas por 2022 (n=10), 2020 (n=9), 2019 (n=8) y 2018 (n=8). En cuanto al organismo desarrollador, la mayoría pertenecía al IETSI (n=41) (ver anexo 05). La mayoría de GPC estaban dirigidas a una población mayor de 18 años (n=29; 64,4%) (ver anexo 06). En relación con la especialidad, la especialidad con más GPC incluidas fue oncología (n=7; 15,5%) (ver anexo 07). La mayoría de las GPC abarcaban más de un tópico (n=19; 42,2%), siendo "manejo" (n=16; 35,5%) el más común (ver anexo 08).

El número de preguntas PICO tuvo una mediana de 7 y un rango de 1 a 12 preguntas. La mayoría de las GPC (n=23; 51,1%) desarrolló una tabla SoF o tablas de perfil de evidencia para todas sus preguntas PICO. En cuanto a la sección EtD, el 53,3% desarrolló un EtD completo para todas sus preguntas (n=24), mientras que un 35,5% (n=16) desarrollo un EtD parcial para todas sus preguntas (ver anexo 09).

RIGOR DEL DESARROLLO Y ADHERENCIA A GRADE

El anexo 10 detalla los resultados generales de la puntuación de adherencia a GRADE tras 45 evaluaciones.

DEFINICIÓN DE LA CERTEZA DE LA EVIDENCIA (Q1-Q3)

Las definiciones utilizadas por las GPC para evaluar la certeza de la evidencia (Q1), la clasificación de la certeza (Q2) y la fuerza de las recomendaciones (Q3) se alinearon en gran medida con las establecidas por el Grupo de Trabajo GRADE. En cuanto a Q1, la mayoría de las GPC se adhirieron a la definición propuesta por GRADE (n=35; 77,8% para cada pregunta). Sin embargo, es notable que, aunque la mayoría de las GPC referenciaron la metodología GRADE, no definieron explícitamente la certeza de la evidencia. En cambio, indicaron adherencia a la metodología GRADE, que abarca nueve aspectos para evaluar la certeza de la evidencia: tipo de estudio, riesgo de sesgo, inconsistencia, evidencia indirecta, imprecisión, sesgo de publicación, tamaño del efecto, relación dosis-respuesta y efectos de confusión (estos tres últimos aplicables a estudios observacionales). En cuanto a Q2 y Q3, una mayor proporción de GPC se refirieron a estos conceptos (Q2: n=42; 93,3% [Q3: n=39; 86,7%]), aunque fueron criticadas por no proporcionar definiciones explícitas.

FORMULACIÓN DE LA PREGUNTA DE SALUD (Q4)

Todas las GPC extraídas estructuraron la(s) pregunta(s) de salud siguiendo el marco PICO (n=45; 100%). Los desarrolladores incorporaron explícitamente cada componente del marco PICO en la pregunta de salud, de acuerdo con el Capítulo 2 del manual GRADE.

SELECCIÓN Y CLASIFICACIÓN DE LA IMPORTANCIA DE LOS RESULTADOS (Q5-Q6)

Casi todas las GPC evaluadas abordaron la clasificación de los resultados por pregunta (Q5) y la graduación de la certeza basada en cada resultado (Q6) (Q5: n=44; 97,8% [Q6: n=45; 100%]). Sin embargo, la guía G51 no cumplió completamente con este criterio para Q5, ya que solo enumeró resultados "importantes" y "críticos", omitiendo "no importantes".

EVALUACIÓN DE LA CERTEZA DE LA EVIDENCIA (Q7)

La mayoría de las GPC (n=43; 95,6%) indicaron que evaluaron la certeza de la evidencia, considerando los dominios descritos en el Capítulo 5 del manual GRADE. Estos dominios incluyen el diseño del estudio, riesgo de sesgo, inconsistencia (heterogeneidad), evidencia indirecta, imprecisión, sesgo de publicación, gran tamaño del efecto (evaluado en estudios observacionales), consideración de cualquier factor de confusión plausible que pueda afectar el efecto demostrado (evaluado en estudios observacionales), y gradiente dosis-respuesta (evaluado en estudios observacionales). Sin embargo, dos GPC (G3 y G17) solo

abordaron parcialmente estos dominios.

DETERMINACIÓN DE LA CERTEZA DE LA EVIDENCIA (Q8-Q9)

La mayoría de las GPC asignaron una calificación de certeza de la evidencia para cada resultado utilizando las categorías GRADE de alta, moderada, baja y muy baja (n=44; 97,8%). Notablemente, una guía (G49) utilizó estas categorías, pero no las informó por resultado, por lo que se registró como 'no'. En cuanto a la certeza general de la evidencia, la mayoría de las GPC (n=32; 71,1%) consideraron este factor en todos los resultados. Sin embargo, algunas GPC (n=13; 28,9%) no calificaron la certeza/calidad general ni detallaron el proceso de calificación en la tabla EtD o en los anexos.

RESUMEN DE HALLAZGOS (Q10-Q11)

La mayoría de las GPC (n=44; 97,8%) presentaron un perfil de evidencia GRADE y/o una tabla de resumen de hallazgos, un método clave para la presentación transparente de las evaluaciones y resultados GRADE (Q10). Todas estas GPC informaron los detalles requeridos por GRADE en estas tablas (Q11), y todas se basaron en revisiones sistemáticas (n=45; 100%). Casi todas las GPC (n=44; 97,8%) describieron claramente la evidencia evaluada, mientras que una guía (G53) no describió el tipo de evidencia incluida para la evaluación de la certeza de cada resultado (Q11b). Sin embargo, todas las GPC (n=45; 100%) describieron de manera transparente las razones para degradar o mejorar la evidencia (Q11c).

ESTABLECIMIENTO DE RECOMENDACIONES BASADAS EN EVIDENCIA (Q12-Q15)

Todas las GPC (n=45; 100%) consideraron los criterios GRADE en su totalidad (balance de consecuencias deseables e indeseables, certeza de la evidencia, valores y preferencias de los afectados, y uso de recursos) al determinar la fuerza de una recomendación (Q12). Para Q13, todas las GPC (n=45; 100%) utilizaron las categorías GRADE (fuerte y débil/condicional) al informar la fuerza de las recomendaciones. Sin embargo, la mayoría de las GPC (n=38; 84,4%) presentaron el marco de evidencia para la decisión (EtD) GRADE, ya sea en la sección EtD o de manera narrativa (Q14). Finalmente, la mayoría de las GPC (n=36; 80%) informaron de manera transparente las razones para deliberar sobre la fuerza (fuerte o débil/condicional) y la dirección (a favor o en contra) de la recomendación, en la sección EtD o en el desarrollo de cada pregunta (Q15).

CALIDAD METODOLÓGICA

Se evaluó la calidad metodológica de las 45 guías que siguieron la adherencia a GRADE utilizando la herramienta AGREE II – Dominio 3 (ver anexo 11 y 12). La calidad metodológica de cada guía se ha presentado en el anexo 10. En general, la mediana de la calidad metodológica (%) para las 45 GPC fue del 75% (IQR = 6,25%).

V. DISCUSIÓN

El propósito de este estudio ha sido evaluar la adherencia al enfoque GRADE en el desarrollo de GPC Peruanas, con el objetivo de determinar cómo se implementa esta metodología y su alineación con las recomendaciones internacionales para la estandarización y mejora de la calidad de la atención sanitaria^{12,24}. El análisis de las GPC denominadas *de novo* peruanas publicadas entre 2018 y 2022 revela una alta (>80%) pero variable adherencia al enfoque GRADE, así como una adecuada calidad metodológica (>60%) según la herramienta AGREE II – Dominio 3. Se utilizó únicamente el Dominio 3 de AGREE II porque este dominio se centra específicamente en la calidad metodológica del desarrollo de las guías, en comparación con los otros dominios que abarcan aspectos más generales o administrativos. La existencia de regulaciones obligatorias para el uso de GRADE en Perú desde 2015²⁴ pudiera haber influido en estandarizar su aplicación entre instituciones peruanas. A pesar de que la aplicación de la metodología GRADE en el contexto latinoamericano es escasa por su complejidad y necesidad de una capacitación adecuada^{16, 25,26}

Comparativamente, nuestro estudio muestra paralelismos y contrastes con la investigación de Barker et al.¹⁵ en Australia, donde solo el 11,2% de las GPC revisadas siguieron la metodología GRADE, mostrando una adopción gradual y desigual entre diferentes organizaciones. A pesar de las diferencias contextuales, ambos estudios enfatizan la importancia de la capacitación y el desarrollo de recursos para una comprensión y aplicación uniforme de GRADE. Nuestro análisis reveló altos índices de adherencia a GRADE con menos variabilidad. Esta diferencia podría deberse a la inclusión de solo GPC denominadas *de novo* que

cumplieron consistentemente con los componentes GRADE, durante un período en el que las regulaciones exigían el uso de estas metodologías^{14,26}. Sin embargo, la variabilidad como elemento constante en la implementación de GRADE destaca la necesidad de estandarizar los procesos de desarrollo de GPC para adaptarse a las necesidades de cada contexto¹⁶.

