



La salud
es de todos

Minsalud

REVISION SISTEMÁTICA DE LA LITERATURA

EXACTITUD DIAGNOSTICA DE LAS PRUEBAS PARA DETECCIÓN DE LA INFECCIÓN CAUSADA POR SARS-CoV-2

ENERO DE 2021



El Instituto de Evaluación Tecnológica en Salud – IETS, es una corporación sin ánimo de lucro, de participación mixta y de carácter privado, con patrimonio propio, creado según lo estipulado en la Ley 1438 de 2011. Su misión es contribuir al desarrollo de mejores políticas públicas y prácticas asistenciales en salud, mediante la producción de información basada en evidencia, a través de la evaluación de tecnologías en salud y guías de práctica clínica, con rigor técnico, independencia y participación. Sus miembros son el Ministerio de Salud y Protección Social – MinSalud, el Ministerio de Ciencia, Tecnología e Innovación – Minciencias, el Instituto Nacional de Vigilancia de Medicamentos y Alimentos – INVIMA, el Instituto Nacional de Salud – INS, la Asociación Colombiana de Facultades de Medicina – ASCOFAME y la Asociación Colombiana de Sociedades Científicas – ACSC.

Autores

Herrera Torres, Ana Milena. Médica y cirujana general. Magíster en epidemiología. Ph.D en patología. Instituto de Evaluación Tecnológica en Salud – IETS.

Cortes Palacio, Katherinne. Médico, ESpecialista en epidemiología clínica. Instituto de Evaluación Tecnológica en Salud – IETS.

Estrada Orozco, Kelly Patricia. Médico, MSc en epidemiología clínica, MSc en biología del comportamiento, PhD en Salud Pública (c). Instituto de Evaluación Tecnológica en Salud – IETS.

Entidad que solicita la evaluación

La revisión sistemática se realiza por solicitud del Ministerio de Salud y Protección Social, en el marco de la pandemia por SARS-CoV-2. Fondo Nacional de Gestión del Riesgo de Desastres mediante Contrato No. 9677-MECOV19-1009-2020.

Conflictos de interés

Los autores declaran, bajo la metodología establecida por el Instituto de Evaluación Tecnológica en Salud - IETS, que no existe ningún conflicto de interés invalidante de tipo financiero, intelectual, de pertenencia o familiar que pueda afectar el desarrollo de este informe.

Declaración de independencia editorial

El desarrollo del estudio, así como sus conclusiones, se realizan de manera independiente, transparente e imparcial por parte de los autores.

Derechos de autor

Los derechos de propiedad intelectual del contenido de este documento son de propiedad del Instituto de Evaluación Tecnológica en Salud. Lo anterior, sin perjuicio de los derechos morales y las citas y referencias bibliográficas enunciadas. En consecuencia, constituirá violación a la normativa aplicable a los derechos de autor, y acarreará las sanciones civiles, comerciales y penales a que haya lugar, su



modificación, copia, reproducción, fijación, transmisión, divulgación, publicación o similares, parcial o total, o el uso del contenido de este sin importar su propósito, sin que medie el consentimiento expreso y escrito del Instituto de Evaluación Tecnológica en Salud.

Citación

Herrera AM, Cortes K, Estrada-Orozco K. Instituto de Evaluación Tecnológica en Salud – IETS. Exactitud diagnóstica de las pruebas para detección de la infección causada por SARS-CoV-2.

Correspondencia

Instituto de Evaluación Tecnológica en Salud – IETS

Carrera 49 A # 91-91

Bogotá, D.C., Colombia.

www.iets.org.co

contacto@iets.org.co

2021.



Tabla de contenido

1.	Introducción.....	8
1.1.	Descripción de la condición	8
1.2.	Intervención	9
2.	Alcance y objetivos.....	14
2.1.	Objetivo General.....	14
3.	Pregunta de la revisión	15
4.	Metodología.....	16
4.1.	Criterios de elegibilidad y fuentes de evidencia de la literatura	16
4.2.	Búsqueda de la información	16
4.3.	Tamización, selección y extracción.....	17
4.4.	Evaluación de calidad.....	18
4.5.	Extracción de datos	19
5.	Resultados	20
5.1.	Búsqueda, tamización y selección de resultados.....	20
5.2.	Calidad metodológica de las RSL y RSL-metanálisis incluidas	21
5.3.	Calidad de la evidencia.....	22
5.4.	Síntesis de la evidencia	28
6.	Discusión.....	43
7.	Conclusiones.....	49
8.	Bibliografía	51
9.	Anexos	54
1	Anexo 1. Bitácoras de búsqueda de evidencia en bases de datos electrónicas... 54	
	Anexo 2. Diagrama PRISMA	59
	Anexo 3. Estudios incluidos.....	60
	1. Lista de artículos incluidos	60
	2. Tabla 4. Resumen de criterios de calidad de las RSL y RSL-metanálisis incluidas	62
	Anexo 4. Lista de artículos excluidos y las causas de exclusión	69
	Anexo 5. Evaluación del riesgo de sesgo de los estudios incluidos.....	75
	Anexo 6. Lista de verificación PRISMA-DTA	87



Anexo 7. Perfiles de evidencia GRADE.....	93
---	----



Lista de abreviaturas y siglas

Abs	Anticuerpos
ARN	Ácido ribonucleico
CDC	Centro de control y prevención de enfermedades (por su nombre en inglés: “center for disease control and prevention”)
CGIA	inmunoensayo de oro coloidal (por su nombre en inglés: “colloidal gold immunoassay”)
CLIA	inmunoensayo quimioluminiscente (por su nombre en inglés: “chemiluminescent immunoassay”)
COVID-19	Enfermedad por coronavirus 2019 (Por su nombre en inglés: “coronavirus disease 2019”)
CRISPR	Repeticiones Palindrómicas Cortas Agrupadas Regularmente Inter-espaciadas (Por su nombre en inglés: “Clustered regularly interspaced short palindromic repeats”)
ELISA	Ensayos inmunoabsorbentes ligados a enzimas (Por su nombre en inglés: “enzyme linked immunosorbent assays”)
FDA	Administración de Drogas y Alimentos de los Estados Unidos (Por su nombre en inglés: “Food and Drug Administration”)
FP	Falso positivo
FN	Falso negativo
IFI	Inmunofluorescencia indirecta
IFL	Inmunoensayo de flujo lateral
IQWiG	Instituto de calidad y eficiencia en salud (Por su nombre en alemán: “Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen”)
ITC	Comparación de tratamiento indirecta (Por su nombre en inglés: “Indirect Treatment Comparison”)
LAMP	Amplificación Isotérmica Mediada por Bucle (Por su nombre en inglés: “loop-mediated isothermal amplification”)
LIPS	Sistema de inmunoprecipitación de luciferasa (Por su nombre en inglés: “luciferase immunoprecipitation system”)
MERS-CoV	Síndrome respiratorio-coronavirus de Oriente Medio
MTC	Comparación de tratamientos múltiples (Por su nombre en inglés: “Mixed Treatment Comparison”)
nAbs	Anticuerpos neutralizantes
NAT	Pruebas de amplificación de ácido nucleico (Por su nombre en inglés: “nucleic acid amplification tests”)
OMS	Organización Mundial de la Salud
PRISMA	Elementos de informes preferidos para revisiones sistemáticas y metaanálisis (Por su nombre en inglés: “Preferred Reporting Items for Systematic Reviews and Meta-Analyses”)
RT-PCR	Reacción en Cadena de la Polimerasa con Transcriptasa Inversa



rRT-PCR	Reacción en Cadena de la Polimerasa con Transcriptasa Inversa en Tiempo Real
RVP	Razón de verosimilitud positiva
RVN	Razón de verosimilitud positiva
SARS-CoV	Síndrome respiratorio agudo severo-coronavirus
SARS-CoV-2	Síndrome de dificultad respiratoria aguda severa causada por coronavirus 2 (Por su nombre en inglés: “severe acute respiratory syndrome coronavirus 2”)
VP	Verdadero positivo
VPP	Valor predictivo positivo
VN	Verdadero negativo
VPN	Valor predictivo negativo



1. Introducción

1.1. Descripción de la condición

COVID-19

Desde su aparición en diciembre de 2019, el síndrome respiratorio agudo severo coronavirus 2 (SARS-CoV-2) ha causado más de 95 millones de infecciones conocidas y más de 2 millones de muertes en todo el mundo (1).

A pesar de la baja letalidad de enfermedad por coronavirus 2019 (COVID-19) (por debajo del 3% con un rango de 2 a 12% en diferentes países) en comparación con el síndrome respiratorio-coronavirus de Oriente Medio (MERS-CoV) (~40%) y el síndrome respiratorio agudo severo-coronavirus (SARS-CoV) (~15%), la tasa de diagnóstico y seguimiento en la comunidad debe aumentarse debido a la prevalencia generalizada y la focalización de muchas personas con enfermedades subyacentes, para evitar más muertes y sus posibles efectos secundarios desconocidos (2,3).

El SARSCoV-2 se transmite a través de las gotículas respiratorias al toser, estornudos, y también contacto directo. Por lo tanto, se propaga rápidamente principalmente entre miembros de la familia, profesionales de la salud y otros contactos cercanos. Los estudios y seguimiento han demostrado que el virus puede ser transmitido fácilmente a individuos sintomáticos o asintomáticos (2,4).

El tiempo entre la infección y la presentación de los síntomas varía de un paciente a otro. El período de incubación del SARS-CoV-2 es de 2 a 14 días según informes recientes del Centro de control y prevención de enfermedades (CDC; por su nombre en inglés) y de 2 a 10 días según la organización mundial de la salud (OMS). Estudios publicados recientemente encontraron que el período de incubación podría ser más largo que el establecido por la OMS y podría extenderse por aproximadamente 24 días (2,4).

Los síntomas clínicos informados son fiebre, tos, fatiga, producción de esputo, dificultad para respirar, dolor de garganta y dolor de cabeza. Algunos de los pacientes también manifiestan síntomas gastrointestinales como diarrea y vómitos. Los ancianos y las personas con enfermedades crónicas desarrollan rápidamente síndrome de dificultad respiratoria aguda, shock séptico y síndrome de coagulación intravascular diseminada, que incluso los lleva a la muerte (2–4).

Es esencial disponer de ensayos que sean rápidos, simples e idealmente con alta sensibilidad y exactitud, permitiendo realizar a gran escala un diagnóstico precoz efectuando el seguimiento del afectado en afán de prevenir la propagación. Las pruebas que están presentes en el mercado deben contar con la aprobación y conformidad de las autoridades regulatorias para asegurar que cuentan con el régimen necesario. Existen dos tipos principales de pruebas: las pruebas de diagnóstico de detección del



Ácido ribonucleico (ARN) viral (moleculares) y las pruebas inmunológicas ya sea de detección de antígenos virales o detección de anticuerpos (serológicas) (5,6).

1.2. Intervención

Pruebas diagnósticas para infección por SARS-CoV-2

Las pruebas para COVID-19 se dividen en dos grandes grupos: pruebas que detectan la presencia del virus SARS-CoV-2 y pruebas que detectan la presencia de anticuerpos contra el SARS-CoV-2. Las pruebas para detectar la presencia de virus suelen utilizar métodos que reconocen y amplifican el ácido nucleico viral del SARS-CoV-2, como la reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) o la amplificación isotérmica (LAMP). La prueba del virus del SARS CoV-2 generalmente se realiza en un entorno de laboratorio especializado utilizando muestras respiratorias, como hisopos nasofaríngeos, pero también se han desarrollado pruebas cerca del paciente (3,4,7).

Los inmunoensayos pueden detectar el antígeno viral (pruebas de antígeno) o la respuesta inmune (pruebas de anticuerpos), respectivamente (7). Las pruebas de anticuerpos contra el SARS CoV-2 (también llamadas pruebas serológicas) se realizan en muestras de sangre o suero y se han desarrollado tanto para el análisis en un laboratorio como en un entorno cercano al paciente en el punto de atención (6,7).

Dado que los anticuerpos se producen como parte de la respuesta inmunitaria del cuerpo a la infección, las pruebas serológicas pueden ser útiles para identificar una infección por SARS-CoV-2 en curso, en recuperación (convaleciente) o una infección previa (6,7). Sin embargo, los ensayos de detección de anticuerpos (Abs) tienen inconvenientes para la detección temprana de casos debido a la ausencia de estos en la primera etapa de la enfermedad. Por lo tanto, en esta etapa, la detección de los antígenos virales mediante inmunoensayos puede ser de gran utilidad diagnóstica (7).

Pruebas para la detección de ARN viral

Pruebas moleculares

En las enfermedades infecciosas, los ácidos nucleicos que pertenecen al patógeno se encuentran en muestras clínicas en un número bajo de copias, por lo que se deben utilizar técnicas de amplificación para detectar la presencia de patógenos (7). Las diferentes pruebas de amplificación de ácido nucleico (NAT) viral con el método RT-PCR y LAMP, entre otras, han sido usadas para detectar la presencia del ARN viral (3,7).

Para la realización de estas pruebas, se deben tomar muestras de las regiones nasofaríngea y orofaríngea, esputo y tracto respiratorio inferior de pacientes con sospecha de infección por COVID-19, independientemente del inicio de los síntomas (3,4,7).



Se han descubierto tres regiones conservadas entre los genomas virales relacionados con el SARS-CoV-2: el gen E (gen de la proteína de la envoltura), el gen RdRP (gen de la ARN polimerasa dependiente de ARN) en la región de marco de lectura abierto ORF1ab, y el gen N (gen de la proteína de la nucleocápside); estas regiones son objeto de las pruebas moleculares disponibles o en desarrollo (3,7).

La LAMP se ha desarrollado como una técnica rápida, precisa, confiable y más económica para amplificar la región objetivo a una temperatura de reacción única en lugar del ciclo térmico requerido en RT-PCR. La ventaja del método LAMP para la RT-PCR es que la cantidad de ADN producido es mucho mayor y un resultado positivo de la prueba se puede ver visualmente sin la necesidad de un paso de análisis adicional. Si bien algunos estudios han reportado que los métodos LAMP presentan más del 97% de sensibilidad en la selección del gen ORF1ab en comparación con la RT-PCR, otros han mostrado que ambos métodos tienen la misma sensibilidad y pueden detectar una muestra diluida 20 veces. También se ha demostrado que la técnica es muy específica, ya que se usan de cuatro a ocho cebadores en el análisis de LAMP para identificar ocho regiones diferentes en el ADN diana (3,4,7).

La prueba de diagnóstico LAMP para COVID-19 es más específica y sensible en comparación con los ensayos convencionales de RT-PCR y no depende de equipos de laboratorio especializados, como un termociclador. Sin embargo, debido a la multiplicidad de cebadores utilizados en este método, optimizar las condiciones de reacción es un gran desafío (3). Por lo tanto, en la actualidad, la RT-PCR sigue siendo el estándar de oro para el diagnóstico molecular del SARS-CoV-2 (3,7).

Pruebas inmunológicas

La sensibilidad de una NAT que se realizará en el período temprano de la enfermedad o en la etapa de recuperación será muy baja y la prueba dará resultados falsos negativos incluso en presencia de infección. En particular, las muestras recolectadas en la etapa tardía de la enfermedad pueden no tener suficiente carga viral y dar lugar a una baja tasa de positivos para RT-PCR (3,5). No obstante, las mismas muestras clínicas también tienen altas cantidades de anticuerpos específicos para el virus. Por lo tanto, detectar los anticuerpos mediante inmunoensayos es un método preferible en esta etapa (2,6).

Para superar todas estas limitaciones de las pruebas moleculares, los inmunoensayos se pueden utilizar como método complementario debido a sus ventajas, como un muestreo y realización comparativamente más fácil, menos requisitos de experiencia técnica y equipo (3,5). Por otro lado, los inmunoensayos también pueden diseñarse no solo para la detección de anticuerpos sino también para la detección de antígenos como alternativa a la RT-PCR (7).

Detección de anticuerpos

Hay dos tipos generales de anticuerpos, los anticuerpos neutralizantes (nAbs) y los no neutralizantes (también conocidos como anticuerpos de unión). La neutralización se



define como la pérdida de infectividad que ocurre cuando un nAb se une a una partícula viral (5,6).

Las pruebas de anticuerpos anti-SARS-CoV-2 disponibles comercialmente utilizan diferentes tecnologías para medir cualitativamente clases de inmunoglobulinas individuales (IgM, IgG o IgA) o anticuerpos totales, pero no diferencian los nAb de los anticuerpos de unión (5). Los anticuerpos IgM dirigidos contra microorganismos se producen típicamente primero después de la infección y se utilizan como medida de una infección reciente. Aproximadamente el 40% de los pacientes producen IgM en su primera semana de infección con el SARS-CoV-2, pero este porcentaje de pacientes aumenta a cerca del 95% en las siguientes dos semanas, es decir, de 12 a 14 días (6).

Los anticuerpos IgG generalmente se desarrollan más tarde después de la IgM y permanecen elevados durante meses o años después de la infección (5). Aunque los anticuerpos IgM pueden detectarse dentro de las dos primeras semanas de síntomas en algunos pacientes, la infección por SARS-CoV-2 parece inusual ya que IgM e IgG aumentan juntas más comúnmente, más de dos semanas después del inicio de los síntomas (5,6).

La IgA secretora es importante para la inmunidad de las mucosas. La IgA también se puede detectar sistémicamente en ciertos tipos de infección, incluido el SARS-CoV-2, pero se sabe comparativamente poco sobre la cinética de la IgA en la sangre. Los componentes del "anticuerpo total" incluyen presumiblemente IgM e IgG y, teóricamente, también otras inmunoglobulinas específicas de antígeno (5,6).

Detección de antígenos virales

Cuatro proteínas estructurales pertenecen a la familia de los coronavirus: proteínas de pico (S), membrana (M), envoltura (E) y nucleocápside (N). La mayoría de los ensayos serológicos utilizan las proteínas S y N como antígenos, ya que estos son los dos antígenos que mostraron la mayor sensibilidad dentro de los ensayos disponibles comercialmente analizados hasta ahora (6,7).

En la infección por SARS-CoV-2, se induce el desarrollo de IgG contra la proteína N que puede detectarse en suero a partir del día 4 después del inicio de la enfermedad y en la mayoría de los pacientes seroconvertidos el día 14, pero al parecer las pruebas que utilizan IgG dirigida a un antígeno S son más sensibles que pruebas basadas en el antígeno N. Después del inicio de los síntomas, las IgM e IgG específicas de N y S aumentan de manera constante y pueden usarse para identificar la infección por SARS-CoV-2 (3,5,7).

Estos inmunoensayos se pueden realizar con diversas muestras clínicas debido a su carga variable de antígeno. Si bien se puede usar sangre completa, suero o plasma como muestra para inmunoensayos basados en anticuerpos, las muestras respiratorias superiores o inferiores también se pueden usar para inmunoensayos basados en antígenos. Sin embargo, la obtención de muestras de sangre es más fácil y sin riesgos



en comparación con las muestras respiratorias, y por lo tanto la mayoría de los estudios contemplan estos ensayos como pruebas serológicas (5,7).

Pruebas serológicas

Las plataformas de diagnóstico clínico más comunes utilizadas para el SARS-CoV-2 incluyen ensayos inmunoabsorbentes ligados a enzimas (ELISA), inmunoensayos quimioluminiscentes (CLIA) y dispositivos de inmunoensayos de flujo lateral (IFL) incluyendo los inmunoensayos cromatográficos con oro coloidal (CGIA) (4–6).

Los IFL generalmente requieren una gota de sangre, suero o plasma, aplicada a una tira reactiva, y los resultados se leen en aproximadamente 15-30 minutos. Estos dispositivos son adecuados para pruebas en el lugar de atención y tienen potencial para implementarse en el campo como parte de grandes estudios de seroprevalencia. Los ELISA vienen en una variedad de formatos diferentes; normalmente, un complejo de antígeno-anticuerpo unido se detecta usando un anticuerpo secundario específico unido a un sustrato que genera una señal colorimétrica o fluorescente (4–6).

Los métodos CLIA son similares al ELISA, pero utilizan sondas químicas que emiten luz en lugar de sustratos enzimáticos. Tanto ELISA como CLIA son métodos de laboratorio clínico adaptables a pruebas de alto rendimiento utilizando suero, plasma o gotas de sangre potencialmente secas (4–6).

Muchas pruebas serológicas diferentes para el SARS-CoV-2 están disponibles comercialmente en un corto período de tiempo. La increíble velocidad de desarrollo ha superado significativamente las evaluaciones rigurosas del rendimiento de las pruebas (5,6).

1.3. Prueba estándar de referencia

Actualmente, el ensayo de RT-PCR es el método estándar de oro (referencia) para diagnosticar el SARS-CoV-2. Desafortunadamente, la sensibilidad de la prueba de ARN en el mundo real no es satisfactoria y también se han reportado casos de falsos negativos debido a problemas con la recolección y transporte de muestras, extracción de ARN, inhibidores de enzimas y el método RT-PCR (3,5). De hecho, las pruebas de RT-PCR tienen muchas limitaciones debido a que requieren una gran carga de trabajo, necesitan operadores hábiles para las pruebas y la recolección de muestras, y necesitan instrumentos costosos y lugares de operación especiales (3,5).

La Administración de Drogas y Alimentos de los Estados Unidos (FDA) concluyó que un resultado negativo de la prueba de RT-PCR no excluye por completo la infección por SARS-CoV-2 y no debería usarse como una única referencia para el diagnóstico (7). Existen estudios que sugieren que el uso combinado de NAT y pruebas serológicas puede mejorar significativamente la sensibilidad diagnóstica, así como la tasa de detección positiva. Si bien los estudios han reportado una sensibilidad que oscila entre el 78,7% y el 100%, se ha informado una tasa de detección positiva del 98,6% utilizando



el ensayo de IgM combinado con NAT en comparación con solo el 51,9% utilizando una única prueba de RT-PCR (5).

1.4. **Justificación de esta revisión sistemática de la literatura**

Dadas las limitaciones del diagnóstico clínico (debido a la similitud de los síntomas de la infección por COVID-19 con los de otros virus) y la disponibilidad de diferentes pruebas moleculares y serológicas con ventajas y desventajas técnicas, es importante resumir los parámetros de exactitud de estos métodos e investigar si son lo suficientemente específicos o sensibles para su uso en la práctica clínica diaria.

Es importante también, abordar el rendimiento diagnóstico de las pruebas para COVID-19, con un enfoque especial en los análisis de las pruebas comerciales disponibles.



2. Alcance y objetivos

Realizar una revisión sistemática de la literatura para recopilar evidencia sobre las características de las pruebas diagnósticas moleculares e inmunológicas disponibles para el SARS-CoV-2, y su rendimiento diagnóstico, incluidos los parámetros de sensibilidad, especificidad, razón de verosimilitud positiva y negativa, y valor predictivo positivo y negativo, siempre que sea posible.

2.1. Objetivo General

El objetivo de esta revisión sistemática fue identificar evidencia que pueda usarse para establecer la exactitud diagnóstica de las pruebas moleculares (amplificación de ácido nucleico) e inmunológicas (detección de antígenos virales y anticuerpos) para la detección de infección por el SARS-CoV-2.



3. Pregunta de la revisión

¿Cuál es la exactitud diagnóstica de las pruebas diagnósticas (moleculares e inmunológicas) que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Tabla 1. Pregunta de evaluación en estructura PICO

P	Hombres y mujeres de cualquier edad, ocupación y ubicación geográfica expuestos al SARS-CoV-2 o con sospecha de infección por SARS-CoV-2
I	<ul style="list-style-type: none">● Pruebas moleculares de amplificación de ácido nucleico (NAT) para SARS-CoV-2 que identifican el ARN viral● Pruebas inmunológicas para medir antígenos virales o anticuerpos (serológicas) contra el virus SARS-CoV-2
C	RT-PCR o RT-PCR más diagnóstico clínico
O	Medidas de exactitud diagnóstica: <ul style="list-style-type: none">● Sensibilidad (S)● Especificidad (E)● Verdadero positivo (VP)● Verdadero negativo (VN)● Falso positivo (FP)● Falso negativo (FN)● Valores predictivos positivos (VPP) y negativo (VPN)● Razones de verosimilitud positivos (RVP) y negativos (RVN)

P: Población I: Intervención C: Comparador O: Desenlaces (del inglés "outcome")



4. Metodología

4.1. Criterios de elegibilidad y fuentes de evidencia de la literatura

Se realizó una búsqueda exploratoria libre preliminar en MEDLINE a través de PubMed (diciembre 16/2020) para evaluar la cantidad de literatura disponible y el tipo de estudios publicados usando los términos: “SARS-Cov-2”, “COVID19”, y “diagnostic tests”. Se logró evidenciar una abundante cantidad de estudios secundarios (revisiones sistemáticas de la literatura-RSL y metanálisis) relacionados con el rendimiento diagnóstico de diferentes pruebas para COVID-19. Teniendo en cuenta este hallazgo, se decidió realizar una revisión sistemática de RSL y metanálisis (“overview”) con el fin obtener la mayor información posible de análisis agrupados de los diferentes estudios publicados desde el inicio de la pandemia.

Criterios de inclusión

Los estudios fueron incluidos en esta revisión sistemática si cumplían con los siguientes criterios de elegibilidad:

- Revisiones sistemáticas o metanálisis de evaluación de cualquier método de diagnóstico molecular o serológico
- Dirigidos al diagnóstico de SARS-CoV-2 (COVID-19)
- Que usen cualquier muestra biológica humana
- Que informan datos sobre la exactitud de la prueba (sensibilidad, especificidad, valores predictivos, razones de verosimilitud, verdaderos positivos y negativos, falsos positivos y negativos)
- Publicados desde el año 2019

Criterios de exclusión

- Estudios preimpresos que no hayan sido revisados por pares.
- Estudios que contemplen la tomografía computarizada de tórax como medio diagnóstico.
- Estudios para todos los coronavirus (no específicos para COVID-19)
- Estudios publicados en caracteres no romanos

4.2. Búsqueda de la información

Se realizó una búsqueda sistemática de la literatura de acuerdo con lo propuesto por el Manual para la elaboración de evaluaciones de efectividad, seguridad y validez diagnóstica de tecnologías en salud del IETS (8). Se consultaron las bases de datos electrónicas MEDLINE, Embase, Cochrane y Epistemonikos; adicionalmente se realizó búsqueda de literatura gris a través de Google Académico. Se utilizó el lenguaje controlado y términos libres como “rapid test”, “diagnostic tests”, “laboratory tests”, “serology”, “serologic tests”, “molecular diagnosis”, “molecular diagnosis techniques”,



“SARS-Cov-2”, “COVID19” con operadores booleanos y truncadores en las estrategias de búsqueda.

- Se accedió a MEDLINE a través de PubMed (<https://www.ncbi.nlm.nih.gov/pubmed/>).
- Se accedió a Embase a través de (www.embase.com)
- Se accedió a la biblioteca de Cochrane a través de (www.cochranelibrary.com).
- Se accedió a Epistemonikos a través de (<https://www.epistemonikos.org/es>)
- Se accedió a Google Académico a través de (<https://scholar.google.es>)

Las estrategias de búsqueda diseñadas para cada una de las bases de datos se presentan en las bitácoras de búsqueda en el Anexo 1.

4.3. Tamización, selección y extracción

Tamización de estudios

La tamización de estudios se realizó empleando el software Rayyan®, donde inicialmente se cargaron todos los documentos identificados como resultado de las búsquedas en las diferentes bases de datos; posteriormente se realizó un proceso de eliminación de duplicados.

La tamización inicial, se hizo mediante la lectura de título y resumen a cargo de 2 revisores independientes (AMHT, KCP), los desacuerdos fueron resueltos por consenso.

Selección de estudios

El proceso completo de tamización y selección utilizado se presenta empleando el diagrama de flujo propuesto en la declaración *PRISMA* (9) (Anexo 2).

Se incluyeron aquellos estudios que cumplían con los criterios de elegibilidad descritos previamente. Para la selección de estudios, se llevó a cabo la recuperación de texto completo de las referencias preseleccionadas en la tamización y se procedió a lectura del texto completo, la cual se realizó de manera independiente por dos revisores (AMHT, KCP). Los desacuerdos fueron resueltos por consenso. Los estudios incluidos en la fase de revisión en texto completo son presentados mediante una lista (Anexo 3). De igual manera, los estudios excluidos se presentan junto con la respectiva justificación de su exclusión (Anexo 4).

Extracción de la información

Se diseñó un formulario de recopilación de datos en Excel®, para obtener información relevante, de los estudios incluidos. Los datos fueron extraídos de forma independiente por dos revisores.

Los datos extraídos de cada estudio incluido, cuando estos estuvieron presentes, fueron:



- Autor principal y año de publicación
- Lugar de estudio (país).
- Cantidad y diseño de los estudios incluidos
- Tamaño total de la población
- Prevalencia de la condición en cada estudio
- Prevalencia de sintomáticos
- Prevalencia de asintomáticos
- Días a la toma de la muestra desde inicio síntomas
- Tipo de prueba diagnóstica (intervención)
- Punto de corte predefinido en cada estudio
- Marca o casa comercial (en aquellos estudios con disponibilidad de la información)
- Comparador (prueba de referencia y sus características)
- Tipo de muestra biológica
- Desenlaces medidos
 - Sensibilidad
 - Especificidad
 - Valores predictivos: positivos y negativos
 - Razones de verosimilitud positivos y negativos
 - Verdaderos positivos, falsos positivos, verdaderos negativos, falsos negativos
- Conclusiones
- Limitaciones

4.4. Evaluación de calidad

La evaluación de calidad se llevó a cabo mediante el uso de la herramienta ROBIS (10) para evaluar la calidad metodológica de cada RSL incluida. La evaluación de las RSL con comparación indirecta de tratamiento (ITC) o con comparación de tratamientos múltiples (MTC) se realizó con la herramienta IQWiG (11) (Anexo 5).

Se hizo uso del estadístico kappa para medir el acuerdo entre los revisores, tomando decisiones simples de inclusión / exclusión según las evaluaciones de calidad. Un valor kappa entre 0,60 y 0,74 se consideró un buen acuerdo y 0,75 o más un excelente acuerdo (12).

Cuando dos o más RSL contemplaron los mismos estudios primarios, se decidió incluir aquella con mejor calidad metodológica y fecha de aceptación para publicación más reciente.

Para complementar la evaluación ROBIS e IQWiG se realizó una valoración específica del informe de las revisiones incluidas. Los mismos dos revisores evaluaron el estándar de informes de cada revisión utilizando la lista de verificación PRISMA-DTA (Elementos de informes preferidos para una revisión sistemática y metanálisis de estudios de precisión de pruebas diagnósticas) (13). El *PRISMA-DTA* consta de 27 ítems, y cada ítem fue respondido con "Sí" (cumplimiento total), "Parcial" (cumplimiento parcial) o "No" (incumplimiento). Se atribuyó una puntuación de "1" por cada "Sí", "0,5" punto por cada "Parcial" y "0" por cada "No". El *PRISMA-DTA* tiene un puntaje máximo de 27. Se consideró que el RSL tiene fallas mayores si la puntuación total es menor de 15, fallas



menores si la puntuación total está entre 15,5 y 21, y fallas mínimas si la puntuación total es entre 21,5 y 27 (Anexo 6).

La certeza de la evidencia se evaluó utilizando el enfoque desarrollado por el GRADE Working Group, que considera cinco criterios (riesgo de sesgo, evidencia indirecta, inconsistencia, imprecisión y sesgo de publicación) para gradar la calidad de la evidencia de los estudios incluidos (9,14–16) (Anexo 7). Estas evaluaciones de calidad de la evidencia fueron realizadas por dos revisores de forma independiente y en casos de discrepancias se consultó un tercer revisor para resolver los desacuerdos.

4.5. Extracción de datos

Los datos relevantes para esta RSL fueron extraídos a partir de las RSL y metanálisis con menor riesgo de sesgo; la cantidad de evidencia disponible permitió incluir más de un estudio en términos de comparaciones y desenlaces.

Debido a que las revisiones utilizaron diferentes enfoques para presentar sus datos, para los datos metanalíticos, se realizó la extracción de las medidas cuantitativas de resumen agrupadas de acuerdo con análisis de heterogeneidad realizado en cada estudio para definir los rangos de rendimiento, pero también se registraron los datos individuales informados en cada revisión. De aquellos estudios que solo reportaban un resumen narrativo, se extrajeron las conclusiones de los autores sobre la exactitud de la prueba.

Las características relacionadas con los parámetros de calidad de los estudios incluidos como el diseño de los estudios primarios, bases de datos buscadas, la herramienta de evaluación de calidad y sesgos, evaluación de heterogeneidad, y tipo de metanálisis se presentan en la tabla 4 del Anexo 3.

Se contemplaron los siguientes subgrupos de análisis: tiempo desde el inicio de los síntomas, subtipos de pruebas tanto moleculares como serológicas, tipo de anticuerpo o antígeno viral detectado (para pruebas serológicas), genes diana para las NAT, y tipos de muestras biológicas usadas.



5. Resultados

5.1. Búsqueda, tamización y selección de resultados

Las búsquedas recuperaron 1309 referencias (PubMed, EMBASE, Cochrane, Epistemonikos) más 27 estudios en la búsqueda complementaria (Google académico). De estos, después de remover duplicados, se tamizaron 1254 por título y resumen, arrojando un total de 62 estudios a ser evaluados en texto completo. Se excluyeron 49 estudios principalmente debido a: que no respondían directamente la pregunta de investigación (algunos sin datos de rendimiento); estudios que usaron TC de tórax como medio diagnóstico para COVID-19; estudios no específicos para SARS-CoV-2 (para coronavirus en general sin hacer análisis diferencial para SARS-CoV-2); y estudios no RSL o metanálisis o que no habían sido evaluados por pares. La lista completa de referencias de los estudios excluidos y las razones de exclusión se presenta en el anexo 4. Los resultados de las estrategias de búsqueda y flujograma *PRISMA* de selección se detallan en los anexos 1 y 2.

Por último, dieciséis estudios fueron seleccionados para evaluación de calidad metodológica, resultando tres estudios excluidos por baja calidad. El acuerdo entre los revisores relacionado con la evaluación de calidad fue excelente (promedio estadística kappa = 0,758).

Características de los estudios incluidos

Un total de trece estudios fueron incluidos para la síntesis y evaluación de la calidad de la evidencia, de estos, cuatro fueron RSL (17–20) y nueve RSL- metaanálisis (21–29). El estudio de Mustafa et al. (26) incluyó evaluación de exactitud diagnóstica de pruebas moleculares para infección por cualquier tipo de coronavirus, sin embargo, realizaron un metanálisis completo y específico por separado para SARS-CoV-2 y por esta razón fue incluido en este análisis.

Doce estudios fueron publicados o aceptados para publicación entre mayo y diciembre de 2020. El estudio de Boger et al. (21) fue publicado en 01/01/2021. Todas las RSL se publicaron en inglés siendo cinco de ellas colaboraciones de autores de diferentes países (18,19,23,24,26), cuatro provenientes de China (22,25,28,29), dos de Europa (Italia y Reino Unido) (20,27), y dos de Sur América (Argentina y Brasil) (17,21).

Seis de los estudios evaluaron el rendimiento de pruebas diagnósticas moleculares (17,19–21,26,27), mientras que nueve de las RSL incluidas evaluaron el desempeño de pruebas serológicas (18,20–25,28,29).

El número de estudios primarios originales incluidos en las RSL que evaluaron las pruebas moleculares fue de 136. De las seis RSL, cinco reportaron algún tipo de información acerca del diseño de los estudios primarios (19–21,26,27). Todos los estudios fueron observacionales; en general, se describe claramente la inclusión de 18 estudios con dos o más grupos de evaluación (con control) (19–21,27), y 28 estudios de un solo grupo (20,21,27). Treinta estudios son descritos como retrospectivos (20,21,26)



y 14 prospectivos (20,26). El total de participantes en los estudios incluidos fue 29281 de los cuales 4477 fueron reportados como controles negativos. En la RSL de Mustafa et al. (26) se incluyeron muestras control pre COVID-19. Dos RSL informaron acerca de características demográficas de los participantes (17,20); el rango de la mediana de edad reportada fue 45-63 años y rango de frecuencia relativa de hombres de 44-60%. El tipo de las muestras biológicas usadas en los estudios fue hisopos nasofaríngeos en 38%, hisopos orofaríngeos 25%, saliva 11,8%, frotis nasal 4,4%, esputo 2,9%, orina 2,9%, heces y frotis rectal 2,9%, sangre 2,2%, lavado broncoalveolar (LBA) 1,5%, y aspirado traqueal 0,7%. El tiempo de toma de muestras desde el inicio de síntomas se reportó en ocho estudios de la RSL de Arevalo-Rodriguez et al. (17), con un rango de 0-14 días en seis de ellos, y hasta 24 días en los dos restantes; y en solo un estudio en la RSL de Dinnes et al. (19) para una mediana de 2 días.

Para las RSL que evaluaron las pruebas inmunológicas, el número de estudios primarios incluidos fue 274, de los cuales, 101 estudios incluyeron un grupo control (18–20,23,24,28), mientras que 93 fueron descritos como series de casos (18–21,23,24,28,29); 50 con diseño retrospectivo (18,20,21,29) y 13 como prospectivos (20,28,29). Se incluyeron 51726 participantes en total en las ocho de las nueve RSL que reportaron este dato (18–21,23,24,28,29). Cinco de las RSL describieron la inclusión de 10484 participantes o muestras como controles negativos (21–23,28,29). Cuatro RSL reportaron características demográficas de los participantes (18,20,23,28); el rango de la mediana de edad informada fue 25-76 años y rango porcentual de hombres de 26-87%. El tipo de pruebas realizadas en los estudios fue ELISA 42,3%, CLIA 34,3%, IFL 35,7%, IFI 5,1%, CGIA 12,8%, y otros 2,5%. Las RSL de Deeks et al. (18) y Dinnes et al. (19) incluyeron en total 7 estudios primarios de pruebas realizadas en el punto de atención. El tiempo de toma de muestras desde el inicio de síntomas se reportó en cuatro RSL (en 109 estudios primarios) (18,22,23,29); el rango fue muy heterogéneo variando desde 0 hasta 62 días.

El resumen de las características básicas de los estudios incluidos se presenta en la tabla 2.

5.2. Calidad metodológica de las RSL y RSL-metanálisis incluidas

Teniendo en cuenta la calidad metodológica general, todas las RSL y RS-metanálisis se calificaron como de calidad moderada.

Todos los estudios realizaron búsquedas en al menos dos bases de datos en inglés. El estudio de Zhang et al.(29) también realizó búsquedas en bases de datos chinas. Ocho estudios reportaron la estrategia de búsqueda completa en el manuscrito principal o en su material suplementario (17–20,24–26,28), y once estudios realizaron una búsqueda bibliográfica complementaria en buscadores de literatura gris o búsqueda manual siguiendo el método de “bola de nieve” (17–26,29).

Doce RSL evaluaron la calidad metodológica de los estudios incluidos a través de la herramienta de evaluación de la calidad de los estudios de precisión diagnóstica-2



(QUADAS-2). El estudio de Riccò et al. (27) solo describe evaluación de sesgo de publicación mediante gráficos en embudo.

De las 4 RSL, evaluadas con la herramienta ROBIS, solo la RSL de Jarrom et al. (20) presentó deficiencias relacionadas con el dominio 1 de criterio de elegibilidad de los estudios y dominio 2 de identificación y selección de los estudios, debido a que en los criterios de elegibilidad se limitaron a incluir estudios publicados sólo en el idioma inglés, que pudo significar dejar por fuera estudios importantes publicados en otro idioma. De igual manera, la RSL de Jarrom et al. generó preocupaciones relacionadas con el dominio 4 de síntesis y resultados debido a que en la metodología refieren la realización de metanálisis, pero no abordan la evaluación de heterogeneidad de los estudios incluidos y realizan un análisis cualitativo de la evidencia.

De las RSL-metanálisis evaluadas con la herramienta IQWiG, solo en una se realizó una clara descripción de la pregunta de investigación; el resto de los estudios solo plantearon objetivos. Siete de los estudios realizaron comparaciones directas y en los restantes no se justificó el uso de comparación indirecta. Solo la RSL-metanálisis de Mustafá et al. (26) explicó la selección de un comparador común. Cinco de las RSL presentaron deficiencias no críticas en la metodología de la revisión sistemática. La mayoría de los estudios utilizaron un modelo bivariado para la realización del metanálisis y en general, la descripción metodológica y ejecución de estos fue de buena calidad.

En la tabla 4 del anexo 3 se presentan los parámetros completos de calidad metodológica de las RSL incluidas en este estudio.

De acuerdo con la lista de chequeo *PRISMA-DTA*, se informaron sólo seis ítems en su totalidad (100%) considerando además que nueve de las trece RSL no reportaron el protocolo ni el registro de este. No más de ocho de las RSL informaron adecuadamente acerca de las estrategias de búsqueda. El riesgo de sesgos y aplicabilidad fue adecuadamente reportado en doce de las RSL. En general, los puntajes más bajos y altos de *PRISMA-DTA* fueron 17,5 y 26,5; el puntaje promedio fue 23,07. Del total de RSL incluidas, cuatro tenían fallas menores en los informes y nueve tenían fallas mínimas (Anexo 6).

5.3. Calidad de la evidencia

Los resultados de la clasificación de la calidad de la evidencia para los tipos de pruebas diagnósticas y subgrupos evaluados se muestran en las tablas del anexo 7.