En nuestro análisis, dos áreas específicas dentro de la metodología GRADE mostraron tasas de adherencia inferiores al 80%, estas fueron definir el término 'certeza de la evidencia' (Q1) y calificar la 'certeza/calidad general de la evidencia' (Q9). Estos problemas conciernen principalmente a la terminología central del marco GRADE. La menor adherencia a estos componentes puede indicar una falta de comprensión clara de la metodología GRADE en sí^{7,27}. A pesar de los altos índices generales de adherencia a GRADE, la aplicación precisa y consistente de los conceptos terminológicos clave sigue siendo un desafío. Esto resalta la necesidad de guías más comprensivas y sesiones de capacitación específicas que se centren particularmente en estos aspectos²⁷.

La adecuada calidad metodológica de las GPC peruanas refleja un avance significativo en su desarrollo, superando los niveles reportados en los estudios de Canelo et al. y Timana et al.^{28,29}, pero es relevante mencionar que estos estudios no se enfocaron en GPC que solamente hayan empleado el enfoque GRADE. Esta mejora se atribuye a las regulaciones recientes y la institucionalización de equipos metodológicos especializados^{5,6,17}. También puede estar influenciada por las diferencias en los contextos regulatorios¹⁶, sugiriendo que el entorno regulatorio peruano puede impactar de manera única esta relación al limitar un número de preguntas PICO por GPC. Estos hallazgos subrayan la necesidad de una mayor

investigación para comprender completamente estas dinámicas. Además, es crucial utilizar la herramienta AGREE II – Dominio 3 y GRADE de manera complementaria, adaptando y estandarizando los métodos de evaluación para ajustarse a contextos específicos.

Es imperativo reconocer las limitaciones de nuestro estudio, incluyendo la posibilidad de que no se capturaron todas las GPC relevantes o que la evaluación de la adherencia al enfoque GRADE no reflejó completamente su implementación. Este estudio se centró exclusivamente en GPC denominadas *de novo*, excluyendo aquellas que fueron adoptadas o adaptadas, lo cual es relevante ya que las metodologías específicas para la adaptación o adopción de guías, como ADAPTE³⁰ o ADOLOPMENT³¹, no se alinean directamente con todos los criterios establecidos por GRADE. Esta selección sugiere que GRADE es particularmente adecuado para el desarrollo de GPC *de novo*. El contenido del instrumento empleado para valorar la adherencia a GRADE fue adaptado por consenso de un equipo de profesionales peruanos para su aplicación en nuestro contexto, pero aún no cuenta con un proceso de validación formal por ello sus resultados deben considerarse exploratorios. Sin embargo. Puede considerarse útil para evaluar el cumplimiento general de los dominios requeridos por GRADE, pero no el detalle de cada uno de estos como los componentes de la certeza o del EtD. La variabilidad en la interpretación de GRADE entre los evaluadores resalta la necesidad de un marco de evaluación más uniforme y futuras validaciones de este instrumento.

VI. CONCLUSIONES

La adherencia al enfoque GRADE y la calidad metodológica de las GPC peruanas que refirieron emplear el enfoque GRADE son altas. A pesar de los desafíos identificados en la implementación efectiva de GRADE, los hallazgos subrayan una oportunidad invaluable para optimizar la estandarización y calidad de las GPC mediante una mayor inversión en capacitación y recursos, así como un compromiso sostenido por parte de las instituciones de salud. Al utilizar este estudio como marco de referencia, las instituciones de salud y los desarrolladores de GPC pueden dirigir sus esfuerzos hacia la optimización de los procesos para incorporar la metodología GRADE en el desarrollo de GPC, abordando tanto los desafíos locales como globales, y avanzando hacia la alineación con los estándares internacionales, fortaleciendo así la toma de decisiones clínicas basadas en evidencia en Perú.

VII. RECOMENDACIONES

- Continuar con la capacitación y el entrenamiento de los equipos de desarrollo de GPC en la metodología GRADE para asegurar una comprensión y aplicación uniforme de sus principios, particularmente en términos de terminología y calificación de la certeza de la evidencia.
- Realizar la evaluación continua de las GPC utilizando herramientas como GRADE CHECK LIST y AGREE II para identificar y corregir inconsistencias en la implementación.
- Promover la colaboración entre los expertos metodológicos de las distintas entidades gubernamentales y académicas para compartir buenas prácticas y recursos.
- Realizar estudios adicionales para entender mejor las dinámicas entre la adherencia a GRADE y la calidad metodológica de las GPC para determinar cómo diferentes contextos regulatorios y operativos impactan esta relación.
- Desarrollar y validar herramientas de evaluación más robustas y específicas que puedan capturar con mayor precisión la implementación de la metodología GRADE y otros enfoques metodológicos.
- Participar en estudios comparativos internacionales para evaluar cómo otros países están implementando GRADE y otras metodologías en el desarrollo de sus GPC.
- Explorar el uso de tecnologías avanzadas como la inteligencia artificial y el aprendizaje automático para apoyar el proceso de desarrollo de GPC, mejorar la revisión de literatura y la síntesis de evidencia, y aumentar la precisión y eficiencia del proceso.

VIII. REFERENCIAS BIBLIOGRÁFICAS

1. Sanabria AJ, Rigau D, Rotaecche R, Selva A, Marzo-Castillejo M, Alonso-Coello P. Sistema GRADE: metodología para la realización de recomendaciones para la práctica clínica. *Aten Primaria*. Elsevier Doyma; 2015;47(1):48–55. DOI: 10.1016/J.APRIM.2013.12.013
2. Graham R, Mancher M, Wolman DM, Greenfield S, Steinberg E. *Clinical Practice Guidelines We Can Trust*. Clinical Practice Guidelines We Can Trust. National Academies Press; 2011; DOI: 10.17226/13058
3. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. *BMJ*; 1999;318(7182):527–30. DOI: 10.1136/BMJ.318.7182.527
4. Gutierrez EL, Piazza M, Gutierrez-Aguado A, Hajar G, Carmona G, Caballero P, et al. Uso de la evidencia en políticas y programas de salud aportes del Instituto Nacional de Salud. *Rev Peru Med Exp Salud Publica*. Instituto Nacional de Salud; 2016;33(3):580–4. DOI: 10.17843/RPMESP.2016.333.2308
5. Timana-Ruiz R. Desarrollo de Guías de Práctica Clínica en el Seguro Social del Perú. *Revista del Cuerpo Médico Hospital Nacional Almanzor Aguinaga Asenjo*. Medical Body of the Almanzor Aguinaga Asenjo National Hospital; 2019;12(2):95–6. DOI: 10.35434/RCMHNA.2019.122.503
6. Médica Peruana A, Fradis Eriberto Gil Olivares C, Alva Diaz C, García-Mostajo JA, Gil-Olivares F, Timana R, et al. Guías de práctica clínica:

evolución, metodología de elaboración y definiciones actuales. *Acta Médica Peruana* [Internet]. Colegio Médico del Perú; 2017 [cited 2024 Mar 31];34(4):317–22. Available from: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1728-59172017000400010&lng=es&nrm=iso&tlng=es

7. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schünemann HJ. What is “quality of evidence” and why is it important to clinicians? *BMJ*. *BMJ*; 2008;336(7651):133–7. DOI: 10.1136/BMJ.39490.551019.BE
8. Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. *JAMA*. *JAMA*; 1999;281(20):1900–5. DOI: 10.1001/JAMA.281.20.1900
9. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the *Cochrane Handbook for Systematic Reviews of Interventions*. *Cochrane Database Syst Rev*. *Cochrane Database Syst Rev*; 2019;10(10):ED000142. DOI: 10.1002/14651858.ED000142
10. Moberg J, Oxman AD, Rosenbaum S, Schünemann HJ, Guyatt G, Flottorp S, et al. The GRADE Evidence to Decision (EtD) framework for health system and public health decisions. *Health Res Policy Syst*. *Health Res Policy Syst*; 2018;16(1). DOI: 10.1186/S12961-018-0320-2
11. Alonso-Coello P, Oxman AD, Moberg J, Brignardello-Petersen R, Akl EA, Davoli M, et al. GRADE Evidence to Decision (EtD) frameworks:

- a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ*. *BMJ*; 2016;353. DOI: 10.1136/BMJ.I2089
12. Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE Guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *J Clin Epidemiol*. *J Clin Epidemiol*; 2016;76:89–98. DOI: 10.1016/J.JCLINEPI.2016.01.032
 13. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. *BMJ*; 2008;336(7650):924–6. DOI: 10.1136/BMJ.39489.470347.AD
 14. Schünemann H, Brozek J, Guyatt G, Oxman A. [Internet]. 2013. The GRADE handbook [cited 2024 Mar 31]. Available from: <https://gdt.gradepro.org/app/handbook/handbook.html>
 15. Barker TH, Dias M, Stern C, Porritt K, Wiechula R, Aromataris E, et al. Guidelines rarely used GRADE and applied methods inconsistently: A methodological study of Australian guidelines. *J Clin Epidemiol*. Elsevier Inc.; 2021;130:125–34. DOI: 10.1016/j.jclinepi.2020.10.017
 16. Cabrera PA, Pardo R. Review of evidence based clinical practice guidelines developed in Latin America and Caribbean during the last decade: an analysis of the methods for grading quality of evidence and topic prioritization. *Global Health*. *Global Health*; 2019;15(1). DOI: 10.1186/S12992-019-0455-0

17. Peruana AM, Jorge C, Mostajo G, García-Mostajo JA, Alva-Díaz C, Suárez V, et al. Desarrollo de guías de práctica clínica basadas en evidencia en Perú desde el sector público. *Acta Médica Peruana* [Internet]. Colegio Médico del Perú; 2017 [cited 2024 Mar 31];34(3):203–7. Available from: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1728-59172017000300007&lng=es&nrm=iso&tlng=es
18. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC medical research methodology*. 2018 Dec;18:1-9.
19. Chong SW, Reinders H. A methodological review of qualitative research syntheses in CALL: The state-of-the-art. *System*. 2021 Dec 1;103:102646.
20. Tricco AC, Lillie E, Zarin W, O’Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. *Ann Intern Med*; 2018;169(7):467–73. DOI: 10.7326/M18-0850
21. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. *Syst Rev*; 2021;10(1). DOI: 10.1186/S13643-021-01626-4
22. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping

- reviews. *JBIEvid Synth*. *JBIEvid Synth*; 2020;18(10):2119–26. DOI: 10.11124/JBIES-20-00167
23. Dias MM, Munn Z, Porritt K, Tufanaru C, Stern C, Aromataris E, et al. Use of GRADE in Australian clinical practice guidelines: a methodological review protocol. *JBIDatabase System Rev Implement Rep*. *JBIDatabase System Rev Implement Rep*; 2018;16(11):2092–6. DOI: 10.11124/JBISRIR-2017-003923
24. Hilton Boon M, Thomson H, Shaw B, Akl EA, Lhachimi SK, López-Alcalde J, et al. Challenges in applying the GRADE approach in public health guidelines and systematic reviews: a concept article from the GRADE Public Health Group. *J Clin Epidemiol*. Pergamon; 2021;135:42–53. DOI: 10.1016/J.JCLINEPI.2021.01.001
25. Movsisyan A, Melendez-Torres GJ, Montgomery P. Users identified challenges in applying GRADE to complex interventions and suggested an extension to GRADE. *J Clin Epidemiol*. Pergamon; 2016;70:191–9. DOI: 10.1016/J.JCLINEPI.2015.09.010
26. MINSA. 2015 [Internet]. Elaboración y uso de guías de práctica clínica del Ministerio de Salud - Informes y publicaciones - Ministerio de Salud - Plataforma del Estado Peruano [cited 2024 Mar 31]. Available from: <https://www.gob.pe/institucion/minsa/informes-publicaciones/314117-elaboracion-y-uso-de-guias-de-practica-clinica-del-ministerio-de-salud>
27. Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, et al. GRADE guidelines 26: informative statements to communicate the findings of systematic reviews of interventions. *J Clin Epidemiol*.

- Elsevier USA; 2020;119:126–35. DOI: 10.1016/j.jclinepi.2019.10.014
28. Timaná R, Alva-Díaz C, Suárez V, Pimentel P, Dongo V. Características y calidad de las guías de práctica clínica en el Seguro Social de Salud del Perú. *Medicina de Familia SEMERGEN*. Elsevier Doyma; 2018;44(8):549–56. DOI: 10.1016/J.SEMERG.2018.06.002
 29. Canelo-Aybar C, Balbin G, Perez-Gomez Á, Florez ID. Guías de práctica clínica en el Perú: evaluación de su calidad usando el instrumento AGREE II. *Rev Peru Med Exp Salud Publica*. Instituto Nacional de Salud; 2016;33(4):732–8. DOI: 10.17843/RPMESP.2016.334.2559
 30. Amer YS, Elzalabany MM, Omar TI, Ibrahim AG, Dowidar NL. The “Adapted ADAPTE”: an approach to improve utilization of the ADAPTE guideline adaptation resource toolkit in the Alexandria Center for Evidence-Based Clinical Practice Guidelines. *J Eval Clin Pract*. *J Eval Clin Pract*; 2015;21(6):1095–106. DOI: 10.1111/JEP.12479
 31. Schünemann HJ, Wiercioch W, Brozek J, Etzeandía-Ikobaltzeta I, Mustafa RA, Manja V, et al. GRADE Evidence to Decision (EtD) frameworks for adoption, adaptation, and de novo development of trustworthy recommendations: GRADE-ADOLOPMENT. *J Clin Epidemiol*. Pergamon; 2017;81:101–10. DOI: 10.1016/J.JCLINEPI.2016.09.009

IX. ANEXO

Anexo 01. Dominios y preguntas de la lista de verificación para la adherencia al enfoque GRADE. Adaptado de Días et al.²¹ y Schünemann et al.²⁴.

- **INSTRUCCIONES PARA EL USO DE *GRADE ADHERENCE CHECK-LIST***

El presente manual busca orientar el uso de la lista de verificación *GRADE Adherence Checklist*.

1. PREPARÁNDOSE PARA UTILIZAR LA LISTA DE VERIFICACIÓN

i. PREPARACIÓN DE LOS DOCUMENTOS A EVALUAR

Antes de aplicar la lista de verificación, los usuarios deben revisar la guía de práctica clínica (GPC) que se evaluará, incluyendo la información relevante relacionada con el proceso de desarrollo de la GPC, que puede estar contenida en sus versiones de informe o en su material suplementario disponible en los sitios web de los organismos desarrolladores. De considerarse necesario se debe realizar esfuerzos adicionales para conseguir dicha información.

ii. EVALUADORES

Un revisor con experiencia declarada en el manejo de la metodología GRADE evaluará cada GPC. Si existen dudas respuestas se solicitará la intervención de un segundo revisor.

iii. NÚMERO DE PREGUNTAS CLÍNICAS

Se armará una lista de todas las preguntas clínicas de la GPC que terminen en la

emisión de recomendaciones terapéuticas y de diagnóstico, y se evaluará la totalidad de PICOs incluidas.

2. ESTRUCTURA Y CONTENIDO DEL LISTA DE VERIFICACIÓN

La lista de verificación consta de 15 preguntas agrupadas en 7 dominios. La pregunta 11 incluye 3 sub-preguntas. (Puede ver detalles de esto en el manual GRADE:

<https://gdt.grade.org/app/handbook/translations/es/handbook.html>)

- **Dominio 1: Definición de la certeza de la evidencia:** Hace referencias a los conceptos propuestos por el *GRADE Working Group* (GRADE-WG) para definir la certeza de la evidencia, cómo categorizar su evaluación y cómo redactar las recomendaciones sobre las intervenciones en la GPC (Preguntas 1-3).
- **Dominio 2: Plantear la pregunta de salud:** Hace referencia al uso de la estructura PICO en el proceso de estructuración de las preguntas sobre las intervenciones en la GPC (Pregunta 4).
- **Dominio 3: Seleccionar y clasificar la importancia de los desenlaces:** Hace referencias al proceso donde los desarrolladores eligieron que desenlaces incluir y la priorizaron entre ellos (Preguntas 5-6).
- **Dominio 4: Evaluar la certeza de la evidencia:** Hace referencias al proceso donde los desarrolladores evaluaron cada dominio (riesgo de sesgo, evidencia directa, consistencia y precisión de los resultados, riesgo de sesgo de publicación, magnitud del efecto, gradiente de dosis-respuesta e influencia de confusión plausible residual) para determinar la confianza en

el conjunto de la evidencia identificada para responder a cada pregunta sobre las intervenciones en la GPC (Preguntas 7).