Toda la evidencia para el rendimiento de las pruebas moleculares fue calificada como muy baja (19–21,26,27). En general, el alto riesgo de sesgo en los estudios primarios, la imprecisión, y el sesgo de publicación fueron las razones principales para la baja calidad de la evidencia mediante los criterios GRADE. El alto riesgo de sesgos en todos los estudios se debió a la falta de claridad en la selección de los pacientes y a la ausencia de grupo control. La imprecisión se debió principalmente a la alta heterogeneidad en el tamaño y tipo de muestras evaluadas y a la poca claridad en la definición de la prueba PCR como prueba índice y como prueba de referencia. El sesgo



de publicación en un estudio fue claramente detectado (27) mientras que en otros tres no fue claro por falta de evaluación de este (19,21,26). La combinación de resultados demasiado heterogéneos dio como resultado hallazgos inconsistentes e incapacidad para sacar conclusiones significativas para una de las RSL (20).

De igual manera, la evidencia relacionada con la exactitud diagnóstica de los ensayos inmunológicos para detección de anticuerpos fue calificada como de muy baja calidad para la mayoría de los estudios (18,20,21,25,28,29) y baja para dos RSL (23,24). La muy baja o baja calidad de la evidencia mediante los criterios GRADE se debió principalmente al alto riesgo de sesgo de los estudios al no describir claramente los criterios de muestreo y selección de los pacientes, la selección de los controles, y la interpretación de los resultados con respecto a la prueba de referencia. Adicionalmente, muchos estudios se describen como retrospectivos o prospectivos sin hacer claridad en el diseño de estos. Otro factor problema fue con frecuencia la alta variabilidad en los estudios primarios, pues en algunos el tamaño efectivo de la muestra era a menudo pequeño y las medidas de rendimiento de las pruebas demasiado heterogéneas llevando a alta inconsistencia. En todos los estudios excepto la RSL de Kontou et al. (23) se detectó un alto riesgo de sesgo de publicación derivado principalmente en las restricciones de idioma en las búsquedas o criterios de elegibilidad de los estudios. La evidencia indirecta y la imprecisión fueron razones menos comunes para la baja calidad a excepción de los estudios de Deeks et al. (18) y Jarrom et al.(20) en el que se informaron rangos muy amplios de valores de sensibilidad.



Tabla 2. Detalles resumidos de los estudios incluidos

Autor/ año	Origen/ país	Número de estudios incluidos	Población total	Edades/Sexo	Número casos / número controles	Prueba diagnóstica evaluada (N estudios primarios)	Comparador- Estándar de referencia	Tipo de muestra biológica (N estudios primarios)	Desenlaces reportados
Arevalo- Rodríguez et al. 2020 (17)	Argentina	34 estudios observacionales	12057	Rango promedio: 2,5-56 años. Rango mediana: 45- 63 años Hombres 5331	10997 pacientes con infección confirmada por SARS-CoV-2 1060 casos con resultados negativos de RT-PCR en su evaluación inicial	RT-PCR	RT-PCR (segundo positivo)	Hisopos nasofaríngeos (13) Hisopos orofaríngeos (7)	FN
Boger et al. 2021 (21)	Brasil	12 estudios observacionales	1479	NR	170 pacientes evaluación 149 pacientes controles	RT-PCR (8)	RT-PCR	Heces e hisopos rectales (4) Orina (4) Sangre (3) Espujo (2) Saliva (2) Hisopos nasofaríngeos y de garganta, y aspirado nasofaríngeo (4)	Sensibilidad, especificidad RVP y RVN
					814 pacientes evaluación 319 pacientes controles	Pruebas serológicas (5) (ELISA, IFL)		IgM/IgG sangre, suero, plasma (9) IgM sangre, suero, plasma (11) IgG sangre, suero, plasma (10)	
Deeks et al. 2020 (18)	Internacional	54 estudios observacionales	15976	Rango mediana: 37- 76 años Hombres 2217-7418 (rango)	8256 muestras fueron de casos de COVID-19 confirmados	Pruebas serológicas De laboratorio (52): ELISA (27) CLIA (19) Otras (6) En el punto de atención (2): IFL, CGIA	RT-PCR/ diagnóstico clínico	Plasma, suero, sangre por punción en el dedo, sangre venosa entera	Sensibilidad, especificidad



Dinnes et al. 2020 (19)	Internacional	18 estudios observacionales	943	NR	596 muestras fueron de casos de COVID-19 confirmados	Pruebas de antígeno en el punto de atención (5): CGIA (4) IFI (4)	RT-PCR/ diagnóstico clínico	Hisopos nasofaríngeos (2) Hisopos nasofaríngeos y orofaríngeos (2) Hisopos y aspirado nasofaríngeos y LBA (1)	Sensibilidad, especificidad
			2255		1175 muestras fueron de casos de COVID-19 confirmados	Pruebas moleculares rápidas en el punto de atención (13)		Hisopos nasofaríngeos (6) Frotis nasales (1) Hisopos nasofaríngeos y orofaríngeos (1) Hisopos nasofaríngeos o nasales (3) Aspirado de tráquea (1)	
Guo et al. 2020 (22)	China	30 estudios observacionales	7917	NR	3856 casos confirmados de SARSCoV-2 368 casos sospechosos con PCR negativa 1167 individuos asintomáticos 2526 controles negativos	Pruebas serológicas ELISA (14) CLIA (8) IFL (15) IFI (2)	RT-PCR/ diagnóstico clínico	Sangre total / suero / plasma	Sensibilidad, especificidad RVP y RVN
Jarrom et al. 2020 (20)	Reino Unido	62 estudios observacionales / 1 RSL	6072	Rango mediana: 45-52 años Hombres 47-60% (5 estudios)	3818 pacientes (16 estudios) para RT-PCR 972 pacientes para amplificación isotérmica	Pruebas moleculares para detección del virus RT-PCR (33) LAMP (5)	RT-PCR	Lavado bronco alveolar, lavado de garganta, hisopos/aspirado nasofaríngeo, lingual, saliva, esputo, orina, heces y/o hisopos rectales, lágrimas/hisopado conjuntival, biopsia con cepillo de fibrobroncoscopio	Sensibilidad, especificidad VPP y VPN
				Rango mediana: 32-48 años	1282 pacientes para pruebas de anticuerpos	Pruebas serológicas ELISA (7) IFL, CLIA, CGIA (17)		Plasma / sangre	



				Hombres 48-68% (4 estudios)					
Kontou et al. 2020 (23)	Internacional	38 estudios observacionales	7848	Rango mediana: 43-76 años Hombres 34-83% (25 estudios)	3522 pacientes COVID-19 4326 individuos sanos o no COVID-19	Pruebas serológicas ELISA (14) IFL (13) CLIA (13) IFI (3)	RT-PCR (13) Otras NAT / diagnóstico clínico (25)	Sangre total / suero / plasma	Sensibilidad, especificidad
Lisboa et al. 2020 (24)	Internacional	40 estudios observacionales	8295	NR	2548 participantes para ELISA 1857 participantes para IFL 3750 participantes para CLIA 140 participantes para otro tipo de pruebas	Pruebas serológicas ELISA (15) IFL (17) CLIA (13) IFI (3)	Patrón de referencia de cultivo viral o RT-PCR	Sangre total / suero / plasma	Sensibilidad, especificidad
Mekonnen et al. 2020 (25)	China	24 estudios observacionales	NA	NR	NR	Pruebas serológicas ELISA (18) IFL (16) CLIA (8) LIPS (1)	RT-PCR	Sangre total / suero / plasma	Sensibilidad, especificidad
Mustafa et al. 2020 (26)	Internacional	29 estudios observacionales (solo para COVID-19)	8742	NR	8742 personas para SARS-CoV-2	Cualquier NAT por coronavirus realizada en muestras del tracto respiratorio LAMP (15)	Cualquier NAT usado en los estudios	Frotis nasal (5) Hisopos nasofaríngeos u orofaríngeos (26) LBA (2) Esputo (2) Otros (3)	Sensibilidad, especificidad
Riccó et al. 2020 (27)	Italia	14 estudios observacionales	1118	NR	376 muestras positivas para SARS-CoV-2	RT-PCR (saliva)	RT-PCR	Saliva versus hisopos nasofaríngeos	Sensibilidad, especificidad RVP y RVN



					742 muestras negativas para SARS-CoV-2				
Wang et al. 2020 (28)	China	27 estudios observacionales	4565	Rango mediana: 25-67 años Hombres 29,5-70% (17 estudios)	2163 participantes casos COVID-19 confirmados 574 participantes con casos sospechosos, pero con PCR negativa 1828 participantes control (PCR negativos, asintomáticos, o donantes sanos)	Pruebas serológicas ELISA (13) IFL (15) CLIA (7)	Cualquier NAT usado en los estudios	Sangre total / suero / plasma	Sensibilidad, especificidad
Zhang et al. 2020 (29)	China	22 estudios observacionales	3767	NR	2282 pacientes con SARSCoV-2 1485 personas sanas o pacientes sin SARSCoV-2	Pruebas serológicas ELISA (3) CGIA (12) CLIA (9)	RT-PCR/ diagnóstico clínico	Sangre total / suero / plasma	Sensibilidad, especificidad RVP y RVN

CLIA: inmunoensayo quimioluminiscente; CGIA: inmunoensayo de oro coloidal; ELISA: Ensayos inmunoabsorbentes ligados a enzimas; IFL: inmunoensayo de flujo lateral; IFI: inmunofluorescencia indirecta; LAMP: Ensayos de amplificación isotérmica; LBA: Lavado broncoalveolar; LIPS: sistema de inmunoprecipitación de luciferasa; NAT: pruebas basadas en ácidos nucleicos; NR: No reportado; RT-PCR: Reacción en Cadena de la Polimerasa con Transcriptasa Inversa; RVP: razon de verosimilitud positiva; RVN: razon de verosimilitud negativa; VPP: valores predictivos positivos; VPV: valores predictivos negativos.



5.4. Síntesis de la evidencia

Exactitud diagnóstica de las pruebas moleculares

Dos RSL reportaron el rendimiento general de NATs. La RSL de Mustafa et al. (26) reportó una sensibilidad general de 90,4% (IC del 95%: 83,7-94,5%) y especificidad de 98,1% (IC del 95%: 95,9-99,2%) para todas las NATs.

Jarrom et al. (20) informaron a partir de 16 estudios con 3818 participantes, una sensibilidad de 87,8% (IC del 95%: 81,5% a 92,2%) para RT-PCR. Mientras que para Mustafa et al. la sensibilidad y especificidad agrupadas de la rRT-PCR fueron 96,2% (IC del 95%: 91-98,4%) y 98,5% (IC del 95%: 95,2-99,6%).

Para ensayos LAMP, los rangos de sensibilidad y especificidad estimados por Jarrom et al. fueron 74,7% a 100% y 87,7% a 100%, respectivamente. En la RSL de Mustafa et al. los LAMP, presentaron una sensibilidad y especificidad de 84,2% (IC del 95%: 75-90,5%) y 97,7% (IC del 95%: 92,8-99,3%), respectivamente.

Para otros tipos de PCR como el CRISPR (repetición palindrómica corta agrupada regularmente inter espaciada), los Mustafa et al. informaron sensibilidad de 82,7% (IC del 95%: 73,1-89,4%) y especificidad de 97,8% (IC del 95%: 93,7-99,3%).

Arevalo-Rodriguez et al. (17) reportaron la estimación agrupada de la tasa de FN de la RT-PCR en la evaluación inicial como 13% (IC del 95%: 9 a 19%). Las estimaciones de estudios individuales de la tasa de falsos negativos oscilaron entre 0,018 y 0,58. De acuerdo con lo encontrado por Arevalo-Rodriguez et al. (17) no hubo diferencias relacionadas con la duración de los síntomas en el momento de la primera prueba de RT-PCR según la información derivada de las medias y medianas de la aparición de los síntomas ni tampoco con los genes diana de la prueba.

Pruebas moleculares rápidas

En cuanto a las pruebas moleculares rápidas, Dinnes et al. (19) reportaron una sensibilidad general de 95,5% (IC del 95%: 88,5% a 98,4%) y especificidad de 98,9% (IC del 95%: 97,3% a 99,5%). Mustafa et al. (26) informaron acerca del rendimiento de la prueba molecular rápida automatizada GeneXpert; la sensibilidad estimada fue 98,9% (IC del 95%: 96,2-99,7%) y la especificidad 95,5% (IC del 95%: 91,8-97,5%).

Tipo de muestra biológica

El análisis del rendimiento de las pruebas moleculares por tipo de muestra biológica realizado por Boger et al. (21) mostró una sensibilidad de 97,2% (IC del 95% 90,3% - 99,7%) en esputo, 73,3% (IC del 95%: 68,1% -78,0%) para hisopos nasofaríngeos/frotis faríngeos, y 62,3% (IC del 95% 54,5% -69,6%) para saliva. La sensibilidad en el resto de las muestras biológicas fue muy baja. La especificidad fue estimada en 98,6% para frotis faríngeo y 90% para esputo, mientras que fue 100% para el resto de las muestras biológicas.



La RSL de Riccò et al. (27) se evaluó el rendimiento de la RT-PCR en muestras de saliva teniendo como estándar de referencia la RT-PCR en hisopos nasofaríngeos; la sensibilidad estimada fue 83,4% (IC del 95%: 73,1 a 90,4) y la especificidad 97,7% (IC del 95%: 93,8-99,2%).

En la RSL de Dinnes et al. (19) en el que evaluaron el rendimiento de las pruebas moleculares rápidas, cuando se usaron muestras nasofaríngeas solamente, la sensibilidad disminuyó a 87,1% (IC del 95%: 71,6% a 94,7%) mientras que la especificidad fue la misma para todas las muestras 98,9%.

Genes objetivo

Con respecto al número y tipos de genes para los cuales está dirigido la NAT, Mustafa et al. (26) estimó que cuando la prueba molecular está dirigida a un solo gen tiene una sensibilidad y especificidad de 82,3% (IC del 95%: 72,4-89,2%) y 97,6% (IC del 95%: 91,9-99,3%); cuando la prueba se dirige a dos o más genes la sensibilidad y especificidad aumentan a 95,6% (IC del 95%: 89,6-98,2%) y 98,5% (IC del 95%: 96,4-99,4%) respectivamente. Las pruebas dirigidas al gen E presentan mayor sensibilidad y especificidad que aquellas dirigidas a los genes N o RdRp siendo estimadas en 97,8% (IC del 95%: 95,6-98,9%) y 98,6% (IC del 95%: 93,9-99,7%) respectivamente.

Las medidas de rendimiento de las pruebas moleculares reportadas en los estudios incluidos se detallan en la tabla 3. Toda la evidencia para las NAT fue calificada como muy baja (GRADE).

Exactitud diagnóstica de las pruebas inmunológicas

Jarrom et al. (20) y Lisboa et al. (24) reportaron estimaciones generales de medidas de rendimiento para pruebas inmunológicas con rangos de sensibilidad de 18,4% a 96,1% y 93,1% a 98,6%, respectivamente. En la RSL de Jarrom et al. la sensibilidad más baja reportada (18,4%) fue la informada para una prueba serológica para aplicación en el punto de atención. El rango de especificidad reportado en estas RSL fue 88,9 a 100% y 93,3 a 100%, respectivamente.

Tipo de anticuerpo (inmunoglobulina)

La exactitud diagnóstica de las pruebas serológicas fue reportada con datos agrupados de acuerdo con el tipo de inmunoglobulina detectada en tres RSL (22,28,29).

La sensibilidad para la detección de IgG fluctuó en el rango de 76% a 85% y la especificidad en de 97 a 99%. Para la IgM los rangos de sensibilidad y especificidad fueron 69 a 97% y 95 a 99%, respectivamente. La detección de ambos anticuerpos, IgG/IgM, tuvo un rango de sensibilidad de 78 a 86% y de especificidad de 97 a 99% en el análisis agrupado de las RSL de Guo et al. (22), Wang et al. (28), y Zhang et al. (29).

Tipo de prueba serológica



Las RSL de Deeks et al. (18), Kontou et al. (23), Mekonnen et al. (25), Wang et al. (28) y Zhang et al. (29) realizaron análisis de medidas de desempeño de las pruebas serológicas de acuerdo con cada tipo de ensayo inmunológico.

Para las pruebas de ELISA, Kontou et al. (23) y Mekonnen et al. (25) reportaron sensibilidades generales de 93,5% (IC del 95%: 90,0 - 97,1%) y 86% (IC del 95%: 82% -89%), respectivamente. Los rangos de sensibilidad estimada para cada anticuerpo fueron 69 a 85,8% para IgG; 71 a 84,5% para IgM; y 69 a 90,5% para IgG/IgM (18,28,29). La especificidad general estuvo en el rango de 96 -100% (23,25) y la agrupada para cada anticuerpo fue 99% (IC del 95%: 96 a 100%) para IgG; 99,7% (IC del 95%: 99,0 a 100%) para IgM; y 100% (IC del 95%: 100 a 100%) para IgG/IgM (24,29).

Para los CLIA, la sensibilidad general estimada fue 92% (IC del 95%: 86–95%) (25) y la agrupada por tipo de anticuerpo estuvo en el rango de 90 a 94,6% para las IgG; 74 a 84% para la IgM; y 90,7 a 97,3% para IgG/IgM (18,23,25,29). La especificidad para los CLIA estuvo en el rango general de 95,4% a 99% (23,25) y por anticuerpo, 99% para IgG o IgM, y 100% para IgG/IgM (29).

Para los IFL se estimó una sensibilidad general de 78% (IC del 95%: 71–83%) (25) estando en el rango de 51, 4 a 76% para IgG o IgM, y de 80 a 86% para IgG/IgM (18,23). La especificidad general se reportó en el rango de 96 a 100% (23–25).

Para los CGIA la sensibilidad para los anticuerpos tipo IgG fue de 83 a 87%, para los tipos IgM 69,5 a 74%, y para ambos IgG/IgM 84 a 91% (18,29). La especificidad reportada por Zhang et al. (29) fue de 99% para las IgG, 97% para las IgM, y 95% para las IgG/IgM.

Wang et al. (28) calcularon un rango de especificidad general de 92% a 100% para los ELISA, CLIA e IFL.

Tipo de muestras biológicas

Una de las RSL incluidas, Boger et al. (21), realizó un análisis del rendimiento diagnóstico de las pruebas serológicas de acuerdo con el tipo de muestra evaluada para los tres tipos de anticuerpos. Los autores encontraron valores de sensibilidad más altos cuando la prueba estaba dirigida a la detección de ambos anticuerpos IgG/IgM independiente del tipo de muestra biológica siendo esta de 84,5% (95% IC: 82,2% - 86,6%) y la especificidad de 91,6% (95% IC: 86,0% -95,4%) en muestras de sangre, suero, y plasma.

Tiempo desde el inicio de los síntomas

El análisis agrupado de la dinámica de las pruebas serológicas después del inicio de los síntomas fue realizado por dos RSL; Deeks et al. (18) y Wang et al. (28).

Para la primera semana o días 0-7 desde el inicio de los síntomas la sensibilidad de las pruebas dirigidas a la IgG fue de 25 a 29,7%; a la IgM fue de 23,2 a 34%; y a IgG/IgM fue 30 a 36%.



Para la segunda semana o días 8-14 la sensibilidad fue 62 a 66,5% para las IgG; 58,4 a 65% para las IgM; y 72,2 a 80% para las IgG/IgM.

Para después de 14 días del inicio de los síntomas Wang et al. (28) reportaron una sensibilidad de 90% para las IgG; 85% para las IgM; y 93% para las IgG/IgM. Adicionalmente Deeks et al. (18) agruparon este tiempo en semanas, estimando para la tercera semana una sensibilidad de 88,2% para la IgG, 75,4% para la IgM, y 91,4% para IgG/IgM. Entre la cuarta y quinta semanas la sensibilidad de la IgG, IgM, e IgG/IgM fue 80%, 68%, y 96%, respectivamente. A partir de la quinta semana la sensibilidad para la IgG fue ~87%, para la IgM ~54%, y para las IgG/IgM ~78%.

De igual manera, Deeks et al. estimaron que para cualquier momento desde el inicio de los síntomas la especificidad de las pruebas dirigidas a la IgG era de 99,1%, a la IgM de 98,7%, y dirigida a ambas, IgG/IgM era de 98,7% (18).

Pruebas para detección de antígenos virales

Dos RSL proporcionaron información acerca de las pruebas inmunológicas dirigidas a la detección de antígenos virales (19,28). El estudio de Dinnes et al. (19) informó acerca del rendimiento de estas pruebas con una sensibilidad de 56,2% (IC del 95%: 29,5 a 79,8%) y especificidad 99,5% (IC del 95%: 98,1% a 99,9%). Cuando estas pruebas fueron realizadas en muestras nasofaríngeas solamente (3 estudios primarios), la sensibilidad y especificidad agrupadas fueron 59,4% (IC del 95%: 50,7% a 67,5%) y 99,6% (IC del 95%: 97,4% a 99,9%).

La RSL de Wang et al. (28) realizó un análisis agrupado de acuerdo con el tipo de antígeno al cual se dirige la prueba obteniendo sensibilidades de 79% (IC del 95%: 68 a 88%) para el antígeno N, 80% (IC del 95%: 62 a 92%) para el antígeno S, y de 86% (IC del 95%: 68 a 91%) para la combinación de ambos antígenos N/S.

Las medidas de rendimiento de las pruebas inmunológicas para detección de anticuerpos o antígenos virales reportadas en los estudios incluidos se detallan en la tabla 3. Toda la evidencia para las pruebas inmunológicas para detección de anticuerpos o antígenos virales fue calificada como muy baja o baja (GRADE).



Tabla 3. Resumen de los resultados informados en las RSL y RSL-metanálisis incluidas

Estudio	Prueba	Sensibilidad		Especificidad	Otras medidas (VP, VN, FP, FN, RVP, RVN, VPP, y VPN)
Arevalo-Rodríguez et al. 2020 (17)	Hallazgos RT-PCR negativos en la evaluación inicial				La tasa de FN fue 13% (IC del 95%: 9 a 19%).
Boger et al. 2021 (21)	Todas la puebas serológicas	Sangre	IgG: 66,1% (95% IC: 62,3% -69,8%)	98,8% (95% IC: 95,8% -99,9%)	RVP: 26,981 (6,240-116,655) RVN: 0,377 (0,128-1,113)
			IgM: 78,8% (95% IC: 75,4% -81,9%)	93,1% (95% IC: 88,2% -96,4%)	RVP: 8,390 (3,367-20,905) RVN: 0,274 (0,072-1,043)
			IgG/IgM: 86,3% (95% IC: 83,3% -88,8%)	90,7% (95% IC: 84,8% -94,8%)	RVP: 8,618 (5,219-14,231) RVN: 0,146 (0,021-1,028)
		Suero	IgG: 73,9% (95% IC: 69,6% -77,9%)		
			IgM: 74,3% (95% IC: 70,1% -78,2%)		
			IgG/IgM: 82% (95% IC: 78% -85%)		
		Sangre, suero, plasma	IgG: 69,4% (95% IC: 66,6% -72,1%)	69,4% (95% IC: 66,6% -72,1%)	RVP: 25,626 (7,131-92,087) RVN: 0,378 (0,128-1,111)
			IgM: 77% (95% IC: 74,5% -79,5%)	93,3% (95% IC: 88,6% -96,5%)	RVP: 7,295 (3,403-15,641) RVN: 0,211 (0,067-0,666)



			IgG/IgM: 84,5% (95% IC: 82,2% -86,6%)	91,6% (95% IC: 86,0% -95,4%)	RVP: 7,604 (3,903-14,817) RVN: 0,170 (0,041-0,697)
	RT-PCR	Hisopos rectales / heces (24,1%, IC del 95%: 16,7% -33,0%)		La especificidad fue del 100% para las muestras de heces, orina, sangre, frotis nasal y de garganta Las muestras de frotis de garganta y de esputo tuvieron especificidades del 98,6% y el 90,0%.	
		Orina (0,0%, IC 95% 0,0% -3,7%)			
		Plasma (7,3%, IC del 95%: 4,1% -11,7%)			
		Esputo (97,2%, IC del 95% 90,3% -99,7%)			
		Saliva (62,3%, IC del 95% 54,5% -69,6%)			
		Aspirado / hisopado nasofaríngeo y frotis faríngeo (73,3%, IC del 95%: 68,1% -78,0%)			
Deeks et al. 2020 (18)	Todas la puebas serológicas	Primeras 3 semanas desde el inicio de los síntomas	IgG: Semana 1: 29,7% (IC del 95%: 22,1 a 38,6) Semana 2: 66,5% (IC del 95%: 57,9 a 74,2) Semana 3: 88,2% (IC del 95%: 83,5 a 91,8)	En cualquier momento de inicio de los síntomas IgG: 99,1% (IC del 95%: 98,3 a 99,6) IgM: 98,7% (IC del 95%: 97,4 a 99,3) IgG/IgM: 98,7% (IC del 95%: 97,2 a 99,4)	
			IgM:		



				Semana 1: 23,2% (IC del 95%: 14,9 a 34,2) Semana 2: 58,4% (IC del 95%: 45,5 a 70,3) Semana 3: 75,4% (IC del 95%: 64,3 a 83,8)		
				IgG/IgM: Semana 1: 30,1% (IC del 95%: 21,4 a 40,7) Semana 2: 72,2% (IC del 95%: 63,5 a 79,5) Semana 3: 91,4% (IC del 95%: 87,0 a 94,4)		
			Semanas 4 -5	IgG: 80,3% (IC del 95%: 72,4 a 86,4) IgM: 68,1% (IC del 95%: 55,0 a 78,9) IgG/IgM: 96,0% (IC del 95%: 90,6 a 98,3).		
			Despues de semana 5	IgG: 86,7% (IC del 95%: 79,6 a 91,7) IgM: 53,9% (IC del 95%: 38,4 a 68,6) IgG/IgM: 77,7% (IC del 95%: 66,0 a 86,2)		
	Tipo de prueba serológica	CLIA	IgG: 94,6%; IgM: 80,9%; IgG/IgM: 97,3%		Las diferencias en la especificidad de IgG e IgM entre los tipos de ensayo fueron pequeñas, las pruebas CLIA y CGIA mostraron una menor especificidad para las IgG/IgM que ELISA e IFL	
		CGIA	IgG: 87,3%; IgM: 69,5%; IgG/IgM: 91,4%			
		ELISA	IgG: 85,8%; IgM: 84,5%; IgG/IgM: 90,5%			



		IFL	IgG: 76%; IgM: 51,4%; IgG/IgM: 85,8%			
Dinnes et al. 2020 (19)	Pruebas de antígenos		56,2% (IC del 95%: 29,5 a 79,8%) En muestras nasofaríngeas solamente (3 estudios): 59,4% (IC del 95%: 50,7% a 67,5%)	99,5% (IC del 95%: 98,1% a 99,9%) En muestras nasofaríngeas solamente: 99,6% (IC del 95%: 97,4% a 99,9%)		
	Pruebas moleculares rápidas		95,5% (IC del 95%: 88,5% a 98,4%) En muestras nasofaríngeas solamente (6 estudios): 87,1% (IC del 95%: 71,6% a 94,7%)	98,9% (IC del 95%: 97,3% a 99,5%) En muestras nasofaríngeas solamente: No diferencia		
Guo et al. 2020 (22)	Todas la puebas serológicas		IgM: 71% (IC del 95%: 69-73%)	IgM: 97% (IC del 95%: 96-97%)	RVP de casos confirmados	IgM: 61,2% (IC del 95%: 53,4% -69,0%)
						IgG: 58,8% (IC del 95%: 49,6% -68,0 %)
						IgM/IgG: 62,1% (52,7% -71,4%)
			IgG: 76% (IC del 95%: 75-78%)	IgG: 97% (IC del 95%: 96-97%)	RVP de casos sospechosos	IgM: 29,0% (IC del 95%: 14,0% - 44,0%)
						IgG: 37,0% (IC del 95%: 20,0% - 55,0%)
						IgM/IgG: 55,0% (IC del 95%: 19,0%-90,0%)



			IgM/IgG: 82% (IC del 95%: 81-84%)	IgM/IgG: 97% (IC del 95%: 96-97%)	RVP de casos asintomáticos	IgM/IgG: 19% (IC del 95%: 10,0% -27,0%)
Jarrom et al. 2020 (20)	Pruebas de detección del virus	RT-PCR	87,8% (IC del 95%: 81,5% a 92,2%)		VPP 57,7% y VPN 99,6% (con prevalencia de 3,0%)	
		LAMP	Rango de 74,7% a 100%	Rango de 87,7% a 100%	VPP 93,5% y VPN 96,1% (con prevalencia de 24,6%)	
	Pruebas de anticuerpos		Rango de 18,4% a 96,1%	Rango de 88,9% a 100%		
Kontou et al. 2020 (23)	Pruebas de anticuerpos	ELISA	93,5% (IC del 95%: 90,0 - 97,1%)	Rango desde 96,1% a 99,5%		
		CLIA	IgG 94,4%; IgM 81%; IgG/IgM: 90,7% (IC del 95%: 75,3 - 100%)	Rango desde 95,4% a 98,4%		
		IFL	IgG o IgM solos: rango 53-66%; IgG/IgM: 80% (IC del 95%: 66,3 - 93,5%)	98,4% (IC del 95%: 96,9 – 99,9%). Rango desde 91,4% a 99,4%		
		IFI	IgG o IgM solos: ~86%	IgG o IgM solos: ~95%		
Lisboa et al. 2020 (24)	Todas la puebas serológicas		Rango desde 93,1% a 98,6%	Rango desde 93,3% a 100%		
				IFL (IgG e IgM): 96,6% (IC del 95%: 94,3% a 98,2%)		



					ELISA (IgM): 99,7% (IC del 95%: 99,0 a 100%)	
Mekonnen et al. 2020 (25)	Puebas serológicas	CLIA	92% (IC del 95%: 86–95%)	IgG: 92% (IC del 95%: 91 a 98%)	99% (IC del 95%: 97-99%)	
				IgM: 84% (IC del 95%: 67 a 93%)		
		ELISA	86% (IC del 95%: 82% -89%)		Rango desde 96% a 100%	
		IFL	78% (IC del 95%: 71–83%)			
Mustafa et al. 2020 (26)	Todos NAT		90,4% (IC del 95%: 83,7-94,5%)		98,1% (IC del 95%: 95,9-99,2%)	
	Tipo de NAT	PCR	82,7% (IC del 95%: 73,1-89,4%)		97,8% (IC del 95%: 93,7-99,3%)	
		rRT-PCR	96,2% (IC del 95%: 91-98,4%)		98,5% (IC del 95%: 95,2-99,6%)	
		LAMP/ amplificación isotérmica	84,2% (IC del 95%: 75-90,5%)		97,7% (IC del 95%: 92,8-99,3%)	
		GeneXpert	98,9% (IC del 95%: 96,2-99,7%)		95,5% (IC del 95%: 91,8-97,5%)	



		Un solo gen objetivo	82,3% (IC del 95%: 72,4-89,2%)	97,6% (IC del 95%: 91,9-99,3%)	
		Más de un gen objetivo	95,6% (IC del 95%: 89,6-98,2%)	98,5% (IC del 95%: 96,4-99,4%)	
		Gen E incluido en la prueba	97,8% (IC del 95%: 95,6-98,9%)	98,6% (IC del 95%: 93,9-99,7%)	
		Gen N incluido en la prueba	93,9% (IC del 95%: 86,5-97,3%)	98,2% (IC del 95%: 95,8-99,3%)	
		Solo gen RdRp como objetivo	77% (IC del 95%: 65,7-85,4%)	97,5% (IC del 95%: 84,6-99,6%)	
Riccó et al. 2020 (27)	RT-PCR saliva		83,4% (IC del 95%: 73,1 a 90,4) Rango desde 31,3% a 100%	97,7% (IC del 95%: 93,8-99,2%) Rango desde 71,4% a 100%	RVP: 20,141 (IC del 95%: 8,207 a 49,430) RVN: 0,203 (IC del 95%: 0,094 - 0,436)
Wang et al. 2020 (28)	Todas las Puebas serológicas		IgG 76% (IC del 95%: 65 a 86%)	IgG 98% (IC del 95%: 96 a 99%)	
			IgM 69% (IC del 95%: 59 a 78%)	IgM 95% (IC del 95%: 91 a 98%)	
			IgG/IgM 78% (IC del 95%: 70 a 85%).	IgG/IgM 97% (IC del 95%: 93 a 99%)	
		Días 0-7 del inicio de síntomas	IgG 25% (IC del 95%: 16 a 36%)		



				IgM 34% (IC del 95%: 25 a 42%)		
				IgG/IgM 36% (IC del 95%: 28 a 43%).		
			Días 8-14 del inicio de síntomas	IgG 62% (IC del 95%: 52 a 71%)		
				IgM 65% (IC del 95%: 36 a 86%)		
				IgG/IgM 80% (IC del 95%: 69 a 99%).		
			Despues de 14 días de inicio de síntomas	IgG 90% (IC del 95%: 86 a 93%)		
				IgM 85% (IC del 95%: 68 a 95%)		
				IgG/IgM 93% (IC del 95%: 80 a 98%).		
	Tipo de prueba serológica	ELISA	IgG 70% (IC del 95%: 55 a 84%)		Rango desde 92% a 100% (ELISA, CLIA, IFL)	
			IgM 78% (IC del 95%: 70 a 85%)			
			IgG/IgM 86% (IC del 95%: 62 a 98%).			



		IFL	IgG 69% (IC del 95%: 50 a 85%)			
			IgM 63% (IC del 95%: 44 a 79%)			
			IgG/IgM 70% (IC del 95%: 61 a 80%).			
	Tipo de antígeno objetivo	Antígeno N	79% (IC del 95%: 68 a 88%)			
		Antígeno S	80% (IC del 95%: 62 a 92%)			
		Antígenos N/S	86% (IC del 95%: 68 a 91%)			
Zhang et al. 2020 (29)	Todas las Puebas serológicas	IgG 85% (IC del 95%: 79 a 90%)		IgG 99% (IC del 95%: 98 a 100%)	RVP	IgG 88,32 (IC del 95%: 38,48 - 229,57)
						IgM 71,41 (IC del 95%: 22,09 - 259,48)
						IgG/IgM 104,14 (IC del 95%: 24,99 - 456,71).
		IgM 74% (IC del 95%: 65 a 81%)		IgM 99% (IC del 95%: 97 a 100%)	RVN	IgG 70,15 (IC del 95%: 0,10 - 0,22)



			IgG/IgM 86% (IC del 95%: 79 a 92%).	IgG/IgM 99% (IC del 95%: 97 a 100%).		IgM 0,27 (IC del 95%: 0,18 - 0,36)
						IgG/IgM 0,14 (IC del 95%: 0,08 - 0,22).
	Tipo de prueba serológica	CGIA	IgG 83% (IC del 95%: 73 a 90%)	IgG 99% (IC del 95%: 96 a 100%)		
			IgM 74% (IC del 95%: 60 a 85%)	IgM 97% (IC del 95%: 93 a 99%)		
			IgG/IgM 84% (IC del 95%: 78 a 90%).	IgG/IgM 95% (IC del 95%: 93 a 98%).		
		CLIA	IgG 90% (IC del 95%: 84 a 95%)	IgG 99% (IC del 95%: 97 a 100%)		
			IgM 74% (IC del 95%: 60 a 85%)	IgM 99% (IC del 95%: 97 a 100%)		
			IgG/IgM 96% (IC del 95%: 91 a 98%).	IgG/IgM 100% (IC del 95%: 100 a 100%).		
		ELISA	IgG 69% (IC del 95%: 48 a 85%)	IgG 99% (IC del 95%: 96 a 100%)		
			IgM 71% (IC del 95%: 40 a 91%)	IgM 100% (IC del 95%: 100 a 100%)		



			IgG/IgM 69% (IC del 95%: 50 a 85%).	IgG/IgM 100% (IC del 95%: 100 a 100%)	
--	--	--	-------------------------------------	---------------------------------------	--

CLIA: inmunoensayo quimioluminiscente; CGIA: inmunoensayo de oro coloidal; Elisa: Ensayos inmunoabsorbentes ligados a enzimas; IFL: inmunoensayo de flujo lateral; IFI: inmunofluorescencia indirecta; LIPS: sistema de inmunoprecipitación de luciferasa; RT-PCR: Reacción en Cadena de la Polimerasa con Transcriptasa Inversa; rRT-PCR: Reacción en Cadena de la Polimerasa con Transcriptasa Inversa en tiempo real; NAT: pruebas basadas en ácidos nucleicos; RVP: razon de verosimilitud positiva; RVN: razon de verosimilitud negativa; VPP: valores predictivos positivos; VPN: valores predictivos negativos; TR: tracto respiratorio; IC: Intervalo de confianza



6. Discusión

El SARS-CoV-2, que ha causado la pandemia de COVID-19, ha infectado a más de 95 millones de personas en todo el mundo y causado poco más de 2 millones de muertes (1). Por tanto, continúa siendo una grave amenaza para la vida humana a nivel global. En la actualidad (enero 20/2021) Colombia presenta una alta prevalencia mayor al 20% en las principales ciudades del país (30,31). A pesar de que se han realizado numerosas y valiosas investigaciones sobre este virus, aún existen muchas ambigüedades acerca de la exactitud diagnóstica de las pruebas disponibles en la actualidad.

Esta revisión se basa en RSL y RSL-metanálisis publicados, o aceptados para publicación disponibles hasta el 18 de diciembre de 2020. Seis de los estudios evaluaron el rendimiento de pruebas diagnósticas moleculares (17,19–21,26,27), mientras que nueve de las RSL incluidas evaluaron el desempeño de pruebas serológicas (18,20–25,28,29). Para todos los estudios incluidos la calidad de la evidencia fue calificada como baja o muy baja, siendo esta última la calificación para la mayoría.

En general, para las pruebas de detección del virus a través de amplificación de ácido nucleico (NAT) (moleculares), la sensibilidad mostró importante imprecisión, pero una alta especificidad diagnóstica. El estudio de Mustafa et al. (26) mostró una sensibilidad y especificidad agrupadas de 90,4% (IC95%: 83,7-94,5%) y 98,9% teniendo como comparador la RT-PCR más criterios clínicos. Jarrom et al. (20) reportó una sensibilidad de 87,8% (IC95%: 81-94%) para la RT-PCR contrastada con una segunda RT-PCR como referencia. Adicionalmente estimaron que el VPN puede ser alto mientras el VPP puede ser bajo en momentos en que la prevalencia en la población analizada es baja; el VPP considerando una prevalencia de 3% se estima en 57,7%, mientras que para una prevalencia de 24,6% el VPP de las NAT sería de 93,5%. Al respecto los autores argumentan que a pesar del hallazgo de un VPN alto (99%) para la RT-PCR, la incertidumbre puede permanecer con un resultado negativo de la prueba, especialmente en el contexto de alta sospecha clínica, y también debe considerarse la posibilidad de un resultado FN (20).

Arevalo-Rodriguez et al. (17) usaron el intervalo predictivo estimado en su propio análisis para demostrar cómo con el rango de tasas de FN encontrado (18-58%) para la RT-PCR, en una población hipotética de 100 personas analizadas y con una estimación de prevalencia del 50%, hasta 27 casos se diagnosticarían erróneamente. De igual manera, Mustafa et al. (26) encontró que bajo el supuesto de una población de 1000 analizados y una prevalencia de 50%, los FN serían alrededor de 48 casos (27-81). Las posibles causas de las pruebas FN incluyen error de laboratorio, error de muestreo y variabilidad en la diseminación viral con la falta o presencia insignificante de ácido nucleico del virus en el tejido recogido en el momento del muestreo (17,20,26).

En cuanto al tipo de muestras para las pruebas moleculares, teniendo en cuenta que las NATs constituyen la principal prueba de diagnóstico para la infección por SARS-CoV-2, la elección del tipo de muestra es un paso importante para un diagnóstico exitoso. El esputo, el aspirado / hisopado nasofaríngeo y el frotis de garganta mostraron ser más



sensibles para detectar el virus (21,27). La RT-PCR en muestras de saliva puede ser altamente específica, pero su sensibilidad es considerablemente menor que en las muestras mencionadas (21,27). Las muestras sanguíneas para RT-PCR presentan sensibilidad mucho menor y mayor imprecisión (21). La RT-PCR en la RSL de Boger et al. (21) mostró sensibilidades agrupadas de 97,2% para muestras de esputo, y de 73% para saliva. La sensibilidad para las muestras de sangre fue muy imprecisa (IC95%: 41-117%). En muchas de las RSL incluidas no fue posible realizar análisis estadísticos sobre los parámetros de especificidad debido al número limitado de estudios de RT-PCR con un grupo control. Comparativamente, teniendo en cuenta los datos reportados por Boger et al. (21), considerando una prevalencia de 50% y el análisis de 1000 muestras, se podría esperar FN en ~14 casos en muestras de esputo, ~133 en muestras de saliva, y ~463 en sangre. En la RSL de Riccò et al. (27) se analizó el rendimiento de la RT-PCR en muestras de saliva teniendo como referencia la misma prueba en muestras de hisopo nasofaríngeo y encontraron una sensibilidad agrupada de 83,4% con un amplio intervalo de confianza (73-90%) y una especificidad de 97,7%. Bajo el mismo supuesto de 1000 muestras y una prevalencia de 50%, se esperaría tener de 48 a 134 resultados FN. Mustafa et al. (26) igualmente encontraron mayor sensibilidad de las pruebas moleculares con el uso de esputo que en muestras respiratorias superiores (95,8% versus 88%). Los resultados mencionados muestran que, entre todos los métodos, la técnica de RT-PCR con muestras de esputo fue el método más sensible para el diagnóstico de COVID-19. La comparación de los resultados de diferentes muestras clínicas demostró que las muestras respiratorias mixtas (tracto respiratorio bajo y alto), son las más adecuadas para lograr índices de sensibilidad más altos. Riccò et al. (27), mencionan que “a pesar de su condición de estándar de oro, las muestras nasofaríngeas están lejos de ser una base óptima para un análisis RT-PCR confiable”. Sin embargo, también sugieren ser cuidadosos con esta interpretación debido a que su análisis no tuvo en cuenta el posible retraso entre la realización de la prueba y el inicio informado de los síntomas, ya que esta información no está disponible regularmente en los estudios primarios (21,26,27).