- **Dominio 5: Determinar la certeza de la evidencia:** Hace referencias al proceso donde los desarrolladores asignaron una categoría (alta, moderada, baja o muy baja) para la certeza de la evidencia por cada desenlace (Preguntas 8-9).
- **Dominio 6: Resumir los hallazgos:** Hace referencia al que los desarrolladores emplearon herramientas adecuadas cómo un perfil de evidencia (EP, *Evidence Profile* por sus siglas en inglés) o tablas de resumen de evidencia (SoF, *Summary of Findings por sus siglas en inglés*) para el resumen de sus hallazgos de cada pregunta sobre las intervenciones en la GPC (Preguntas 10-11).
- **Dominio 7: Establecer recomendaciones basadas en evidencias:** Hace referencia al empleo de los criterios GRADE del marco de la evidencia a la decisión (EtD, *Evidence to Decition* por sus siglas en inglés) para determinar la fuerza y dirección de las recomendaciones (12-15).

3. ESCALA DE RESPUESTAS

La respuesta a cada pregunta es dicotómica (Sí/No). Cada pregunta de la lista de verificación contará con criterios específicos que deberán emplearse como apoyo para definir la respuesta.

De manera general se podrá emplear como respuesta las categorías “No reportado” si se cumple lo siguiente:

- No reportado (NR): Si es que no hay reporte explícito en el documento.

4. EVALUACIÓN GLOBAL

El *GRADE ADHERENCE CHECK-LIST* presenta una evaluación cualitativa y descriptiva de los componentes mínimos que debe poseer una GPC para determinar si realizó el proceso de elaboración de recomendaciones empleando el enfoque GRADE. Estos resultados pueden emplearse para comparar dos o más GPC mediante una comparativa porcentual de la adherencia a cada dominio. Sin embargo, no proporciona una interpretación descriptiva y solo se limita a los aspectos generales de la metodología GRADE.

5. LISTA DE VERIFICACIÓN

Dominios y preguntas	Criterios	Respuesta (0=No / 1=Sí / 3=NR)	Justificación de la respuesta (incluyendo en dónde se encontró lo buscado)
D1. Definición de la certeza de la evidencia			
Q1. ¿Se definió el término “certeza de la evidencia” (también conocida como la confianza en los efectos estimados) de acuerdo con las definiciones utilizadas por el GRADE-WG?	<p>Donde buscar:</p> <p>Probablemente Ud. encuentre las definiciones en la sección de glosario de términos, introducción o métodos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores de la GPC han empleado una definición de la certeza de la evidencia acorde o similar a la presentada en el glosario</p>		

	<p>del manual GRADE (capítulo 9):</p> <ul style="list-style-type: none">• Certeza de la evidencia: “es la seguridad que se tiene que el estimado del efecto sea adecuado para soportar una recomendación o que el estimado del efecto está cerca del efecto verdadero (que sea correcto)”. <p>Los términos o expresiones alternativos a "certeza de la evidencia" suelen ser apropiados; las siguientes frases pueden usarse indistintamente: certeza de la evidencia, fuerza de la evidencia, calidad de la evidencia, confianza en las estimaciones (del efecto).</p>		
--	--	--	--

	<p>Cuando asignar No:</p> <p>Empleo de una definición incongruente con la definición planteada por el GRADE-WG</p>		
<p>Q2. ¿Las definiciones para la clasificación de certeza (como alta, moderada, baja y / o muy baja) estuvo de acuerdo con las definiciones utilizadas por el GRADE-WG?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. encuentre las definiciones en la sección de glosario de términos, introducción o métodos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han empleado una definición de la clasificación de la certeza de la evidencia acorde con el glosario del manual GRADE (capítulo 9):</p> <ul style="list-style-type: none"> • Certeza alta: Estamos muy seguros de que el efecto real se aproxima al 		

	<p>de la estimación del efecto.</p> <ul style="list-style-type: none">• Certeza moderada: Tenemos una confianza moderada en la estimación del efecto: es probable que el efecto real se acerque a la estimación del efecto, pero existe la posibilidad de que sea sustancialmente diferente.• Certeza baja: Nuestra confianza en la estimación del efecto es limitada: el efecto real puede ser sustancialmente diferente de la estimación del efecto.• Certeza muy baja: Tenemos muy poca confianza en la		
--	---	--	--

	<p>estimación del efecto: es probable que el efecto real sea sustancialmente diferente de la estimación del efecto.</p> <p>Cuando asignar No:</p> <p>Empleo de una definición incongruente con la definición planteada por el GRADE-WG</p>		
<p>Q3. ¿Las definiciones para redactar cada una de las categorías de la fortaleza (fuerte y débil/condicional) y dirección (a favor o en contra) de la recomendación estuvo de acuerdo con las definiciones utilizadas por el GRADE-WG?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. encuentre las definiciones en la sección de glosario de términos, introducción o métodos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han empleado una definición de la clasificación de la fuerza de la recomendación acorde</p>		

	<p>con el glosario del manual GRADE (capítulo 9):</p> <ul style="list-style-type: none">• Recomendación fuerte: aquella donde los efectos deseables de una intervención superan a sus efectos indeseables (recomendación fuerte a favor para una intervención) o que los efectos indeseables de una intervención superan a sus efectos deseables (recomendación fuerte en contra una intervención).• Recomendación débil/condicional: aquella donde los efectos deseables probablemente superan a los efectos indeseables (recomendación débil a favor para una intervención) o los		
--	--	--	--

	<p>efectos indeseables probablemente superen los efectos deseables (recomendación débil en contra una intervención) pero existe una incertidumbre apreciable.</p> <p>Cuando asignar No:</p> <p>Empleo de una definición incongruente con la definición planteada por el GRADE-WG</p>		
D2. Plantear la pregunta de salud			

<p>Q4. ¿Se formuló cada pregunta de salud de acuerdo con la estructura PICO (paciente/población, intervención, comparación, desenlaces) o un equivalente?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre las especificaciones en la sección de métodos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han planteado la pregunta de salud incluyendo de manera explícita cada componente de la PICO, acorde con el capítulo 2 del manual GRADE.:</p> <ul style="list-style-type: none"> • Población: a población a quienes se supone que se aplican las recomendaciones • Intervención: la intervención terapéutica, 		
---	--	--	--

	<p>diagnóstica o de otro tipo;</p> <ul style="list-style-type: none"> • Comparador: la intervención alternativa, la intervención en el grupo control • Desenlace: el resultado o los resultados de interés. <p>debe considerarse satisfactorio el empleo de una pregunta narrativa estructurada para cada recomendación como:</p> <p><i>¿Se debe usar [intervención] versus [comparación] para [problema de salud]?, ¿Se debe usar [intervención] versus [comparación] en [población]?,</i></p> <p>cuando se presume que el comparador es “no brindar la</p>		
--	--	--	--

	<p>intervención” podría aceptarse la siguiente estructura:</p> <p><i>Se debe usar [intervención] para [problema de salud]?</i></p> <p><i>¿Se debe usar [intervención] en [población]?</i></p> <p>Cuando asignar No:</p> <p>o se empleó la estructura PICO o se empleó una estructura diferente.</p>		
<p>D3. Seleccionar y clasificar la importancia de los desenlaces</p>			
<p>Q5. ¿Para la toma de decisiones, se clasificaron los desenlaces por cada pregunta cómo no importante, o críticos para los pacientes?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre la argumentación en la sección de métodos, tabla de resumen de hallazgos o anexos de la GPC.</p> <p>Cuando asignar Sí:</p>		

	<p>Los desarrolladores han presentado de manera explícita y transparente el proceso de clasificación de la totalidad de desenlaces por cada pregunta, en tres categorías de acuerdo con su importancia para la toma de decisiones. Estos resultados se clasifican numéricamente en una escala del 1 a 9, acorde con el capítulo 3 del manual GRADE:</p> <ul style="list-style-type: none">• Crítico (7 a 9 en la escala)• Importante pero no crítico (4 a 6 en la escala)• Importancia limitada o “no importante” (1 a 3 en la escala). <p>Cuando asignar No:</p>		
--	--	--	--

	<p>No priorización de desenlaces o empleo de una terminología diferente a la planteada por el GRADE-WG.</p>		
<p>Q6. ¿La gradación de la certeza se realizó en base a cada desenlace?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. encuentre la argumentación en la tabla de resumen de hallazgos o anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han gradado la certeza por cada uno de los desenlaces seleccionados acorde con el capítulo 6 del manual GRADE.</p> <p>Cuando asignar No:</p>		

	La gradación de la certeza no se realizó por cada desenlace.		
D4. Evaluar la certeza de la evidencia			
<p>Q7. ¿Se consideraron todos los aspectos que contribuyeron al enfoque GRADE para evaluar la certeza de la evidencia para cada desenlace?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de métodos y en la tabla SoF.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores de GPC han declarado que evaluaron la certeza de la evidencia considerando todos los dominios propuestos en el capítulo 5 del manual GRADE:</p> <ul style="list-style-type: none"> • Diseño del estudio • Riesgo de sesgo • Inconsistencia (Heterogeneidad) • Evidencia indirecta • Imprecisión • Sesgo de publicación 		

	<ul style="list-style-type: none"> • Gran magnitud de efecto • Todo factor de confusión plausible reduciría el efecto demostrado o aumentaría el efecto si no se observara ningún efecto. • Gradiente dosis-respuesta <p>Cuando asignar No:</p> <p>No declaración explícita del uso de los dominios propuestos por el GRADE-WG.</p>		
D5. Determinar la certeza de la evidencia			
<p>Q8. ¿Se evaluó la certeza de la evidencia por cada desenlace expresado mediante las categorías que</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en el EP, la tabla SoF o anexos de la GPC.</p> <p>Cuando asignar Sí:</p>		

<p>propone GRADE, en consideración de los factores evaluados en la Q7?</p>	<p>Cuando se haya respondido “Sí” a las preguntas Q6 y Q7, y además se muestre que para cada desenlace se haya evaluado la certeza y consignado alguna de las 4 categorías que propone GRADE: <i>Alto, Moderado, Bajo, o Muy Bajo</i></p> <p>Cuando asignar No:</p> <p>No descripción acorde a la terminología propuesta por el GRADE-WG o empleo otro sistema de categorización.</p>		
<p>Q9. ¿Se calificó la “certeza/calidad general de la evidencia”?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de la tabla EtD o anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han calificado la “certeza/calidad</p>		

	<p>general” como la calificación combinada de la certeza de la evidencia para responder la pregunta de investigación, considerando todos los desenlaces calificados como críticos (aquellos que están impulsando una recomendación). Para ello, se considera la menor certeza encontrada en los desenlaces críticos.</p> <p>Cuando asignar No:</p> <p>No descripción detallada del proceso de calificación de la certeza/calidad general.</p>		
D6. Resumir los hallazgos			
<p>Q10. ¿Se resumió los resultados usando tablas SoF o el EP?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en el desarrollo de las preguntas o los anexos de la GPC.</p>		

	<p>Cuando asignar Sí:</p> <p>Los desarrolladores han presentado los resultados en tablas SoF o EP.</p> <p>Esto debe incluir los juicios que sustentan la calificación de la calidad y los efectos de las alternativas de manejo en los desenlaces de interés, acorde con el capítulo 4 del manual GRADE.</p> <p>Cuando asignar No:</p> <p>No se realizan tablas SoF ni EP.</p>		
<p>Q11. ¿Se proporcionaron todos los detalles requeridos por GRADE en el perfil de evidencia, resumen de evidencia o SoF table?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en el desarrollo de las preguntas o los anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Las tablas SoF o EP incluye en su formato/estructura los siguientes detalles:</p>		

	<ul style="list-style-type: none">• Número y tipo de estudios incluidos.• Certeza de la evidencia de acuerdo con los dominios GRADE (listados en la Q7).• Resultados estadísticos clave. Esto debe incluir medidas de efecto para desenlaces absolutos (como la diferencia de riesgos o el NNT). En algunos casos esto puede presentarse de manera narrativa.• Categorización de desenlaces. <p>Para mayor detalle revisar el capítulo 4 del manual GRADE</p>		
--	--	--	--

	<p>Cuando asignar No:</p> <p>No hubo una descripción con los detalles requeridos en las tablas SoF ni EP.</p>		
<p>Q11-a. ¿El EP o tabla SoF se basó en revisiones sistemáticas?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de métodos, tabla SoF, el EP o anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Cada una de las filas de la tabla SoF o EP corresponde a una revisión sistemática. Si bien puede que la revisión haya incluido un solo estudio o ningún estudio.</p> <p>Cuando asignar No:</p> <p>No se basó en revisión sistemática.</p>		
<p>Q11-b. ¿Se describió claramente el tipo</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de</p>		

<p>de evidencia que se evaluó?</p>	<p>métodos, tabla SoF, el EP o anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han descrito claramente el tipo de evidencia (experimental u observacional) incluida para la evaluación de la certeza por cada desenlace acorde con el capítulo 4 del manual GRADE.</p> <p>Cuando asignar No:</p> <p>No se describió el tipo de evidencia evaluada.</p>		
<p>Q11-c. ¿Se describieron de manera transparente las razones para disminuir o incrementar la certeza de la evidencia?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de métodos, tabla SoF, el EP o anexos de la GPC.</p> <p>Cuando asignar Sí:</p>		

	<p>Los desarrolladores han presentado un resumen detallado de los argumentos por los que se disminuyó o aumentó el nivel de la certeza de evidencia, acorde con el capítulo 4 del manual GRADE</p> <p>Cuando asignar No:</p> <p>No presentaron los argumentos.</p>		
<p>D7. Establecer recomendaciones basadas en evidencia</p>			
<p>Q12. ¿Se consideraron los criterios GRADE para determinar la fuerza y dirección de una recomendación)?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de métodos, tabla EtD, o texto en el que explicita en qué se basó la recomendación.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han considerado de manera explícita cada uno de los</p>		

	<p>criterios del proceso para determinar la fuerza y dirección de la recomendación, acorde con el capítulo 6.2 al 6.3 del manual GRADE y especificando la perspectiva empleada al desarrollar la tabla EtD.</p> <ul style="list-style-type: none">• Para la perspectiva individual (Por ejemplo, desde la perspectiva del paciente o una institución en específico) se emplea la mayoría de los criterios:<ul style="list-style-type: none">○ Equilibrio de consecuencia s deseables e indeseables○ Certeza de la evidencia		
--	---	--	--

	<ul style="list-style-type: none">○ Valores- preferencias de los afectados (pacientes)○ Equidad○ Aceptabilidad○ Factibilidad <ul style="list-style-type: none">• Para la perspectiva poblacional o ampliada (Por ejemplo, desde la perspectiva del sistema de salud) adicionalmente se debe discutir quién asumirá los costos de la adquisición e implementación de la intervención evaluad. Por tanto, se deberá incorporar el criterio:		
--	---	--	--

	<p style="text-align: center;">○ Uso de recursos (Costos)</p> <p>Para recomendaciones diagnósticas se considerará adicionalmente los criterios: (disponibles en https://gdt.gradepro.org/)</p> <ul style="list-style-type: none"> • Precisión de la prueba • Efectos deseables/indeseables de la prueba • Certeza de evidencia de la exactitud de la prueba • Certeza de evidencia de los efectos de la prueba • Certeza de evidencia de los efectos del manejo • Certeza de evidencia de la exactitud del 		
--	--	--	--

	<p>resultado/manejo clínico de la prueba</p> <p>Cuando asignar No:</p> <p>No se consideran los criterios GRADE para determinar la fuerza de la recomendación.</p>		
<p>Q13. ¿La fuerza de las recomendaciones (a favor o en contra) se expresó usando dos categorías (fuerte o débil/condicional)?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la tabla de resumen de recomendaciones.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han presentado la fuerza de recomendación como fuerte o débil/condicional acorde con el capítulo 6.1 del manual GRADE.</p> <p>Ud. podría ver detalles de estas definiciones en los criterios De la pregunta Q3.</p>		

	<p>Cuando asignar No:</p> <p>No se categoriza en <i>fuerte</i> o <i>débil/condicional</i>.</p>		
<p>Q14. ¿Se utilizó el marco EtD en el proceso argumentativo del desarrollo de recomendaciones?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la tabla de resumen de recomendaciones, en la sección “de la evidencia a la recomendación”, en el desarrollo de cada pregunta, sección EtD o anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>En la formulación de las recomendaciones, los desarrolladores han presentado de manera</p>		

	<p>explícita los argumentos para cada criterio del marco EtD (ya sea en la sección EtD, o con párrafos de manera narrativa), acorde con el capítulo 6.5 del manual GRADE.</p> <p>Se entiende como argumentativo al proceso de presentar opinión y/o información para discutir cada uno de los criterios del marco EtD.</p> <p>Cuando asignar No:</p> <p>No se presenta la tabla EtD o no se declaran los argumentos en las secciones previamente sugeridas.</p>		
--	--	--	--

<p>Q15. ¿Se reportaron de forma transparente las razones para deliberar la fuerza y dirección de las recomendaciones?</p>	<p>Donde buscar:</p> <p>Probablemente Ud. lo encuentre en la sección de matriz de resultados de votación en la sección EtD, en la sección “de la evidencia a la recomendación”, en el desarrollo de cada pregunta, en la “dirección y fuerza de la recomendación” o los anexos de la GPC.</p> <p>Cuando asignar Sí:</p> <p>Los desarrolladores han reportado de manera explícita los determinantes para deliberar la fuerza (fuerte o débil/condicional) y dirección (a favor o en contra) de la recomendación). Para más información Ud. puede revisar el capítulo 6.5 del manual GRADE.</p>		
---	---	--	--

	<p>Se entiende como deliberativo al proceso de seleccionar una respuesta a cada una de las preguntas del marco EtD con la finalidad de señalar la fuerza y dirección de la recomendación.</p> <p>Cuando asignar No:</p> <p>No se reportaron las razones determinantes.</p>		
--	---	--	--

Anexo 02. Estrategias de búsqueda. Filtro desde 01 de enero del 2018 a 31 de diciembre del 2022.