Las pruebas moleculares rápidas para uso en el punto de atención como la ID NOW (Abbott, Chicago, IL) o la Xpert Xpress (Cepheid, Sunnyvale, CA), tal como ocurre con las RT-PCR convencionales, mostraron alta imprecisión en la sensibilidad. En la RSL de Dinnes et al. (19) la sensibilidad agrupada se estimó en 68-100%. En este análisis, las especificidades fueron consistentemente altas (92% a 100%, con límites superiores de IC del 95% del 99% o 100% en cada estudio). La estimación esperada de FN para 1000 muestras con una prevalencia de 50% sería de 0 a 160 casos para ambas pruebas rápidas. Sin embargo, según el análisis de Dinnes et al., la prueba Xpert Xpress presentó mejor sensibilidad (99,4%) y en el caso hipotético de una prevalencia de 10% se esperaría solo un (n=1) caso FN. Esta sensibilidad fue 22,6 puntos porcentuales más alta que la de ID NOW. Los autores argumentaron que debido a que los estudios que evaluaron ambas pruebas tenían limitaciones metodológicas similares, esta diferencia posiblemente se mantenga cuando esas limitaciones no están presentes. La diferencia de especificidad entre las pruebas es pequeña (ID NOW es un 2,8% más específica en comparación con Xpert Xpress), pero potencialmente importante, especialmente si se



utiliza en un entorno de baja prevalencia. Sin embargo, esto no sería un problema si las pruebas positivas fueran confirmadas por un ensayo de RT-PCR de laboratorio (19).

Con respecto a otros tipos de NAT, la sensibilidad igualmente muestra alta imprecisión. Para las pruebas LAMP, Jarrom et al. (20) encontraron un rango de sensibilidad agrupada de 74,7-100% y especificidad de 88,7-100% en cinco estudios analizados que incluían diferentes tipos de PCR como referencia. Mustafa et al. (26) incluyeron 15 estudios que evaluaron el rendimiento de pruebas LAMP pero también todos los estudios utilizaron una PCR diferente como estándar de referencia. Las RT-LAMP o ensayos isotérmicos, evaluados en los estudios, resultaron en una sensibilidad más baja (84,2%) que GeneXpert 98,9% u otras NAAT (93,8%), pero todas con alta especificidad. Los autores consideran que, debido a la alta heterogeneidad de los estudios y la falta de estandarización del estándar de referencia utilizado, no era apropiado realizar un análisis conjunto de la precisión del diagnóstico, lo que significa que se pueden extraer conclusiones limitadas sobre las pruebas LAMP en función de los datos disponibles (20,26).

Finalmente, en relación con el gen diana, en el estudio de Mustafa et al. (26) se evidencio que las pruebas dirigidas al gen N o al gen E (97,8% y 93,9%) tenían una sensibilidad más alta que otras pruebas, mientras que el gen RdRp, siempre evaluado con RT-LAMP, tenía una sensibilidad significativamente menor (77%). Adicionalmente, las pruebas dirigidas a más de un gen tenían mejor sensibilidad que las pruebas dirigidas a un solo gen (95,6% versus 82,3%).

Para los resultados de rendimiento de ensayos inmunológicos que constituyen las pruebas de detección de anticuerpos (serológicas) o de detección de antígenos virales, se contó con un mayor número de estudios secundarios, pero con mucha heterogeneidad en los estudios primarios reportados. Todos los estudios contemplaron la RT-PCR como prueba de referencia (18–25,28,29). RSL como la de Jarrom et al. (20) y Lisboa et al. (24) reportaron rangos de sensibilidad muy diferentes para las pruebas de anticuerpos; 18-96% y 93-98,6% respectivamente. Mientras que los rangos de especificidad fueron menos variables; 88-100% y 93-100% respectivamente.

Por lo general, la respuesta inmune del cuerpo a un patógeno tarda de una a dos semanas en ocurrir. En este contexto, el uso de pruebas serológicas para la detección en la fase inicial / aguda de la enfermedad puede ser un desafío. Adicionalmente, con los estudios realizados se ha logrado evidenciar, según las mediciones cinéticas de estos, que la IgM comienza a ser detectada a los días 5, alcanza su pico entre los 5 y 12 días, y luego desciende lentamente durando hasta un mes. Las IgG comienzan a ser detectadas desde el día 14 y alcanzan concentraciones máximas después del día 20 aproximadamente cuando los anticuerpos IgM están desapareciendo o han desaparecido (23–25,28,29). En consonancia con esto, es de esperarse que para las pruebas serológicas, en general, la sensibilidad fuera mayor cuando se evaluó la combinación de anticuerpos IgG e IgM en un rango de 78-86% (18,21,22,29). Las especificidades de todas las pruebas serológicas para la detección de anticuerpos totales (IgG/IgM) en estos estudios estuvo en el rango de 97-99% (18,21,22,29). Boger



et al. (21) encontraron mejor RVN cuando se evaluaron los anticuerpos totales. Con las estimaciones suministradas por este estudio, en una muestra de 1000 pruebas bajo una prevalencia de 50% los FN y FP que se obtendrían evaluando anticuerpos totales serían ~77 y ~42, respectivamente.

Es precisamente el tiempo desde el inicio de los síntomas que es importante para considerar la realización de este tipo de pruebas. Los estudios de Deeks et al.(18) y Wang et al.(28) hicieron un análisis muy completo teniendo en cuenta estos tiempos. Para la primera semana (0-7 días), de nuevo la mejor sensibilidad reportada fue para la detección de total de anticuerpos (IgG/IgM) con un rango de 30-36%. Para la segunda (8-14 días) y tercera (14-21 días) semanas las sensibilidades fueron 72-80% y 91-93%, respectivamente. Adicionalmente, Deeks et al. (18) proporcionó análisis para las semanas 4-5 y después de la quinta semana encontrando sensibilidad para anticuerpos totales se mantiene elevada entre la cuarta y quinta semana (96%) pero disminuye después (77,7%). Para esta RSL la especificidad reportada para IgG/IgM fue ~98% en cualquier momento de la toma de muestra. La utilidad de esta prueba de detección de anticuerpos totales como complemento a la RT-PCR negativa cuando existe alta sospecha, se observa con los potenciales FP y FN ante la situación hipotética de una prevalencia de 50% y una muestra de 1000 pacientes, calculados en ~10 y ~139 en la segunda semana (8-14 días) (18).

Al no contar con pruebas serológicas para anticuerpos totales, la mejor sensibilidad reportada para cualquiera de los dos tipos de anticuerpos, IgG o IgM ocurre alrededor de la tercera semana (88-90%) para la IgG y (85-91%) para la IgM (18,28). A partir de la cuarta semana la sensibilidad de las pruebas dirigidas a la IgM disminuye considerablemente (~54%) mientras que para la IgG se mantiene alta (~86%). Como la sensibilidad de las pruebas se evaluó principalmente en pacientes hospitalizados, los autores argumentan que no está claro si las pruebas tienen la capacidad de detectar niveles más bajos de anticuerpos probablemente en pacientes COVID-19 no hospitalizados (18). Zhang et al. (29) sugieren además que el rendimiento de la detección de IgG/IgM anti SARS-CoV-2 para portadores asintomáticos no es claro y que posiblemente los hallazgos de los estudios no son aplicables a poblaciones asintomáticas.

En cuanto a inmunoensayos para la detección de antígenos virales, la RSL de Dinnes et al. (19) evaluó estudios muy heterogéneos de pruebas realizadas en el punto de atención detectando una sensibilidad general agrupada de 56% que aumentó a 59% cuando solo se trataba de muestras nasofaríngeas (3 estudios). La especificidad fue de 99% para todos los tipos de muestras. Los autores sugieren que el valor promedio de sensibilidad debe interpretarse con precaución, ya que observaron grandes diferencias en la sensibilidad según la carga viral y posiblemente las diferencias en la distribución de muestras con carga viral alta y baja entre los estudios pueden haber afectado las estimaciones de precisión general. Combinado con limitaciones metodológicas y otros factores desconocidos, no es posible afirmar con certeza si alguna prueba es superior a las demás (19). El estudio de Wang et al. (28) evaluó la sensibilidad de estas pruebas de acuerdo con el tipo de antígeno diana, mostrando la sensibilidad más alta cuando la prueba se dirige a ambos antígenos N/S (86%).



Con respecto al tipo de prueba, los reportes de sensibilidad para la detección de anticuerpos fueron muy heterogéneos (18,23,25,28,29). De las evaluaciones estudiadas, se observaron algunas diferencias por tecnología de prueba en las que los métodos CLIA parecieron más sensibles que los ELISA o los IFL basados en CGIA para IgG/IgM. Las sensibilidades reportadas en los diferentes estudios para anticuerpos totales estuvieron en el rango de 91-97% para los CLIA, 70-86% para los IFL (pero 84-91% para los CGIA), y 69-90,5% para los ELISA. Las especificidades fueron en general muy similares y altas 95-100%. Wang et al. (28) sugieren que a pesar de que los IFL presentan una sensibilidad relativamente más baja que los ELISA o CLIA, son convenientes y pueden proporcionar un tiempo de respuesta rápido. La elección de la metodología de la prueba serológica debe basarse en el entorno de la prueba y la población de pacientes. Las pruebas IFL, incluyendo las CGIA, podrían resultar útiles en la sala de emergencias y ámbito ambulatorio en puntos de atención en lugar de simplemente ser abandonadas por su desempeño relativamente deficiente. En contraste, con estos resultados, Lisboa et al. (24) argumentaron que las pruebas serológicas disponibles tienen una utilidad limitada en el diagnóstico del COVID-19 agudo dado que de aquellos a los que se les hace la prueba en la primera semana de la aparición de los síntomas, en promedio, entre el 44% y 87% serán identificados falsamente como no infectados. Los autores también plantean que estos hallazgos también deberían dar una pausa a los gobiernos que están contemplando el uso de pruebas serológicas, en particular, pruebas en el lugar de atención, para emitir "certificados" o "pasaportes" de inmunidad. Por ejemplo, si se aplica un IFL a una población con una prevalencia real de SARS-CoV-2 del 10%, por cada 1000 personas analizadas, a 31 que nunca tuvieron COVID-19 se les informará incorrectamente que son inmunes y a 34 personas que tuvieron COVID-19 se le dirá incorrectamente que nunca se infectaron (24).

Esta revisión sistemática se planeó inicialmente como una actualización de una RSL previa con el objetivo de brindar la nueva evidencia disponible. Sin embargo, debido al creciente número de publicaciones potencialmente con mejor calidad metodológica de las que inicialmente se habían incluido para responder la pregunta de la revisión, el enfoque evolucionó hacia una revisión sistemática de RSL y metanálisis publicados o aceptados para su publicación después de pasar por el proceso de evaluación por pares. Por lo tanto, este manuscrito refleja la actualización de nuestras búsquedas bibliográficas enfocadas a estudios secundarios con información actualizada hasta enero de 2021.

Los puntos fuertes de este estudio son la inclusión de un buen número de revisiones sistemáticas, una selección sólida, la extracción de datos y la evaluación metodológica crítica. Sin embargo, es necesario resaltar algunas debilidades al interpretar los hallazgos de esta RSL de revisiones sistemáticas.

Como se señaló en las secciones de evaluación de la calidad metodológica de las RSLs y de la calidad de la evidencia, se encontró deficiencia metodológica en el diseño de los estudios primarios, un alto riesgo de sesgo en el muestreo y selección de pacientes, así como en la interpretación de resultados con respecto a la prueba de referencia, y una heterogeneidad considerable en las medidas de rendimiento de las pruebas explicada



de manera insuficiente por los metanálisis realizados. Las fuentes sugeridas de heterogeneidad, como el tipo de pacientes y muestra recolectada, el tipo de prueba, y el tiempo hasta la aparición de los síntomas, fueron informadas parcialmente o en absoluto en los estudios primarios. Adicionalmente, el sesgo de publicación estuvo muy presente en todas las RSL puesto que la mayoría de los estudios con número considerable de muestra provienen de China y Estados Unidos y son publicados en los idiomas chino o inglés.

Esta variabilidad en los datos de las pruebas de COVID-19 y el desafío de proporcionar una estimación combinada con un significado clínico útil continúa siendo la principal limitación en el desarrollo de este tipo de revisiones sistemáticas que probablemente se ven afectadas por varios problemas no informados, la falta de métodos estandarizados para las pruebas de COVID-19 y, en general, el conocimiento limitado sobre la enfermedad.



7. Conclusiones

Según los resultados de las RSL incluidas, el rendimiento de las pruebas diagnósticas para COVID-19 es muy heterogéneo y pueden ser condicionado por múltiples factores. No está claro si las limitaciones de los estudios primarios pudieron conducir a estimaciones excesivas o insuficientes de la exactitud de las pruebas, por lo que todos los resultados informados deben interpretarse con un alto grado de precaución. La mayoría de los estudios utilizaron un muestreo sesgado basado en la presencia o ausencia de infección por COVID-19 confirmada y seleccionaron muestras de las enviadas a los laboratorios para las pruebas de rutina de RT-PCR con poco o ningún detalle de los participantes que proporcionaron las muestras en relación con el estado de los síntomas o el tiempo desde el inicio de estos.

Las conclusiones principales que pueden hacerse dadas estas limitaciones son las siguientes:

- Los hallazgos de refuerzan la necesidad de realizar NAT repetidas en pacientes con sospecha de infección, ya sea por razones clínicas o epidemiológicas, dado que una alta tasa de los pacientes con COVID-19 pueden tener un resultado inicial de RT-PCR negativo y no recibirían un tratamiento clínico adecuado y podrían requerir pruebas repetidas en algún momento de su hospitalización o complementar con otras pruebas diagnósticas como los ensayos inmunológicos.
- La RT-PCR sigue siendo el estándar de oro para el diagnóstico de COVID-19 en muestras de esputo. Sin embargo, según el tipo de muestra y el estadio de la enfermedad, son preferibles otros métodos. Se recomienda una combinación de pruebas de diagnóstico clínicas, moleculares y serológicas para lograr una sensibilidad y especificidad adecuadas. Siempre que sea posible, se deben utilizar pruebas moleculares dirigidas a 2 genes para mejorar el rendimiento analítico.
- El uso de muestras salivales para estudios de RT-PCR en casos sospechosos de COVID-19 es actualmente cuestionable para propósitos clínicos y no puede sustituir la recolección convencional de muestras respiratorias. Aún se requieren más estudios. Para examinar la sensibilidad y la especificidad, en lugar de la concordancia de la prueba, se debe definir un estándar de referencia optimizado que pueda usarse de manera consistente en estudios futuros.
- Las pruebas moleculares rápidas podrían usarse como reemplazo de la RT-PCR del laboratorio, si es lo suficientemente exacta. La evidencia incluida hasta la fecha sugiere que algunas pruebas moleculares rápidas pueden tener niveles de precisión que se aproximan a los de la RT-PCR basada en laboratorio. Sin embargo, muchos de los datos provienen de estudios de dos grupos con sobre muestreo sesgado de casos y preocupaciones sobre la aplicabilidad de la evidencia. No hay certeza de si alguna prueba se desempeña suficientemente bien para esta función.
- Alternativamente, las pruebas en serie (durante varios días) o combinaciones de diferentes pruebas rápidas (por ejemplo, una prueba de antígeno seguida de una prueba molecular rápida) en la misma muestra pueden proporcionar una



estrategia de prueba útil; sin embargo, se necesitarían pruebas adicionales del rendimiento de dicha estrategia de diagnóstico. En ausencia de pruebas adicionales, y en entornos de baja prevalencia, los resultados positivos y negativos de cualquiera de las pruebas moleculares rápidas en el lugar de atención deberían ser seguidas con una RT-PCR basada en laboratorio.

- Las pruebas moleculares rápidas, o de detección de antígenos virales o anticuerpos en el punto de atención podrían usarse como triaje a la RT-PCR, que permite la detección temprana y el manejo rápido (autoaislamiento, cuarentena o intervención terapéutica) de los que dan positivo esperando el resultado de la RT-PCR de laboratorio. Según la evidencia del rendimiento, las pruebas rápidas solo podrían realizar este tipo de función de clasificación en entornos de mayor prevalencia (es decir, 20% o más) debido al riesgo (relativo) de resultados falsos positivos en condiciones de baja prevalencia.
- En pacientes que presentan síntomas de sospecha aguda de COVID-19, las pruebas de anticuerpos no tienen ningún papel por sí mismas como prueba principal para usar en el diagnóstico de COVID-19 cuando los pacientes se presentan durante la primera semana desde el inicio de los síntomas, ya que su sensibilidad es demasiado baja. Por lo tanto, puede ser útil el uso de pruebas de anticuerpos en pacientes con sospechas fuertes de COVID-19 pero RT-PCR negativos después de dos semanas desde el inicio de los síntomas.
- De ser posible, de acuerdo con los hallazgos, debe preferirse cualquier tipo de prueba serológica dirigida a la detección de anticuerpos totales (IgG/IgM) considerando los tiempos desde la exposición confirmada o inicio de síntomas o cualquier tipo de prueba de detección de antígenos virales dirigida a ambos antígenos (N/S).
- Los resultados destacan que las pruebas serológicas podrían desempeñar un papel importante en el diagnóstico de sospechas de infecciones por COVID-19 durante la etapa posterior de la enfermedad. En la práctica clínica, las pruebas serológicas de COVID-19 podrían contribuir a la comprensión del estado inmunológico de la población. Sin embargo, no es seguro si una respuesta inmune detectable indica que un paciente es inmune o que ya no es infeccioso. La duración de los aumentos de anticuerpos aún no se conoce completamente y hay pocos datos más allá de los 35 días posteriores al inicio de los síntomas.



8. Bibliografía

1. Coronavirus Update (Live): 95,960,392 Cases and 2,048,061 Deaths from COVID-19 Virus Pandemic - Worldometer [Internet]. [cited 2021 Jan 17]. Available from: https://www.worldometers.info/coronavirus/?utm_campaign=homeAdvegas1?
2. Saeed H, Osama H, Madney YM, Harb HS, Abdelrahman MA, Ehrhardt C, et al. COVID-19; current situation and recommended interventions. *Int J Clin Pract* [Internet]. 2020 Dec 5;(li):0–3. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/jmv.25688>
3. Mohamadian M, Chiti H, Shoghli A, Biglari S, Parsamanesh N, Esmaeilzadeh A. COVID-19; Virology, Biology and Novel Laboratory Diagnosis. *J Gene Med* [Internet]. 2020 Dec 10;0–2. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/jgm.3303>
4. Manigandan S, Wu MT, Ponnusamy VK, Raghavendra VB, Pugazhendhi A, Brindhadevi K. A systematic review on recent trends in transmission, diagnosis, prevention and imaging features of COVID-19. *Process Biochem* [Internet]. 2020;98(August):233–40. Available from: <https://doi.org/10.1016/j.procbio.2020.08.016>
5. Sidiq Z, Hanif M, Dwivedi KK, Chopra KK. Benefits and limitations of serological assays in COVID-19 infection. *Indian J Tuberc* [Internet]. 2020;(xxxx):11–4. Available from: <https://doi.org/10.1016/j.ijtb.2020.07.034>
6. Tantuoayir MM, Rezaei N. Serological Tests for COVID-19: Potential Opportunities. *Cell Biol Int* [Internet]. 2020 Dec 2;cbin.11516. Available from: <http://dx.doi.org/10.1002/cbin.11516>
7. Alpdagtas S, Ilhan E, Uysal E, Sengor M, Ustundag CB, Gunduz O. Evaluation of current diagnostic methods for COVID-19. *APL Bioeng*. 2020;4(4):041506.
8. Díaz D, Peña E, Pinzón C. Manual metodológico para la elaboración de evaluaciones de efectividad, seguridad y validez diagnóstica de tecnologías en salud. Instituto de Evaluación Tecnológica en Salud - IETS. 2018.
9. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* [Internet]. 2009 Dec 4;339(jul21 1):b2700–b2700. Available from: <https://www.bmj.com/lookup/doi/10.1136/bmj.b2700>
10. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69(2016):225–34.
11. Kiefer C, Sturtz S, Bender R. Indirect Comparisons and Network Meta-Analyses. *Dtsch Arztebl Int*. 2015;112(47):803–8.
12. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons; 2019.



13. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies. *JAMA* [Internet]. 2018 Jan 23;319(4):388. Available from: <https://jamanetwork.com/journals/jama/fullarticle/2670259>
14. Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: Grading a Body of Evidence on Diagnostic Tests. *J Gen Intern Med* [Internet]. 2012 Jun;27(S1):47–55. Available from: <http://link.springer.com/10.1007/s11606-012-2021-9>
15. Hartmann KE, Matchar DB, Chang S. Chapter 6: Assessing Applicability of Medical Test Studies in Systematic Reviews. *J Gen Intern Med* [Internet]. 2012 Jun;27(S1):39–46. Available from: <http://link.springer.com/10.1007/s11606-011-1961-9>
16. Matchar DB. Chapter 1: Introduction to the Methods Guide for Medical Test Reviews. *J Gen Intern Med* [Internet]. 2012 Jun;27(S1):4–10. Available from: <http://link.springer.com/10.1007/s11606-011-1798-2>
17. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, et al. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. *PLoS One* [Internet]. 2020;15(12):e0242958. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33301459>
18. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database Syst Rev* [Internet]. 2020 Jun 25;2. Available from: <http://doi.wiley.com/10.1002/14651858.CD013652>
19. Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* [Internet]. 2020 Aug 26;8:CD013705. Available from: <http://doi.wiley.com/10.1002/14651858.CD013705>
20. Jarrom D, Elston L, Washington J, Prettyjohns M, Cann K, Myles S, et al. Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review. *BMJ Evidence-Based Med*. 2020;0(0):bmjebm-2020-111511.
21. Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. *Am J Infect Control*. 2020;49.
22. Guo CC, Mi JQ, Nie H. Seropositivity rate and diagnostic accuracy of serological tests in 2019-nCoV cases: A pooled analysis of individual studies. *Eur Rev Med Pharmacol Sci*. 2020;24(19):10208–18.
23. Kontou PI, Braliou GG, Dimou NL, Nikolopoulos G, Bagos PG. Antibody Tests in Detecting SARS-CoV-2 Infection: A Meta-Analysis. *Diagnostics* [Internet]. 2020 May 19;10(5):319. Available from: <https://www.mdpi.com/2075-4418/10/5/319>
24. Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui LP, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: Systematic review and meta-analysis. *BMJ*. 2020;370.



25. Mekonnen D, Mengist HM, Derby A, Nibret E, Munshea A, He H, et al. Diagnostic accuracy of serological tests and kinetics of severe acute respiratory syndrome coronavirus 2 antibody: A systematic review and meta-analysis. *Rev Med Virol*. 2020;(June).
26. Mustafa Hellou M, Górski A, Mazzaferri F, Cremonini E, Gentilotti E, De Nardo P, et al. Nucleic acid amplification tests on respiratory samples for the diagnosis of coronavirus infections: a systematic review and meta-analysis. *Clin Microbiol Infect* [Internet]. 2020 Nov; Available from: <https://doi.org/10.1016/j.cmi.2020.11.002>
27. Riccò M, Ranzieri S, Peruzzi S, Valente M, Marchesi F, Balzarini F, et al. RT-qPCR assays based on saliva rather than on nasopharyngeal swabs are possible but should be interpreted with caution: Results from a systematic review and metaanalysis. *Acta Biomed*. 2020;91(3):1–15.
28. Wang H, Ai J, Loeffelholz MJ, Tang YW, Zhang W. Meta-analysis of diagnostic performance of serology tests for COVID-19: impact of assay design and post-symptom-onset intervals. *Emerg Microbes Infect* [Internet]. 2020;9(1):2200–11. Available from: <https://doi.org/10.1080/22221751.2020.1826362>
29. Zhang ZL, Hou YL, Li DT, Li FZ. Diagnostic efficacy of anti-SARS-CoV-2 IgG/IgM test for COVID-19: A meta-analysis. *J Med Virol*. 2020;(June):1–9.
30. Instituto nacional de salud INS. Estudio Nacional de Seroprevalencia [Internet]. Available from: <https://www.ins.gov.co/estudio-nacional-de-seroprevalencia/reporte.html#curso>
31. Research and data: Hannah Ritchie, Esteban Ortiz-Ospina, Diana Beltekian, Edouard Mathieu, Joe Hasell, Bobbie Macdonald, Charlie Giattino and MR, Web development: Breck Yunits, Ernst van Woerden, Daniel Gavrilov, Matthieu Bergel, Shahid Ahmad, Jason Crawford and MG. Colombia: Coronavirus Pandemic Country Profile [Internet]. Available from: <https://ourworldindata.org/coronavirus/country/colombia?country=~COL>



9. Anexos

1 Anexo 1. Bitácoras de búsqueda de evidencia en bases de datos electrónicas.

Tipo de búsqueda	Nueva	
Base de datos	EMBASE/MEDLINE (OVID)	
Fecha de búsqueda	17/12/2020	
Rango de fecha de búsqueda	2019-	
Otros límites	Ninguno	
Estrategia de búsqueda (resultados)	#	Búsquedas
	1	SARS-CoV-2.sh.
	2	COVID-19. sh
	3	Sars virus.sh
	4	(sars and virus).af
	5	(sars and cov).af
	6	(sars virus OR sars cov AND "2").af
	7	4 OR 5 OR 6
	8	(sars-cov-2 OR covid19 OR severe acute respiratory syndrome coronavirus 2 OR ncov OR 2019 ncov OR sars cov 2).af
	9	1 OR 2 OR 3 OR
	10	(COVID serological testing OR COVID-19 nucleic acid testing OR COVID-19 testing OR diagnostic tests, routine OR Laboratory proficiency testing OR Self-testing OR Point-of-care testing OR Molecular diagnostic techniques OR Serologic test OR Polymerase chain reaction OR Nucleic acids OR Pathology, molecular).sh
	11	(serolog* AND diagnos*).af
	12	(serolog* AND test*).af
	13	11 OR 12
	14	((molecular OR moleculars) AND test*).af
	15	((molecular OR moleculars) AND diagnos*).af
	16	14 OR 15
	17	((test* or diagnos*) AND nucleic acids).af
	18	10 OR 13 OR 16 OR 17
	19	(diagnos* OR detect*).af
	20	9 AND 19
	21	18 AND 20
	22	Systematic review.af
	23	(meta analysis OR metaanalysis OR meta-analysis).af
	24	22 OR 23
	25	21 AND 24
	26	25 limited to 2019-
		Resultados
		2086
		3107
		3719
		49047
		75120
		76664
		87168
		102608
		115966
		736620
		310610
		324709
		407112
		2127574
		1438566
		2789847
		192649
		3730101
		20272180
		45056
		12757
		1257713
		1253378
		1863757
		1750
		975

Tipo de búsqueda	Nueva	
Base de datos	PubMed (NLM)	



Fecha de búsqueda	17/12/2020		
Rango de fecha de búsqueda	2019-		
Otros límites	Ninguno		
Estrategia de búsqueda (resultados)	#	Búsquedas	Resultados
	1	"SARS-CoV-2"[MeSH Terms]	35497
	2	"COVID-19"[MeSH Terms]	42706
	3	("sars virus"[MeSH Terms] OR ("sars"[All Fields] AND "virus"[All Fields]) OR "sars virus"[All Fields] OR ("sars"[All Fields] AND "cov"[All Fields]) OR "sars cov"[All Fields] AND "2"[All Fields]	28448
	4	"sars-cov-2"[All Fields]	50139
	5	"covid19"[All Fields]	72982
	6	"severe acute respiratory syndrome coronavirus 2"[Supplementary Concept] OR "severe acute respiratory syndrome coronavirus 2"[All Fields] OR "ncov"[All Fields] OR "2019 ncov"[All Fields] OR "covid 19"[All Fields] OR "sars cov 2"[All Fields] OR (("coronavirus"[All Fields] OR "cov"[All Fields]) AND 2019/11/01:3000/12/31[Date - Publication])	84143
	7	1 OR 2 OR 3 OR 4 OR 5 OR 6	85725
	8	"COVID-19 Serological Testing"[MeSH Terms] OR "COVID-19 Nucleic Acid Testing"[MeSH Terms] OR "COVID-19 Testing"[MeSH Terms]	2947
	9	"diagnostic tests, routine"[MeSH Terms]	12697
	10	"Laboratory Proficiency Testing"[MeSH Terms] OR "Self-Testing"[MeSH Terms] OR "Point-of-Care Testing"[MeSH Terms]	9696
	11	"Molecular Diagnostic Techniques"[MeSH Terms]	18183
	12	"Serologic Tests"[MeSH Terms]	178981
	13	"polymerase chain reaction"[MeSH Terms]	451334
	14	((("nucleic acids"[MeSH Terms] OR ("nucleic"[All Fields] AND "acids"[All Fields]) OR "nucleic acids"[All Fields] OR ("nucleic"[All Fields] AND "acid"[All Fields]) OR "nucleic acid"[All Fields]) AND "diagnos*[All Fields]) OR ((("nucleic acids"[MeSH Terms] OR ("nucleic"[All Fields] AND "acids"[All Fields]) OR "nucleic acids"[All Fields] OR ("nucleic"[All Fields] AND "acid"[All Fields]) OR "nucleic acid"[All Fields]) AND "test*[All Fields])	313211
	15	("serolog*[All Fields] AND "diagnos*[All Fields]) OR ("serolog*[All Fields] AND "test*[All Fields])	97201
	16	((("molecular"[All Fields] OR "moleculars"[All Fields]) AND "test*[All Fields]) OR ((("molecular"[All Fields] OR "moleculars"[All Fields]) AND "diagnos*[All Fields])	632112
	17	"molecular diagnostic techniques"[MeSH Terms] OR ("molecular"[All Fields] AND "diagnostic"[All Fields] AND "techniques"[All Fields]) OR "molecular diagnostic techniques"[All Fields] OR ("molecular"[All Fields] AND "testing"[All Fields]) OR "molecular testing"[All Fields]	107243
	18	("pathology, molecular"[MeSH Terms] OR ("pathology"[All Fields] AND "molecular"[All Fields]) OR "molecular pathology"[All Fields] OR ("molecular"[All Fields] AND "diagnostic"[All Fields]) OR "molecular diagnostic"[All Fields]) AND ("research design"[MeSH Terms] OR ("research"[All	50226



		Fields] AND "design"[All Fields]) OR "research design"[All Fields] OR "test"[All Fields])	
19		("serologic"[All Fields] OR "serological"[All Fields] OR "serologically"[All Fields]) AND ("diagnosable"[All Fields] OR "diagnosi"[All Fields] OR "diagnosis"[MeSH Terms] OR "diagnosis"[All Fields] OR "diagnose"[All Fields] OR "diagnosed"[All Fields] OR "diagnoses"[All Fields] OR "diagnosing"[All Fields] OR "diagnosis"[MeSH Subheading])	76036
20		8 OR 9 OR 10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19	1746742
21		"systematic review"[Publication Type] OR "systematic reviews as topic"[MeSH Terms] OR "systematic review"[All Fields] OR ("review"[Publication Type] OR "review literature as topic"[MeSH Terms] OR "review"[All Fields]) AND "classification"[MeSH Terms] OR ("meta analysis"[Publication Type] OR "meta analysis as topic"[MeSH Terms] OR "meta analysis"[All Fields]) OR ("meta analysis as topic"[MeSH Terms] OR ("meta analysis"[All Fields] AND "topic"[All Fields]) OR "meta analysis as topic"[All Fields] OR "metaanalysis"[All Fields])	321017
22		7 AND 20 AND 21	242

Tipo de búsqueda	Nueva		
Base de datos	COCHRANE/Cochrane reviews		
Fecha de búsqueda	17/12/2020		
Rango de fecha de búsqueda	2019-		
Otros límites	Revisiones sistemáticas de la literatura		
Estrategia de búsqueda (resultados)	#	Búsquedas	Resultados
	1	(SARS-CoV-2):ti,ab,kw	162
	2	(COVID-19):ti,ab,kw	3393
	3	(covid19):ti,ab,kw	228
	4	(severe acute respiratory syndrome coronavirus 2):ti,ab,kw	508
	5	(2019 ncov):ti,ab,kw	123
	6	(covid 19):ti,ab,kw	3396
	7	(sars cov 2):ti,ab,kw	262
	8	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7	3513
	9	MeSH descriptor: [Molecular Diagnostic Techniques] explode all trees	56
	10	MeSH descriptor: [Serologic Tests] explode all trees	1569
	11	MeSH descriptor: [Polymerase Chain Reaction] explode all trees	2065
	12	#9 OR #10 OR #11	3652
	13	(serolog*):ti,ab,kw	4209
	14	(molecular*):ti,ab,kw	17085
	15	(nucleic acid*):ti,ab,kw	895
	16	#13 OR #14 OR #15	21898
	17	(test*):ti,ab,kw	377741
	18	(diagnos*):ti,ab,kw	229262
19	(technique*):ti,ab,kw	93374	



20	(detect*):ti,ab,kw	84238
21	#17 OR #18 OR #19 OR #20	623864
22	#16 AND #21	11758
23	MeSH descriptor: [Diagnosis] explode all trees	332785
24	#22 OR #23	341902
25	#8 AND #24 in Cochrane Reviews	6

Tipo de búsqueda	Nueva	
Base de datos	Epistemonikos	
Fecha de búsqueda	17/12/2020	
Rango de fecha de búsqueda	2019-	
Otros límites	Revisiones sistemáticas de la literatura	
Estrategia de búsqueda (resultados)	#	Búsquedas
	1	(title:(SARS-CoV-2) OR abstract:(SARS-CoV-2)) OR (title:(COVID-19) OR abstract:(COVID-19)) OR (title:(covid19) OR abstract:(covid19)) OR (title:(severe acute respiratory syndrome coronavirus 2) OR abstract:(severe acute respiratory syndrome coronavirus 2)) OR (title:(covid 19) OR abstract:(covid 19)) OR (title:(sars cov 2) OR abstract:(sars cov 2))
	2	(title:(Serologic Tests) OR abstract:(Serologic Tests)) OR (title:(Molecular Diagnostic Techniques) OR abstract:(Molecular Diagnostic Techniques)) OR (title:(Nucleic Acid Testing) OR abstract:(Nucleic Acid Testing)) OR (title:(Point-of-Care Testing) OR abstract:(Point-of-Care Testing)) OR (title:(Self-Testing) OR abstract:(Self-Testing)) OR (title:(polymerase chain reaction) OR abstract:(polymerase chain reaction)) OR (title:(serological testing) OR abstract:(serological testing))
	3	1 AND 2
		Resultados
		2669
		954
		86

Tipo de búsqueda	Cualquier tipo de documento	
Base de datos	Google scholar	
Fecha de búsqueda	18/12/2020	
Rango de fecha de búsqueda	2019-	
Otros límites	Ninguno	
Estrategia de búsqueda (resultados)	Con todas las palabras: ("COVID19" OR "SARS Cov 2") AND ("diagnostics*" OR "detect*" OR "test*" OR "serology*" OR "molecular" OR "nucleic acid") AND ("systematic review" OR "meta analysis")	
	Con todas las palabras: ("COVID 19" OR "SARS Cov 2") AND ("diagnóstico*" OR "detectiON" "prueba" OR "serológi*" OR "molecular" OR "acido nucleico") AND ("revisión sistemática" OR "meta análisis")	
		26
		1

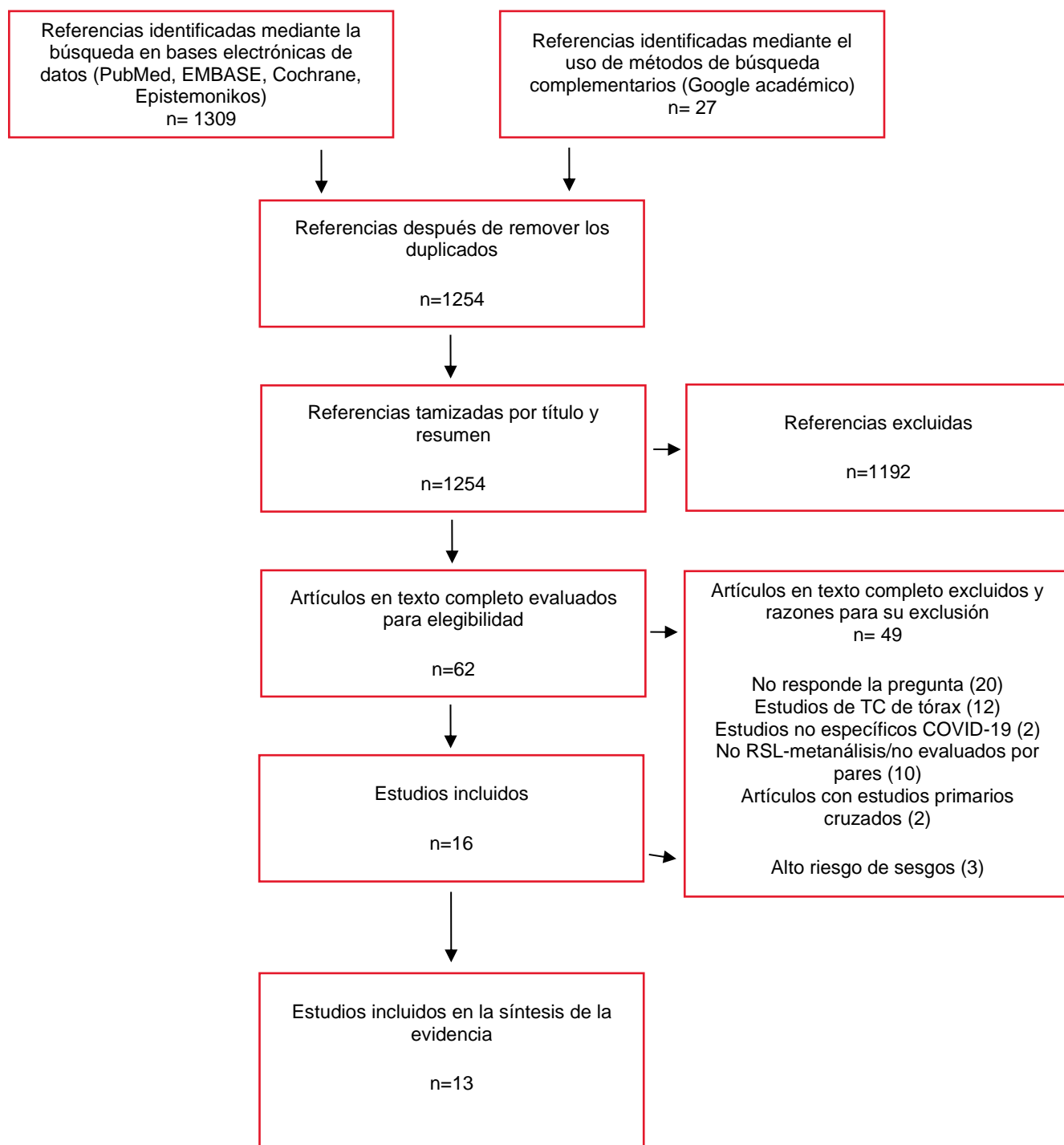


La salud
es de todos

Minsalud



Anexo 2. Diagrama PRISMA





Anexo 3. Estudios incluidos

1. Lista de artículos incluidos

1. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, et al. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. PLoS One [Internet]. 2020;15(12):e0242958. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33301459>
2. Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. Am J Infect Control. 2020;49.
3. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. Cochrane Database Syst Rev [Internet]. 2020 Jun 25;2. Available from: <http://doi.wiley.com/10.1002/14651858.CD013652>
4. Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. Cochrane Database Syst Rev [Internet]. 2020 Aug 26;8:CD013705. Available from: <http://doi.wiley.com/10.1002/14651858.CD013705>
5. Guo CC, Mi JQ, Nie H. Seropositivity rate and diagnostic accuracy of serological tests in 2019-nCoV cases: A pooled analysis of individual studies. Eur Rev Med Pharmacol Sci. 2020;24(19):10208–18.
6. Jarrom D, Elston L, Washington J, Prettyjohns M, Cann K, Myles S, et al. Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review. BMJ Evidence-Based Med. 2020;0(0):bmjebm-2020-111511.
7. La Marca A, Capuzzo M, Paglia T, Roli L, Trenti T, Nelson SM. Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide to molecular and serological in-vitro diagnostic assays. Reprod Biomed Online [Internet]. 2020;41(3):483–99. Available from: <https://doi.org/10.1016/j.rbmo.2020.06.001>
8. Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui LP, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: Systematic review and meta-analysis. BMJ. 2020;370.
9. Mekonnen D, Mengist HM, Derby A, Nibret E, Munshea A, He H, et al. Diagnostic accuracy of serological tests and kinetics of severe acute respiratory syndrome coronavirus 2 antibody: A systematic review and meta-analysis. Rev Med Virol. 2020;(June).
10. Mustafa Hellou M, Górská A, Mazzaferri F, Cremonini E, Gentilotti E, De Nardo P, et al. Nucleic acid amplification tests on respiratory samples for the diagnosis of coronavirus infections: a systematic review and meta-analysis. Clin Microbiol Infect [Internet]. 2020 Nov;(January). Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1198743X20306959>



11. Riccò M, Ranzieri S, Peruzzi S, Valente M, Marchesi F, Balzarini F, et al. RT-qPCR assays based on saliva rather than on nasopharyngeal swabs are possible but should be interpreted with caution: Results from a systematic review and metaanalysis. *Acta Biomed.* 2020;91(3):1–15.
12. Wang H, Ai J, Loeffelholz MJ, Tang YW, Zhang W. Meta-analysis of diagnostic performance of serology tests for COVID-19: impact of assay design and post-symptom-onset intervals. *Emerg Microbes Infect.* 2020;9(1):2200–11.
13. Zhang ZL, Hou YL, Li DT, Li FZ. Diagnostic efficacy of anti-SARS-CoV-2 IgG/IgM test for COVID-19: A meta-analysis. *J Med Virol.* 2020;(June):1–9.



2. Tabla 4. Resumen de criterios de calidad de las RSL y RSL-metanálisis incluidas

Estudio/ Fecha publicación (dd/mm/aa)	Objetivo de la RSL	Tipo de estudios y tipo de participantes incluidos	Bases de datos buscadas	Evaluación de sesgos y calidad	Evalación de heterogeneidad	Metanálisis	*Kappa/ acuerdo en evaluación metodológica
Arevalo- Rodríguez et al. (17) 10/12/2020	Obtener a través de una RSL, una estimación resumida de la proporción de falsos negativos relacionados con la detección de SARS-CoV-2 mediante ensayos de RT-PCR en el primer encuentro de atención médica (prueba inicial).	No se especifican las clases de estudios observacionales incluidos. Pacientes con sospecha de infección por criterios clínicos o epidemiológicos. El rango de tiempo desde el inicio de síntomas (0-14 días) fue reportado en 8 estudios. Dos estudios reportaron hasta el día 24. En todos los casos, la presencia de infección se confirmó con cualquier ensayo de RT-PCR en tiempo real que se repitió después de un resultado negativo.	MEDLINE; Embase; LILACS. Recursos adicionales: Base de datos de la OMS; El repositorio centralizado LOVE (Living Overview of Evidence) desarrollado por Epistemonikos; El Proyecto de Acceso Abierto COVID-19 Evidencia Viva sobre COVID-19, desarrollado en el Instituto de Medicina Social y Preventiva de la Universidad de Berna	Herramienta QUADAS-2. El dominio más afectado por un alto riesgo de sesgo fue el dominio de flujo y tiempo; algunos estudios no proporcionaron información sobre el intervalo de tiempo para la administración de un nuevo ensayo de RT-PCR	Resultados de proporciones estimadas (con IC del 95%) en un diagrama de bosque para evaluar la variabilidad entre estudios. Los datos presentaron heterogeneidad considerable entre los estudios; el intervalo de predicción del 90% osciló entre 0,02 y 0,54 (tau cuadrado = 1,39).	No se realizó Subgrupos de interés: Gen objeto, tipo de muestra biológica, y país de origen de los estudios primarios.	0,558 Moderada
Boger et al. (21) 01/01/2021	Realizar una RSL con metanálisis para recopilar evidencia sobre las características de todas las pruebas de diagnóstico disponibles para el SARS-CoV-2, incluidos los parámetros de sensibilidad, especificidad, cocientes de probabilidad positiva y negativa y curvas ROC	Un estudio de casos y controles (Evaluación de RT-PCR) y once estudios observacionales de series retrospectivas. Todos los estudios presentaron un grupo de evaluación (pacientes diagnosticados de COVID-19), mientras que solo 5 estudios utilizaron un grupo control (pacientes negativos para COVID-19). Los pacientes del grupo de prueba fueron diagnosticados previamente mediante la técnica de PCR (estándar de oro).	Pubmed; Scopus. Recursos adicionales: Google Scholar	Herramienta QUADAS-2. Los estudios se calificaron como de calidad metodológica general moderada	Se realizaron análisis de sensibilidad para todos los metanálisis con resultados de alta heterogeneidad ($I^2 > 50\%$); sin embargo, no se encontraron diferencias adicionales en comparación con los efectos de los análisis originales.	Se realizaron metanálisis evaluando los parámetros de exactitud y metanálisis según el tipo de muestra. Subgrupos de interés: tipo de muestra biológica, y tipo de anticuerpo para las pruebas serológicas.	0,588 Moderado



Deeks et al. 2020 (18) 25/06/2020	Evaluar la exactitud diagnóstica de las pruebas de anticuerpos para determinar si una persona que se presenta en la comunidad o en atención primaria o secundaria tiene infección por SARS-CoV-2, o ha tenido previamente una infección por SARSCoV-2, y la exactitud de las pruebas de anticuerpos para su uso en estudios de seroprevalencia.	<p>Se identificaron seis estudios que reclutaron casos sospechosos de COVID-19 antes de que se determinara si los pacientes tenían o no COVID-19.</p> <p>Los otros cuarenta y ocho estudios reclutaron de forma retrospectiva pacientes cuando ya se sabía si tenían o no COVID-19.</p> <p>Veintinueve (29) estudios utilizaron diseños de dos o varios grupos con una selección separada de casos de COVID-19 y participantes sanos o participantes sin COVID-19 con otra enfermedad.</p> <p>Diecinueve (19) estudios incluyeron solo un grupo de solo casos de COVID-19, por lo que solo permitieron estimar la sensibilidad.</p> <p>Ventinueve (21) estudios reportaron recolección de las muestras antes del día 21 del inicio de los síntomas; dos (2) después del día 21; veintitres (23) reportaron tiempos mixtos, y seis (6) no reportaron esta información.</p>	El Registro de estudios Cochrane COVID-19 (covid-19.cochrane.org/) y la base de datos de evidencia viva COVID-19 de la Universidad de Berna. Registros Embase obtenidos a través de Martha Knuth para los Centros para el Control y la Prevención de Enfermedades (CDC); Base de datos descargable de artículos de investigación COVID-19. Repositorios: El Centro de Coordinación e Información de Evidencia para Políticas y Prácticas (Centro EPPi) y el Instituto Noruego de Salud Pública.	Herramienta QUADAS-2. El riesgo de sesgo era alto en 89% estudios sobre cómo se seleccionaron los participantes, en 26% relacionados con la aplicación de la prueba índice, 31% por preocupaciones sobre el estándar de referencia y 54% por cuestiones relacionadas con el flujo de participantes y el tiempo. Ningún estudio tuvo bajo riesgo en todos los dominios.	Las estimaciones de sensibilidad y especificidad de los estudios para IgG, IgM e IgG/IgM, mostraron una heterogeneidad sustancial con estimaciones de sensibilidad desde 0% al 100% para los tres marcadores. Para IgA, anticuerpos totales, IgA/IgG, IgA/IgM, mostraron una heterogeneidad similar con un menor número de estudios incluidos en el análisis.	No se realizó	Subgrupos de interés: tiempo de inicio de síntomas, y tipo de anticuerpo para las pruebas serológicas.	0,462 Moderado
Dinnes et al. 2020 (19) 26/08/2020	Evaluar la exactitud diagnóstica de las pruebas rápidas de antígeno y de base molecular en el lugar de atención para determinar si una persona que se presenta en la comunidad o en atención primaria o secundaria tiene una infección por SARS-CoV-2.	<p>Los estudios seleccionaron predominantemente muestras de las enviadas a los laboratorios para pruebas de RT-PCR de rutina con detalles limitados de los participantes que proporcionaron las muestras.</p> <p>Siete fueron estudios de dos grupos y nueve de un solo grupo. Dos estudios incluyeron solo muestras con</p>	El Registro de estudios Cochrane COVID-19 (covid-19.cochrane.org/) y la base de datos de evidencia viva COVID-19 de la Universidad de Berna. Registros Embase obtenidos a través de Martha Knuth para los Centros para el Control y la Prevención de Enfermedades (CDC); Base de datos descargable de artículos de investigación COVID-19. Repositorios: El	Herramienta QUADAS-2. Se consideró que el riesgo de sesgo era alto en 50% estudios por cómo seleccionaron las muestras y en 72% porque consideraron que una RT-PCR negativa era suficiente para confirmar la ausencia de infección por COVID-19.	Se investigó la heterogeneidad relacionada con la carga viral, la marca de la prueba y el tipo de muestra. Hubo una heterogeneidad considerable en la sensibilidad de las pruebas de antígeno, con resultados entre los estudios que variaron del 0% al 94%. Para las pruebas moleculares, la	No se realizó	Subgrupos de interés: tipos de pruebas de antígenos virales y tipos de pruebas de anticuerpos, tipo de muestra biológica, carga viral.	0,504 Moderado



		SARS-CoV-2 confirmado, por lo que solo permitieron estimar la sensibilidad. Todos los estudios definieron la presencia o ausencia de infección por COVID-19 según la RT-PCR, con un único (n = 17) o dos (n = 1) resultados de RT-PCR negativos utilizados para confirmar la ausencia de infección. Solo un (1) estudio reportó tiempo desde el inicio de los síntomas en 2 días.	Centro de Coordinación e Información de Evidencia para Políticas y Prácticas (Centro EPPI) y el Instituto Noruego de Salud Pública.		heterogeneidad en la sensibilidad varió del 68% al 100%		
Guo et al. 2020 (22) 01/10/2020	Evaluar sistemáticamente la tasa positiva de detección de anticuerpos (IgM, IgG) en 2019-nCoV, así como su eficiencia diagnóstica integral mediante un metanálisis cuantitativo con el fin de proporcionar pruebas novedosas para el diagnóstico clínico y el tratamiento de 2019-nCoV.	La población de estudio estuvo compuesta por poblaciones asiáticas y europeas. Se utilizó RT-PCR como referencia para el diagnóstico de 2019-nCoV. Se incluyeron 22 estudios para la síntesis cuantitativa de la tasa de seropositividad y 16 para el desempeño diagnóstico. Diez (10) estudios reportaron el tiempo de recolección de las muestras desde el inicio de los síntomas en un rango de 0-39 días; dos (2) de estos estudios reportaron más de 30 días.	PubMed, medRxiv, bioRxiv, EMBASE	Herramienta QUADAS-2. Las puntuaciones de los 8 estudios fueron superiores a 4 puntos, lo que indica que la calidad de los estudios incluidos fue aceptable. El gráfico en embudo de mostró que no hubo sesgo de publicación claro en los estudios de diagnóstico agrupados en general (prueba de IgM: p = 0,096; prueba de IgG: p = 0,231; prueba de IgM / IgG: p = 0,154).	El análisis del coeficiente de correlación de Spearman no demostró heterogeneidad causada por el efecto umbral entre todos los análisis.	Las medidas del efecto combinado incluyeron la tasa positiva, sensibilidad, especificidad, razón de probabilidad positiva y negativa, razón de probabilidades de diagnóstico (DOR). Subgrupos de interés: tipos de pruebas de anticuerpos, tipo de anticuerpo, tiempo desde el inicio de los síntomas.	1,000 Excelente
Jarrom et al. 2020 (20) 01/10/2020	Identificar, evaluar y resumir la evidencia publicada sobre el rendimiento diagnóstico y la efectividad de las pruebas de virus y anticuerpos del SARS-CoV-2 en el diagnóstico y manejo del COVID-19 actual o anterior.	Para la evaluación de pruebas moleculares, se incluyó una RSL y metanálisis de 68 estudios primarios, tres estudios de validación de pruebas, 6 estudios de series retrospectivas, y dos series prospectivas. Para las evaluaciones de pruebas serológicas, se	Medline, Embase, Cochrane Library, base de datos de la Red Internacional de Agencias para la Evaluación de Tecnologías Sanitarias (INAHTA) y Open Gray. Recursos adicionales: The	Herramienta QUADAS-2. Para las pruebas de virus, la mayoría de los estudios tenían un riesgo alto o poco claro de sesgo con respecto a la selección de pacientes. Para las pruebas de anticuerpos, el método de selección de pacientes se consideró	No se evaluó la heterogeneidad estadística. Se menciona alta heterogeneidad de los estudios solamente.	El metanálisis de sensibilidad y / o especificidad se realizó solo para estudios adecuados utilizando un modelo binomial bivariado de efectos aleatorios.	0,288 Débil



		incluyeron un estudio de validación de pruebas, 6 estudios de series retrospectivas, cinco series prospectivas, y un estudio de cohorte.	Health Technology Wales COVID-19 Evidence Digest	poco claro en el 36% de los estudios y elevado en otro 52%.		Subgrupos de interes: tipos de pruebas de anticuerpos, tipo de anticuerpo, gen objeto de las pruebas moleculares.	
Kontou et al. 2020 (23) 19/05/2020	Resumir la evidencia disponible sobre el desempeño de todas las pruebas de anticuerpos disponibles para el SARS-CoV-2.	<p>Un total de 21 estudios informaron datos para casos de COVID-19 y para controles, mientras que 17 estudios informaron datos solo para los casos.</p> <p>Treinta (30) estudios reportaron tiempo de aparición de los síntomas durante la recolección de las muestras para un rango de 2-43 días.</p>	PubMed, medRxiv y bioRxiv.	Herramienta QUADAS-2. Las pruebas de Egger y Begg no detectaron sesgo de publicación u otros efectos de estudios pequeños	Se evaluó la heterogeneidad estadística mostrando ser considerable.	<p>Se usó el método metaanalítico bivariado modificado para el análisis de pruebas diagnósticas. El metanálisis se realizó en los casos en los que había dos o más estudios disponibles. La metarregresión en los casos en que se dispuso de 5 o más estudios.</p> <p>Subgrupos de interes: tipos de pruebas de anticuerpos, tipo de anticuerpo, tipo de antígeno.</p>	1,000 Excelente
Lisboa et al. 2020 (24) 23/06/2020	Evaluar la calidad de la evidencia disponible, comparar sensibilidades y especificidades agrupadas de diferentes pruebas e identificar estudios, pruebas y características de las pacientes asociadas con la exactitud de la prueba.	<p>Entre los estudios incluidos, el 50% no fueron revisados por pares.</p> <p>El 80% de los estudios utilizaron un diseño de casos y controles y el 10% incluyó poblaciones de pacientes ambulatorios.</p> <p>La gravedad de la enfermedad se informó en el 40% y la sensibilidad estratificada por el tiempo desde el inicio de los</p>	Ovid-Medline, medRxiv y bioRxiv	Herramienta QUADAS-2. Para el dominio de selección de pacientes, un riesgo alto o poco claro de sesgo en el 98%, relacionado con un diseño de casos y controles y sin un muestreo consecutivo o aleatorio. Para el dominio de la prueba índice, el 73% mostró un riesgo alto o poco claro de sesgo porque no estaba claro si la prueba serológica se	Como los modelos eran bivariados, no se usaron el estadístico I ² . Las curvas ROC con regiones de predicción del 95% y diagramas de bosque mostraron importante heterogeneidad de los estudios.	<p>Metanálisis bivariado con un efecto aleatorio a nivel de prueba únicamente.</p> <p>Subgrupos de interes: tipos de pruebas de anticuerpos, tipo de anticuerpo, tipo de antígeno, y origen de la muestra (tipo de paciente)</p>	1,000 Excelente



		síntomas se informó en el 45%.		interpretó ciega al estándar de referencia			
Mekonnen et al. 2020 (25) 26/09/2020	Determinar la exactitud de las pruebas de diagnóstico serológicas y evaluar la cinética de la producción de anticuerpos a lo largo del tiempo en pacientes con COVID-19.	Los artículos evaluaron los kits de pruebas serológicas desarrollados por 47 empresas (31 comerciales conocidas, 9 comerciales anónimas y 7 internas). Quince (62,5%) de los artículos incluidos informaron los tipos de muestras de RT - qPCR utilizadas. Trece (54%) artículos utilizaron muestras de sangre de control recolectadas en el período anterior a COVID-19.	MEDLINE a través de PubMed, Scopus, medRxiv y bioRxiv	Herramienta QUADAS-2. La mayoría de los artículos no proporcionaron información clara sobre el cegamiento al leer el índice y las pruebas estándar de oro. Además, no se proporcionó información sobre la estrategia de muestreo (aleatoria, consecutiva o de casos y controles).	No se evaluó la heterogeneidad estadística; se evaluó visualmente mediante diagramas de bosque: la exactitud de las pruebas de diagnóstico fue heterogénea	Metanálisis bivariado de la exactitud de las pruebas de diagnóstico mediante un enfoque de modelo mixto lineal generalizado. Debido a la naturaleza inherente de la heterogeneidad en la precisión; Se utilizó un enfoque de efectos aleatorios. Subgrupos de interés: tipos de pruebas de anticuerpos, tipo de anticuerpo, tipo de antígeno, y tiempo desde el inicio de los síntomas.	0,696 Bueno
Mustafa et al. 2020 (26) 04/11/2020	Resumir los estudios que evalúan la exactitud de las pruebas de diagnóstico de las NAT realizadas en muestras respiratorias para el diagnóstico de infecciones respiratorias agudas superiores o inferiores causadas por coronavirus (un aparatado para COVID-19)	Los estudios evaluaron principalmente a pacientes con una enfermedad inespecífica similar a la influenza o sospecha de COVID-19. Doce estudios fueron prospectivos (12/29 estudios COVID-19) y el resto fueron retrospectivos o mixtos, por lo general utilizando muestras almacenadas para el análisis.	PubMed, Web of Science, Cochrane Library, Embase y Open Grey. Se realizó una actualización de PubMed y medRxiv para incluir estudios que examinan las NAAT para COVID-19 hasta el 31 de agosto de 2020.	Herramienta QUADAS-2. Los estudios tuvieron un mayor riesgo de sesgo que el riesgo de mala aplicabilidad. Los procedimientos de selección de pacientes tuvieron en su mayoría un riesgo de sesgo alto o poco claro. Las pruebas índices tenían un alto riesgo de sesgo; no estaba claro si las pruebas índice se interpretaron sin conocimiento de los resultados del estándar de referencia	Las posibles fuentes de heterogeneidad se incluyeron como covariables en el modelo de metarregresión. Al analizar las tasas de concordancia entre diferentes NAT para diagnosticar COVID-19, la heterogeneidad podría explicarse por varios factores relacionados con la muestra tomada y el tipo y métodos de la prueba de PCR.	Se realizó mediante el modelo bivariado, un método de metarregresión jerárquica que incorpora sensibilidad y especificidad. Subgrupos de interés: tipos de pruebas moleculares, tipo de muestra biológica, número y tipo de genes objeto.	Acuerdo total



Riccó et al. 2020 (27) 07/09/2020	Explorar la confiabilidad de las pruebas de diagnóstico RT-PCR para SARS-CoV-2 basadas en muestras de saliva	Seis de los estudios incluidos eran preimpresos, mientras que ocho estudios habían recibido un proceso completo de revisión por pares. En general, 9 estudios tuvieron un diseño de casos y controles, ya que incluyeron casos positivos y negativos de COVID-19, mientras que 5 estudios incluyeron solo pacientes que habían recibido un diagnóstico previo de COVID-19	PubMed / MEDLINE y EMBASE, así como servidores de preimpresión, incluido medrxiv.org	El sesgo de publicación se evaluó mediante gráficos en embudo y análisis de regresión. El análisis de regresión confirmó un alto sesgo de publicación significativo	Evaluación visual de curvas ROC y cálculo del estadístico I^2 . La heterogeneidad fue sustancial tanto para las estimaciones de especificidad ($I^2 = 74\%$, $p < 0.01$) como para la sensibilidad (I^2 , 79%)	Los datos de sensibilidad y especificidad se calcularon para cada estudio y luego se agruparon en un modelo de efectos aleatorios. Subgrupos de interés: tipo de estudio y sincronización en la recolección de la muestra.	1,000 Excelente
Wang et al. 2020 (28) 16/09/2020	Evaluar la eficacia diagnóstica y las características de las pruebas serológicas actuales para COVID-19.	16 estudios de casos y controles; 7 estudios longitudinales; y 4 estudios de cohortes. El 37,0% de los estudios se realizaron en China.	PubMed, biblioteca Cochrane, EBSCO y OVID	Herramienta QUADAS-2. El 85,1% de los estudios mostró un alto riesgo de sesgo en la selección de pacientes; estos artículos no evitaron el diseño de casos y controles o longitudinal.	Se proporcionaron diagramas de bosque de estimaciones puntuales e IC95%. Se trazaron curvas ROC para evaluar la heterogeneidad (efecto umbral) entre los estudios. No hubo heterogeneidad entre los estudios	La análisis de datos se realizó utilizando la aproximación de Laplace anidada integrada bivariada bayesiana. Subgrupos de interés: tipos de pruebas de anticuerpos, tipo de anticuerpo, tipo de antígeno, y tiempo desde el inicio de los síntomas.	1,000 Excelente
Zhang et al. 2020 (29) 18/06/2020	Resumir la eficacia diagnóstica de la prueba IgG / IgM anti-SARSCoV-2 en cada estudio, cuyos resultados pueden ayudar en el diagnóstico de SARSCoV2	Solo un estudio prospectivo, el resto (21) fueron retrospectivos. Trece (13) estudios reportaron tiempo de aparición de los síntomas durante la recolección de las muestras para un rango de 20-62 días.	Bases de datos PubMed, Web of Science, Embase, CNKI (China) y Wanfang (China)	Herramienta QUADAS-2. Los gráficos en embudo de Deeks se dibujaron para evaluar el riesgo de sesgo de publicación. Se mostró un alto riesgo de sesgo en el 95,45% de los estudios para la selección de pacientes, en el 36,36% para la prueba índice, en el 13,64% para el flujo y el tiempo y en el 4,55% para el estándar de referencia.	Se evaluó usando el estadístico I^2 y el test Q. Se mostró un alto riesgo de sesgo en el 95,45% de los estudios para la selección de pacientes, en el 36,36% para la prueba índice, en el 13,64% para el flujo y el tiempo y en el 4,55% para el estándar de referencia.	Se utilizó un modelo bivariado de efectos aleatorios para estimar las medidas de rendimiento diagnóstico agrupadas y un IC del 95%. Subgrupos de interés: tipos de pruebas de	1,000 Excelente



				tiempo y en el 4,55% para el estándar de referencia.		anticuerpos, tipo de anticuerpo	
--	--	--	--	--	--	---------------------------------	--