Bases de datos	Estrategias de búsqueda	Resultados
MEDLINE through Pubmed	((guideline [MeSH Terms]) OR (guideline*[title])) AND ((Peru [MeSH Terms]) OR (Peru*[title/abstract]))	35
SCOPUS	((INDEXTERMS("guideline ")) OR (TITLE("guideline*"))) AND ((INDEXTERMS("Peru ")) OR (TITLE-ABS("Peru*")))	228
Embase	((exp guideline /) OR (guideline*.ti.)) AND ((exp Peru /) OR (Peru*.ti,ab.))	311
Scielo	(ti:(guía de práctica clínica)) AND (ab:(Perú))	125
LILACS	(MH "Clinical Practice Guideline") AND (Perú)	47

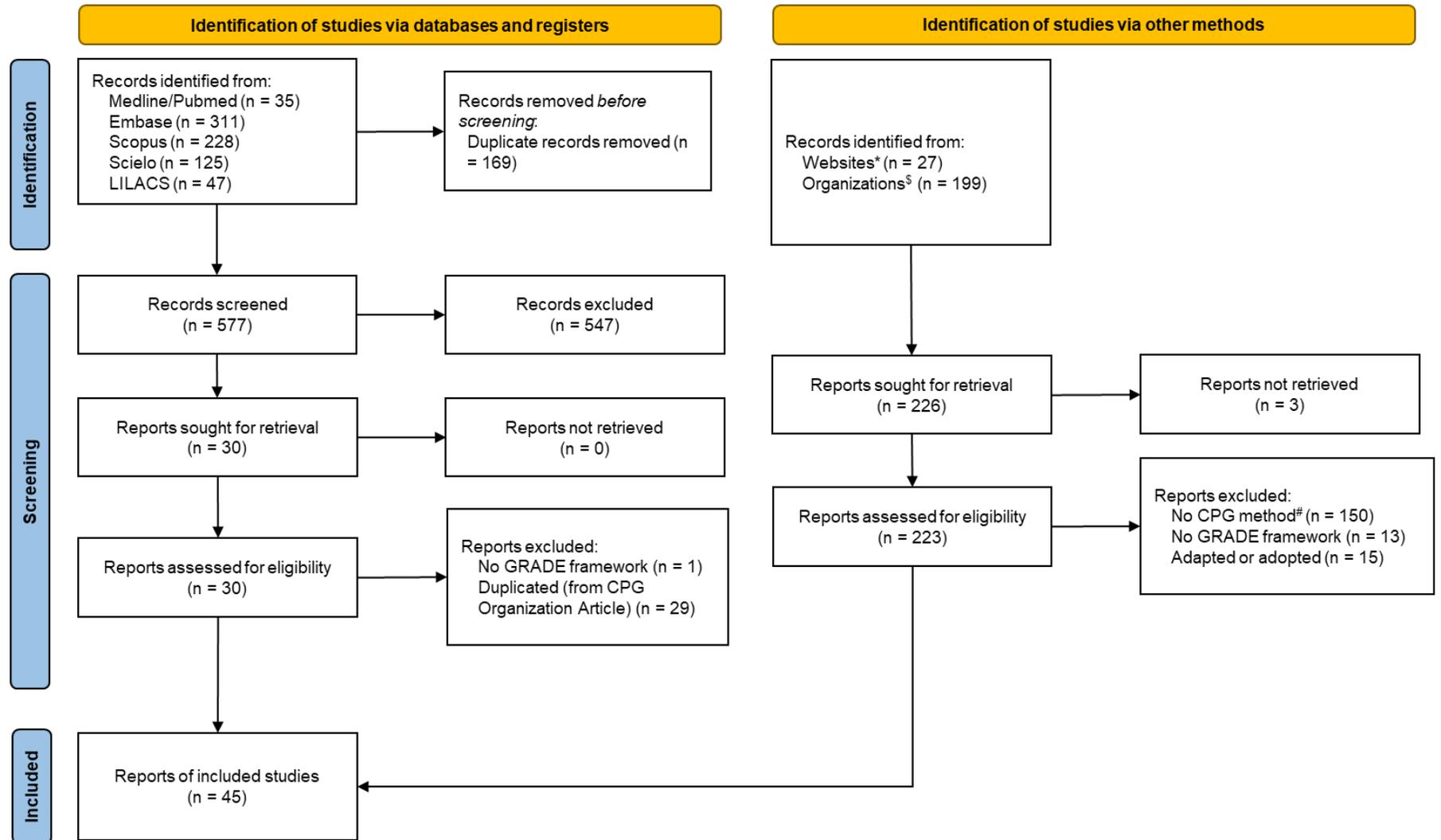
Anexo 03. Expertos metodólogos en el campo de la elaboración de GPC quienes participaron de validación de cara de la lista de verificación *GRADE ADHERENCE CHECK-LIST*.

Este trabajo se desarrolló entre 2021 y 2022, y se realizaron reuniones tanto virtuales como presenciales. El consenso se logró a través de discusiones colaborativas entre los participantes. En caso de discrepancias, se buscó resolverlas mediante consenso, y si no se alcanzaba, se procedió a una votación por mayoría simple para tomar las decisiones. Es importante señalar que ninguno de los participantes solicitó manifestar por escrito una opinión contraria a las decisiones tomadas durante este proceso.

Nombres	Filiación institucional	ORCID
Carlos Antonio Vicente Salazar Ordoñez	<ul style="list-style-type: none"> • Departamento de medicina, Universidad Peruana Cayetano Heredia, UPCH, Lima, Perú. • Maestro en Epidemiología clínica, UPCH, Lima, Perú. 	0000-0001-9959-8097.
Carlos Alva-Díaz	<ul style="list-style-type: none"> • Red de Eficacia Clínica y Sanitaria, REDECS, Lima, Perú • Grupo de Investigación en Neurociencia, Efectividad Clínica y Salud Pública, Universidad Científica del Sur, Lima, Perú. • Maestro en Epidemiología clínica, UPCH, Lima, Perú. 	0000-0003-3584-7298

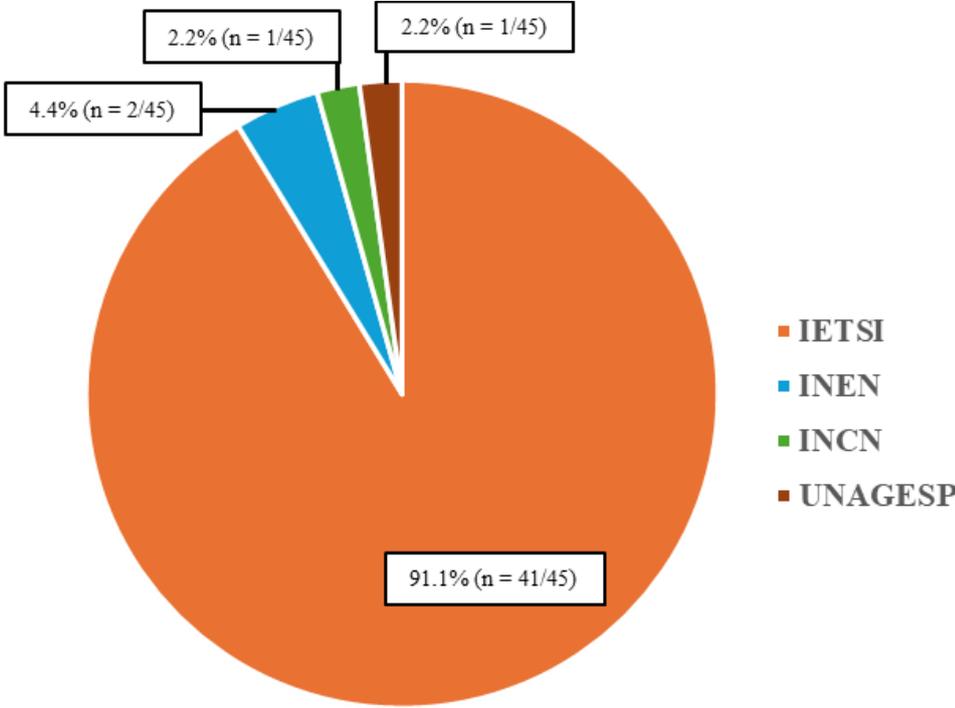
<p>Alvaro Taype- Rondan</p>	<ul style="list-style-type: none"> • Instituto de evaluación de tecnologías en salud e investigación, IETSI, Essalud, Lima, Perú. • Maestro en ciencias de la investigación epidemiológica, UPCH, Lima, Perú. 	<p>0000-0001- 8758-0463</p>
<p>Wendy Nieto- Gutiérrez</p>	<ul style="list-style-type: none"> • Instituto de evaluación de tecnologías en salud e investigación, IETSI, Essalud, Lima, Perú. • Maestra en ciencias de la investigación epidemiológica, UPCH, Lima, Perú. 	<p>0000-0001- 8012-1520</p>
<p>Fradis Gil- Olivares</p>	<ul style="list-style-type: none"> • Red de Eficacia Clínica y Sanitaria, REDECS, Lima, Perú. • Unidad de Guías de Práctica Clínica, AUNA, Lima, Perú. • Maestro en Epidemiología clínica y bioestadística, UCSUR, Lima, Perú. 	<p>0000-0002- 9231-6023</p>

Anexo 04. Flujograma PRISMA para la selección de estudios.



*Incluye la plataforma BIGG (Base de datos internacional de directrices GRADE) y la base de datos Epistemonikos. \$ Incluye las siguientes organizaciones peruanas IETSI (Instituto de Evaluación de Tecnologías en Salud e Investigación), INSN (Instituto Nacional de Salud del Niño), INMP (Instituto Nacional Materno Perinatal), INCN (Instituto Nacional de Ciencias Neurológicas), MINSA-INS (Ministerio de Salud - Instituto Nacional de Salud) e INEN (Instituto Nacional de Enfermedades Neoplásicas). # Según el IOM², no cumple con los estándares de una GPC si no realiza una búsqueda sistemática de evidencia, no se adhiere a una metodología de toma de decisiones y/o no emite recomendaciones basadas en evidencia.

Anexo 05. Número de GPC según institución peruana que declararon uso de la metodología GRADE entre 2018-2022 (n = 45).



Anexo 06. Número de GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022, según categorías de grupos de edad (n = 45). Se reportan las categorías que emplearon por los autores en cada GPC.

Grupo etario	Frecuencia absoluta	Porcentaje (%)
> 18 años	28	62,2
No distinción de edad	12	26,7
> 50 años	1	2,2
> 16 años	1	2,2
< 32 años	1	2,2
> 15 años	1	2,2
> 14 años	1	2,2

Anexo 07. Número de GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022, según especialidad médica (n = 45). Se reportan las categorías que emplearon por los autores en cada GPC.

Especialidad	Frecuencia absoluta	Porcentaje (%)
Oncología	7	15,5
Cardiología	6	13,3
Gastroenterología	4	8,9
Nefrología	4	8,9
Oftalmología	3	6,7
Neurología	2	4,4
Urología	2	4,4
Obstetricia	2	4,4
Reumatología	2	4,4
Cirugía general	2	4,4
Odontología	2	4,4
Neonatología	1	2,2
Ginecología	1	2,2
Pediatría	1	2,2
Enfermedades infecciosas	1	2,2
Nutrición	1	2,2
Hematología	1	2,2

Psiquiatría	1	2,2
Medicina interna	1	2,2

Anexo 08. Número de GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022, según tópico (n = 45). Se reportan las categorías que emplearon por los autores en cada GPC.

Tópico	Frecuencia absoluta	Porcentaje (%)
Manejo	16	35,5
Diagnóstico y tratamiento	9	20
Diagnóstico y manejo	5	11,1
Tratamiento	3	6,7
Evaluación inicial y manejo	2	4,4
Prevención y manejo	1	4,35
Detección y manejo	1	2,2
Detección, diagnóstico y tratamiento	1	2,2
Diagnóstico, manejo y seguimiento	1	2,2
Prevención, diagnóstico y manejo	1	2,2
Prevención y tratamiento	1	2,2
Prevención y manejo	1	2,2
Detección, diagnóstico y tratamiento	1	2,2
Detección	1	2,2

Prevención, diagnóstico y tratamiento	1	2,2
--	---	-----

Anexo 09. Mediana del número de preguntas PICO de las GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022.

Descripción	Mediana	Rango	Rango intercuartílico
Número total de preguntas PICO	7	11 (1-12)	3
Preguntas PICO que desarrollan tablas de perfil de evidencia o SoF	6	11 (0-11)	3
Preguntas PICO que desarrollan parcialmente marcos EtD	0	11 (0-11)	6
Preguntas PICO que desarrollan completamente los marcos EtD	4	10 (0-10)	6

Anexo 10. Resultados individuales de la adherencia a GRADE y de la calidad metodológica utilizando el dominio 3 de la herramienta AGREE II para 45 GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022.

N	GPC ID #	Preguntas																		Adherencia a GRADE	Calidad metodológica (AGREE II - Dominio 3)	
		Definición de certeza de la evidencia			Planteamiento de la pregunta clínica	Selección y clasificación la importancia de los resultados		Evaluación de la Certeza de la evidencia	Determinación de la Certeza de la evidencia		Resumen de los hallazgos					Establecer recomendaciones basadas en evidencia						
		Q1	Q2	Q3		Q4	Q5		Q6	Q7	Q8	Q9	Q10	Q11	Q11 a	Q11 b	Q11 c	Q12	Q13			Q14
1	G1	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	NO	88.9%	75%	
2	G2	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	94.4%	72.9%	
3	G3	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	75%	
4	G4	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	NO	88.9%	79.2%	
5	G5	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	94.4%	79.2%	
6	G8	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	79.2%	
7	G9	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	94.4%	81.3%	
8	G10	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	87.5%	
9	G11	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	81.3%	
10	G12	YES	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	NO	YES	88.9%	81.3%	
11	G13	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	81.3%	
12	G14	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	83.3%	
13	G16	YES	YES	YES	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	NO	NO	83.3%	81.3%	
14	G17	YES	YES	NO	YES	YES	YES	YES	YES	NO	YES	YES	YES	YES	YES	YES	YES	YES	NO	83.3%	70.8%	
15	G18	NO	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	94.4%	79.2%	

38	G56	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	72.9%	
39	G57	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	75%
40	G58	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	72.9%
41	G59	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	72.9%
42	G60	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	72.9%
43	G61	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	75%
44	G62	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	72.9%
45	G63	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	100%	75%
		77.8%	93.3%	86.7%	100%	97.8%	100%	95.6%	97.8%	71.1%	100%	100%	100%	97.8%	100%	100%	100%	100%	84.4%	80.0%			

Q1. ¿Se definió el término "certeza de la evidencia" (también conocido como confianza en los efectos estimados) de acuerdo con las

definiciones utilizadas por el Grupo de Trabajo GRADE (GRADE-WG)?

Q2. ¿Estuvieron las definiciones para clasificar la certeza (como alta, moderada, baja y/o muy baja) en consonancia con las

definiciones utilizadas por el GRADE-WG?

Q3. ¿Estuvieron las definiciones para redactar cada una de las categorías de fuerza (fuerte y débil/condicional) y dirección (a favor o

en contra) de la recomendación de acuerdo con las definiciones utilizadas por el GRADE-WG?

Q4. ¿Se formuló cada pregunta de salud según la estructura PICO (paciente/población, intervención, comparación, resultados) o

equivalente?

Q5. Para la toma de decisiones, ¿se clasificaron los resultados de cada pregunta cómo no importantes, importantes o críticos para los

pacientes?

Q6. ¿Se realizó la clasificación de la certeza en base a cada resultado?

Q7. ¿Se consideraron todos los aspectos que contribuyen al enfoque GRADE al evaluar la certeza de la evidencia para cada resultado?

Q8. ¿Se evaluó la certeza de la evidencia para cada resultado utilizando las categorías propuestas por GRADE, considerando los factores evaluados en Q7?

Q9. ¿Se clasificó la "certeza/calidad general de la evidencia"?

Q10. ¿Se resumieron los resultados utilizando tablas de Resumen de Hallazgos (SoF) o perfiles de evidencia (EP)?

Q11. ¿Se proporcionaron todos los detalles requeridos por GRADE en el perfil de evidencia, resumen de evidencia o tabla SoF?

Q11-a. ¿Se basó el EP o la tabla SoF en revisiones sistemáticas?

Q11-b. ¿Se describió claramente el tipo de evidencia evaluada?

Q11-c. ¿Se describieron transparentemente las razones para disminuir o aumentar la certeza de la evidencia?

Q12. ¿Se consideraron los criterios GRADE para determinar la fuerza y dirección de una recomendación?

Q13. ¿Se expresó la fuerza de las recomendaciones (a favor o en contra) utilizando dos categorías (fuerte o débil/condicional)?

Q14. ¿Se utilizó el marco de Evidencia para la Decisión (EtD) en el proceso argumentativo de desarrollo de las recomendaciones?

Q15. ¿Se informaron transparentemente las razones para deliberar sobre la fuerza y dirección de las recomendaciones?

Anexo 11. Evaluación de la calidad metodológica con el instrumento AGREE-II

Dominio 3 de las 45 GPC peruanas que declararon uso de la metodología GRADE entre 2018-2022.

N	GPC ID#	P7	P8	P9	P10	P11	P12	P13	P14	Total	%AGREE-II Dominio 3
1	G1	7	5	3	7	6	6	5	5	44	75.00%
2	G2	6	5	3	7	6	6	5	5	43	72.92%
3	G3	6	5	4	7	7	5	5	5	44	75.00%
4	G4	7	5	4	7	6	6	6	5	46	79.17%
5	G5	7	5	4	7	6	6	6	5	46	79.17%
6	G8	7	5	4	7	6	6	6	5	46	79.17%
7	G9	7	5	5	7	6	6	6	5	47	81.25%
8	G10	5	4	7	7	7	7	7	6	50	87.50%
9	G11	7	6	4	7	6	6	6	5	47	81.25%
10	G12	7	6	4	7	6	6	6	5	47	81.25%
11	G13	7	6	4	7	6	6	6	5	47	81.25%
12	G14	7	6	5	7	6	6	6	5	48	83.33%
13	G16	7	6	4	7	6	6	6	5	47	81.25%
14	G17	6	5	4	5	6	6	5	5	42	70.83%
15	G18	7	5	4	5	6	6	6	7	46	79.17%
16	G19	6	5	4	5	6	6	5	5	42	70.83%
17	G21	6	5	4	5	6	6	5	5	42	70.83%
18	G22	7	5	5	6	5	6	5	6	45	77.08%
19	G23	7	6	5	7	7	6	5	5	48	83.33%
20	G24	6	5	5	6	7	6	5	5	45	77.08%
21	G25	7	5	4	6	5	5	6	4	42	70.83%
22	G27	6	5	4	5	5	5	5	5	40	66.67%
23	G29	7	6	4	6	6	5	6	5	45	77.08%
24	G30	7	6	4	6	6	5	6	5	45	77.08%
25	G31	7	6	4	6	6	6	5	5	45	77.08%
26	G32	6	5	5	5	6	5	5	5	42	70.83%
27	G33	7	6	4	5	6	5	6	5	44	75.00%
28	G34	6	5	4	5	6	6	5	5	42	70.83%
29	G35	6	6	4	6	6	5	6	5	44	75.00%
30	G36	6	5	4	5	6	6	5	5	42	70.83%
31	G37	7	6	4	6	6	5	6	5	45	77.08%
32	G39	6	5	4	5	6	6	5	5	42	70.83%
33	G49	6	6	4	6	6	5	5	6	44	75.00%
34	G51	6	5	5	5	6	5	6	7	45	77.08%
35	G53	6	5	5	5	6	5	6	7	45	77.08%
36	G54	6	5	4	5	6	6	5	6	43	72.92%
37	G55	7	5	4	5	6	6	5	6	44	75.00%
38	G56	6	5	4	5	6	6	5	6	43	72.92%

39	G57	7	5	4	5	6	6	5	6	44	75.00%
40	G58	6	5	4	5	6	6	5	6	43	72.92%
41	G59	6	5	4	5	6	6	5	6	43	72.92%
42	G60	6	5	4	5	6	6	5	6	43	72.92%
43	G61	7	5	4	5	6	6	5	6	44	75.00%
44	G62	6	5	4	5	6	6	5	6	43	72.92%
45	G63	6	5	4	5	6	6	6	6	44	75.00%

P7. Se utilizaron métodos sistemáticos para buscar evidencia.

P8. Los criterios para seleccionar la evidencia están claramente descritos.

P9. Las fortalezas y limitaciones del cuerpo de evidencia están claramente descritas.

P10. Los métodos para formular las recomendaciones están claramente descritos.

P11. Se han considerado los beneficios y riesgos/efectos secundarios potenciales al formular las recomendaciones.

P12. Existe un vínculo explícito entre las recomendaciones y la evidencia de apoyo.

P13. La guía ha sido revisada externamente por expertos antes de su publicación.

P14. Se proporciona un procedimiento para actualizar la guía.

Anexo 12. Preguntas de la lista de verificación para valorar la rigurosidad del desarrollo de las GPC según el dominio 3 de AGREE II.

- **INSTRUCCIONES PARA EL USO DE AGREE-II Dominio 3: Rigor de la elaboración**

- El presente dominio está compuesto por las siguientes indicadores:

P7. Se utilizaron métodos sistemáticos para buscar evidencia: Este ítem evalúa si se han implementado procedimientos estructurados y reproducibles para la búsqueda de literatura científica y estudios relevantes, con el fin de garantizar una recopilación exhaustiva y objetiva de la evidencia.

- P8. Los criterios para seleccionar la evidencia están claramente descritos: Este ítem revisa si se especificaron de manera clara y transparente los criterios que se utilizaron para incluir o excluir estudios o fuentes de información, facilitando así la comprensión y la replicación del proceso de selección.

- P9. Las fortalezas y limitaciones del cuerpo de evidencia están claramente descritas: Evalúa si se proporciona una evaluación completa y balanceada de la calidad de la evidencia, señalando tanto los aspectos sólidos como las limitaciones (por ejemplo, sesgos, variabilidad, tamaño muestral).

- P10. Los métodos para formular las recomendaciones están claramente

descritos: Este ítem revisa si se explica con detalle el proceso mediante el cual se desarrollaron las recomendaciones, incluyendo las técnicas empleadas para la toma de decisiones, como el consenso o el uso de algoritmos específicos.

- P11. Se han considerado los beneficios y riesgos/efectos secundarios potenciales al formular las recomendaciones: Aquí se evalúa si las recomendaciones emitidas han tenido en cuenta tanto los posibles beneficios como los riesgos o efectos secundarios de las intervenciones o tratamientos sugeridos.
- P12. Existe un vínculo explícito entre las recomendaciones y la evidencia de apoyo: Este ítem verifica si se ha establecido de manera clara la relación directa entre las recomendaciones propuestas y la evidencia científica que las respalda, ofreciendo transparencia en el razonamiento detrás de cada recomendación.
- P13. La guía ha sido revisada externamente por expertos antes de su publicación: Este ítem revisa si la guía ha sido sometida a una revisión por pares o expertos externos que no participaron en su elaboración, para asegurar la validez y la solidez de sus recomendaciones.
- P14. Se proporciona un procedimiento para actualizar la guía: Evalúa si la guía incluye un plan o mecanismo formal para su revisión y actualización

periódica, lo que garantiza que las recomendaciones sigan siendo relevantes y basadas en la evidencia más reciente.

El presente instrumento está graduado mediante una escala de 7 puntos (desde el 1 «Muy en desacuerdo» hasta el 7 «Muy de acuerdo»).

- Puntuación 1 («Muy en desacuerdo»). La puntuación 1 debe darse si no hay información relevante al ítem del AGREE II o si el concepto se presenta de forma muy vaga.
- Puntuación 7 («Muy de acuerdo»). La puntuación 7 debe darse si la calidad de la información es excepcional y cuando todos los criterios y consideraciones definidos en el Manual del Usuario (https://www.agreetrust.org/wp-content/uploads/2013/06/AGREE_II_Spanish.pdf) se han alcanzado por completo.
- Puntuación entre 2 y 6. La puntuación entre 2 y 6 se asigna cuando la información respecto al ítem del AGREE II no cumple por completo con todos los criterios o consideraciones. La puntuación se asignará dependiendo del grado de cumplimiento o calidad de la información.

La puntuación total del dominio se calcula sumando todos los puntos de los ítems individuales del dominio y estandarizando el total, como un porcentaje sobre la máxima puntuación posible para ese dominio. Por ejemplo, si 4 evaluadores dan las siguientes puntuaciones para un dominio con 3 ítems la calificación se realizará como sigue la siguiente figura:

	Ítem 1	Ítem 2	Ítem 3	Total
Evaluador 1	5	6	6	17
Evaluador 2	6	6	7	19
Evaluador 3	2	4	3	9
Evaluador 4	3	3	2	8
TOTAL	16	19	18	53

Puntuación máxima posible = 7 (Muy de acuerdo) x 3 (ítems) x 4 (evaluadores) = 84
Puntuación mínima posible = 1 (Muy en desacuerdo) x 3 (ítems) x 4 (evaluadores) = 12
Puntuación estandarizada del dominio sería

$$\frac{\text{Puntuación obtenida} - \text{Puntuación mínima posible}}{\text{Puntuación máxima posible} - \text{Puntuación mínima posible}}$$

$$53 - 12 / 84 - 12 \times 100 = 41 / 72 \times 100 = 0,5694 \times 100 = 57 \%$$

Si se excluyen ítems, debe modificarse adecuadamente los cálculos de la puntuación máxima y mínima posible del dominio.