CLIA: inmunoensayo quimioluminiscente; CGIA: inmunoensayo de oro coloidal; Elisa: Ensayos inmunoabsorbentes ligados a enzimas; IFL: inmunoensayo de flujo lateral; IFI: inmunofluorescencia indirecta; LIPS: sistema de inmunoprecipitación de luciferasa; RT-PCR: Reacción en Cadena de la Polimerasa con Transcriptasa Inversa; NAT: pruebas basadas en ácidos nucleicos; RVP: razon de verosimilitud positiva; RVN: razon de verosimilitud negativa; VPP: valores predictivos positivos; VPN: valores predictivos negativos; TR: tracto respiratorio

*Acuerdos de las evaluaciones de calidad metodologica con las herramientas IQWiG y ROBIS de acuerdo con el diseño del estudio realizadas por los dos revisores (AMHT y KCP)



Anexo 4. Lista de artículos excluidos y las causas de exclusión

Estudios que no responden la pregunta de investigación

1. Alnor A, Sandberg MB, Gils C, Vinholt PJ. Laboratory Tests and Outcome for Patients with Coronavirus Disease 2019: A Systematic Review and Meta-Analysis. J Appl Lab Med [Internet]. 2020 Sep 1;5(5):1038–49. Available from: <https://academic.oup.com/jalm/article/5/5/1038/5861535>
2. Bwire GM, Majigo M V., Njiro BJ, Mawazo A. Detection profile of SARS-CoV-2 using RT-PCR in different types of clinical specimens: A systematic review and meta-analysis. J Med Virol. 2020;
3. Elshazli RM, Toraih EA, Elgaml A, El-Mowafy M, El-Mesery M, Amin MN, et al. Diagnostic and prognostic value of hematological and immunological markers in COVID-19 infection: A meta-analysis of 6320 patients. PLoS One [Internet]. 2020;15(8 August):1–20. Available from: <http://dx.doi.org/10.1371/journal.pone.0238160>
4. Fakheran O, Dehghannejad M, Khademi A. Saliva as a diagnostic specimen for detection of SARS-CoV-2 in suspected patients: A scoping review. Infect Dis Poverty. 2020;9(1):1–7.
5. Hanson KE, Caliendo AM, Arias CA, Englund JA, Hayden MK, Lee MJ, et al. Infectious Diseases Society of America Guidelines on the Diagnosis of Coronavirus Disease 2019 (COVID-19): Serologic Testing. Clin Infect Dis [Internet]. 2020 Sep 12;1–44. Available from: <https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1343/5904785>
6. Jefferson T, Ea S, Brassey J, Oxford HCU De. CULTIVOS VIRALES PARA LA EVALUACIÓN DE LA INFECTIVIDAD DE LA COVID-19: una revisión sistemática. 2020;1–8.
7. Jefferson T, Spencer EA, Brassey J, Heneghan C. Viral cultures for COVID-19 infectious potential assessment – a systematic review. Clin Infect Dis [Internet]. 2020 Dec 3; Available from: <https://academic.oup.com/cid/advance-article/doi/10.1093/cid/ciaa1764/6018217>
8. Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A. The role of biomarkers in diagnosis of COVID-19 – A systematic review. Life Sci [Internet]. 2020;254(May):117788. Available from: <https://doi.org/10.1016/j.lfs.2020.117788>
9. La Marca A, Capuzzo M, Paglia T, Roli L, Trenti T, Nelson SM. Testing for SARS-CoV-2 (COVID-19): a systematic review and clinical guide to molecular and serological in-vitro diagnostic assays. Reprod Biomed Online [Internet]. 2020 Sep;41(3):483–99. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1472648320303187>



10. Lawandi A, Danner RL. Antibody tests have higher sensitivity at ≥ 15 days after symptom onset and 99% specificity for detecting SARS-CoV-2. *Ann Intern Med*. 2020;173(10):JC57.
11. Li WT, Ma J, Shende N, Castaneda G, Chakladar J, Tsai JC, et al. Using machine learning of clinical data to diagnose COVID-19: A systematic review and meta-analysis. *BMC Med Inform Decis Mak*. 2020;20(1):1–13.
12. Mallett S, Allen AJ, Graziadio S, Taylor SA, Sakai NS, Green K, et al. At what times during infection is SARS-CoV-2 detectable and no longer detectable using RT-PCR-based tests? A systematic review of individual participant data. *BMC Med*. 2020;18(1):1–17.
13. Mattiuzzi C, Henry BM, Sanchis-Gomar F, Lippi G. Sars-cov-2 recurrent rna positivity after recovering from coronavirus disease 2019 (COVID-19): A meta-analysis. *Acta Biomed*. 2020;91(3):1–7.
14. Mohammadi A, Esmaeilzadeh E, Li Y, Bosch RJ, Li JZ. SARS-CoV-2 detection in different respiratory sites: A systematic review and meta-analysis. *EBioMedicine*. 2020; 59:1–6.
15. Silva EH, Lekamwasam S. Confirmatory and supportive laboratory investigations in SARS-CoV-2 infection; a systematic review. *J Prev Epidemiol [Internet]*. 2020 Jul 6;5(1): e03–e03. Available from: <https://doi.org/10.34172/jpe.2020.03>
16. Stegeman I, Ochodo EA, Guleid F, Holtman GA, Yang B, Davenport C, et al. Routine laboratory testing to determine if a patient has COVID-19. *Cochrane Database Syst Rev [Internet]*. 2020 Nov 19;19. Available from: <http://doi.wiley.com/10.1002/14651858.CD013787>
17. Uygun-Can B, Acar-Bolat B. Clinical Properties and Diagnostic Methods of COVID-19 Infection in Pregnancies: Meta-Analysis. *Biomed Res Int*. 2020;2020.
18. van Doorn AS, Meijer B, Frampton CMA, Barclay ML, de Boer NKH. Systematic review with meta-analysis: SARS-CoV-2 stool testing and the potential for faecal-oral transmission. *Aliment Pharmacol Ther*. 2020;52(8):1276–88.
19. Wang J gan, Cui H rong, Tang H bo, Deng X li. Gastrointestinal symptoms and fecal nucleic acid testing of children with 2019 coronavirus disease: a systematic review and meta-analysis. *Sci Rep [Internet]*. 2020;10(1):1–9. Available from: <https://doi.org/10.1038/s41598-020-74913-0>
20. Wong MC, Huang J, Lai C, Ng R, Chan FKL, Chan PKS. Detection of SARS-CoV-2 RNA in fecal specimens of patients with confirmed COVID-19: A meta-analysis. *J Infect [Internet]*. 2020 Aug;81(2):e31–8. Available from: <https://linkinghub.elsevier.com/retrieve/pii/S0163445320303947>.

Estudios acerca de la TC de tórax como prueba diagnóstica



21. Adams HJA, Kwee TC, Yakar D, Hope MD, Kwee RM. Systematic Review and Meta-Analysis on the Value of Chest CT in the Diagnosis of Coronavirus Disease (COVID-19): Sol Scientiae, Illustra Nos. Am J Roentgenol [Internet]. 2020 Dec;215(6):1342–50. Available from: <https://www.ajronline.org/doi/10.2214/AJR.20.23391>
22. Bellini D, Panvini N, Carbone I, Rengo M, Wang CL, Mileto A. Diagnostic Yield of Computed Tomography for the Identification of Coronavirus Disease 2019 Using Repeated Reverse Transcriptase Polymerase Chain Reaction Testing or Confirmed True-Negative State as Reference Standard: Systematic Review and Meta-Analysis. J Comput Assist Tomogr. 2020;44(6):812–20.
23. Hossein H, Ali KM, Hosseini M, Sarveazad A, Safari S, Yousefifard M. Value of chest computed tomography scan in diagnosis of COVID-19; a systematic review and meta-analysis. Clin Transl Imaging [Internet]. 2020 Dec 12;8(6):469–81. Available from: <https://doi.org/10.1007/s40336-020-00387-9>
24. Huang EPC, Sung CW, Chen CH, Fan CY, Lai PC, Huang YT. ¿Can computed tomography be a primary tool for COVID-19 detection? Evidence appraisal through meta-analysis. Crit Care. 2020;24(1):4–6.
25. Islam N, Salameh J-P, Leeflang MMG, Hooft L, McGrath TA, van der Pol CB, et al. Thoracic imaging tests for the diagnosis of COVID-19. Cochrane Database Syst Rev [Internet]. 2020 Nov 26;2020(9). Available from: <http://doi.wiley.com/10.1002/14651858.CD013639.pub2>
26. Kim H, Hong H, Yoon SH. Diagnostic Performance of CT and Reverse Transcriptase Polymerase Chain Reaction for Coronavirus Disease 2019: A Meta-Analysis. Radiology [Internet]. 2020 Sep;296(3):E145–55. Available from: <http://pubs.rsna.org/doi/10.1148/radiol.2020201343>
27. Lv M, Wang M, Yang N, Luo X, Li W, Chen X, et al. Chest computed tomography for the diagnosis of patients with coronavirus disease 2019 (COVID-19): a rapid review and meta-analysis. Ann Transl Med. 2020;8(10):622–622.
28. Shao JM, Ayuso SA, Deerenberg EB, Elhage SA, Augenstein VA, Heniford BT. A systematic review of CT chest in COVID-19 diagnosis and its potential application in a surgical setting. Color Dis. 2020;22(9):993–1001.
29. Simoni P, Bazzocchi A, Boitsios G, De Leucio A, Preziosi M, Aparisi Gómez MP. Chest computed tomography (CT) features in children with reverse transcription-polymerase chain reaction (RT-PCR)-confirmed COVID-19: A systematic review. J Med Imaging Radiat Oncol. 2020;64(5):649–59.
30. Taha MS, Milad P, Ahmed ANA, Hamdy TAH. Sensitivity of Chest Computed Tomography Scan in Diagnosis of SARS-CoV-2: A Meta-Analysis Study. SSRN Electron J. 2020;
31. Waller J V., Allen IE, Lin KK, Diaz MJ, Henry TS, Hope MD. The Limited Sensitivity of Chest Computed Tomography Relative to Reverse Transcription Polymerase Chain Reaction for Severe Acute Respiratory Syndrome Coronavirus-2 Infection:



A Systematic Review on COVID-19 Diagnostics. *Invest Radiol*. 2020;55(12):754–61.

32. Xu B, Xing Y, Peng J, Zheng Z, Tang W, Sun Y, et al. Chest CT for detecting COVID-19: a systematic review and meta-analysis of diagnostic accuracy. *Eur Radiol* [Internet]. 2020 Oct 15;30(10):5720–7. Available from: <http://link.springer.com/10.1007/s00330-020-06934-2>

Estudios para todos los coronavirus (no específico para COVID-19)

33. Floriano I, Silvinato A, Bernardo WM, Reis JC, Soledade G. Accuracy of the Polymerase Chain Reaction (PCR) test in the diagnosis of acute respiratory syndrome due to coronavirus: a systematic review and meta-analysis. *Rev Assoc Med Bras*. 2020;66(7):880–8.
34. Lin C, Ye R, Xia YL. A meta-analysis to evaluate the effectiveness of real-time PCR for diagnosing novel coronavirus infections. *Genet Mol Res*. 2015;14(4):15634–41.

Estudios no elegibles (no publicados/no RSL-metanálisis)

35. Caini S, Bellerba F, Corso F, Díaz-Basabe A, Natoli G, Paget J, et al. Meta-analysis of diagnostic performance of serological tests for SARS-CoV-2 antibodies up to 25 April 2020 and public health implications. *Eurosurveillance* [Internet]. 2020;25(23):1–5. Available from: <http://dx.doi.org/10.2807/1560-7917.ES.2020.25.23.2000980>
36. Castro R, Luz PM, Wakimoto MD, Veloso VG, Grinsztejn B, Perazzo H. COVID-19: a meta-analysis of diagnostic test accuracy of commercial assays registered in Brazil. *Brazilian J Infect Dis* [Internet]. 2020;24(2):180–7. Available from: <https://doi.org/10.1016/j.bjid.2020.04.003>
37. Chauhan N, Soni S, Gupta A, Jain U. New and developing diagnostic platforms for COVID-19: A systematic review. *Expert Rev Mol Diagn* [Internet]. 2020;20(9):971–83. Available from: <https://doi.org/10.1080/14737159.2020.1816466>
38. Karam M, Althuwaikh S, Alazemi M, Abul A, Hayre A, Alsaif A, et al. Chest CT versus RT-PCR for the Detection of COVID-19: Systematic Review and Meta-analysis of Comparative Studies. *medRxiv*. 2020;
39. Khatami F, Saatchi M, Zadeh SST, Aghamir ZS, Shabestari AN, Aghamir SMK. The exact place of PCR and chest CT in screening, staging, and diagnosis of Covid-19: A meta-analysis. *Res Sq*. 2020;1–22.
40. Lee RA, Herigon JC, Benedetti A, Pollock NR, Denkinger C. Performance of saliva, oropharyngeal swabs, and nasal swabs for SARS-CoV-2 molecular detection: A systematic review and meta-analysis. *medRxiv* [Internet]. 2020;preprint:2020.11.12.20230748. Available from: <http://medrxiv.org/content/early/2020/11/13/2020.11.12.20230748.abstract>



41. Suhandynata RT, Hoffman MA, Kelner MJ, McLawhon RW, Reed SL, Fitzgerald RL. Multi-Platform Comparison of SARS-CoV-2 Serology Assays for the Detection of COVID-19. *J Appl Lab Med*. 2020;5(6):1324–36.
42. Tu Y, Iqbal J, Care U. Evaluation of ID NOW and RT-PCR for Detection of SARS-CoV-2 in an Ambulatory Population. *medRxiv*. 2020;1–33.
43. Urwin SG, Lendrem BC, Suklan J, Green K, Graziadio S, Buckle P, et al. FebriDx point-of-care test in patients with suspected COVID-19: a pooled diagnostic accuracy study. *medRxiv* [Internet]. 2020;2020.10.15.20213108. Available from: <https://doi.org/10.1101/2020.10.15.20213108>
44. Vengesai A, Maritha Kasambala, Tariro Mduluza-Jokonya, Thajasvarie Naicker, Francisca Mutapi, Takafira Mduluza HM. A systematic review on the diagnostic accuracy of serological tests for COVID-19. *Res Sq*. 2020;1–41.

Artículos con estudios primarios cruzados

45. Moura DTH de, McCarty TR, Ribeiro IB, Funari MP, Oliveira PVAG de, Miranda Neto AA de, et al. Diagnostic Characteristics of Serological-Based COVID-19 Testing: A Systematic Review and Meta-Analysis. *Clinics* [Internet]. 2020 Aug 6;75(13):1–11. Available from: <https://www.clinicsjournal.com/article/diagnostic-characteristics-of-serological-based-covid-19-testing-a-systematic-review-and-meta-analysis/>
46. Riccò M, Ferraro P, Gualerzi G, Ranzieri S, Henry BM, Said Y Ben, et al. Point-of-Care Diagnostic Tests for Detecting SARS-CoV-2 Antibodies: A Systematic Review and Meta-Analysis of Real-World Data. *J Clin Med*. 2020;9(5):1515.

Estudios con alto riesgo de sesgos

47. Duarte ML, Santos LR dos, Contencas AC de S, Iared W, Peccin MS, Atallah ÁN. Reverse-transcriptase polymerase chain reaction versus chest computed tomography for detecting early symptoms of COVID-19. A diagnostic accuracy systematic review and meta-analysis. *Sao Paulo Med J* [Internet]. 2020 Oct;138(5):422–32. Available from: <https://www.ssrn.com/abstract=3588538>
48. Mahendiratta S, Batra G, Sarma P, Kumar H, Bansal S, Kumar S, et al. Molecular diagnosis of COVID-19 in different biologic matrix, their diagnostic validity and clinical relevance: A systematic review. *Life Sci* [Internet]. 2020;258(June):118207. Available from: <https://doi.org/10.1016/j.lfs.2020.118207>
49. Kucirka LM, Lauer SA, Laeyendecker O, Boon D, Lessler J. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction-Based SARS-CoV-2 Tests by Time Since Exposure. *Ann Intern Med*. 2020;173(4):262–7.



La salud
es de todos

Minsalud



Anexo 5. Evaluación del riesgo de sesgo de los estudios incluidos

Evaluación de calidad de RS con metanálisis con la herramienta IQWiG

Estudios incluidos

Estudio		Böger et al. 2021 (21)	Guo et al. 2020 (22)	Kontou et al. 2020 (23)	Lisboa et al. 2020 (24)
Criterio	Evaluador	Juicio			
1. ¿Se ha establecido <i>a priori</i> la pregunta?	AMHT	No (objetivo)	No (objetivo)	No (objetivo)	No (objetivo)
<ul style="list-style-type: none"> Descripción clara de la pregunta. Transferencia a una hipótesis estadística. Explicación de las desviaciones a partir del plan establecido originalmente. 	KCP	No	No	No	No (Parcialmente)
2. ¿Se ha explicado suficientemente la justificación para el uso de una comparación indirecta?	AMHT	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)
	KCP	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)	NA (No se hizo comparación indirecta)
3. ¿Se ha explicado suficientemente la selección de un comparador común en lugar de una comparación directa?	AMHT	NA (Se realizó comparación directa)	NA (Se realizó comparación directa)	NA (Se realizó comparación directa)	NA (Se realizó comparación directa)
	KCP	NA (Comparación directa)	NA (Comparación directa)	NA (Comparación directa)	NA (Comparación directa)
4. ¿Se ha realizado una revisión sistemática y exhaustiva de la literatura, y esta se ha descrito en detalle?	AMHT	Si	No (falta información detallada)	Si	Si
<ul style="list-style-type: none"> ¿Para la intervención de interés primario? ¿Para el comparador común? 	KCP	Si	Si (Parcialmente)	Si (Parcialmente)	Si
5. ¿Se han usado y descrito criterios de inclusión y exclusión inicialmente definidos?	AMHT	Si	Si	Si	Si
	KCP	Si	Si	Si	Si
6. ¿Hubo un reporte completo de todos los datos relevantes de los estudios?	AMHT	No	Si	Si	Si
<ul style="list-style-type: none"> Características de todos los estudios incluidos. Evaluación de todos los estudios incluidos. Gráficos de la red, descripción de la geometría de la red. <ul style="list-style-type: none"> Para todos los desenlaces relevantes, comparaciones y subgrupos: <ul style="list-style-type: none"> Resultados individuales de todos los estudios (las estimaciones del efecto y sus correspondientes intervalos de confianza). Estimaciones del efecto e intervalos de confianza a partir de los metanálisis en parejas. 	KCP	Si (Parcialmente)	Si	Si	Si
7. ¿Se han investigado los supuestos principales, y se han tratado adecuadamente los resultados de esta investigación?	AMHT	Si	Si	No (No se describe este análisis)	Si
<ul style="list-style-type: none"> Similitud (transitividad). Homogeneidad. Consistencia (coherencia). 	KCP	Si (Parcialmente)	Si (Parcialmente)	No	Si
8. ¿Se han usado herramientas estadísticas adecuadas, y ellas se han descrito con suficiente detalle?	AMHT	Si (los necesarios de acuerdo con el análisis propuesto)	Si (los necesarios de acuerdo con el análisis propuesto)	Si	Si
<ul style="list-style-type: none"> Uso de comparaciones indirectas ajustadas. Tratamiento de estudios con múltiples grupos. Detalles técnicos (especialmente, cuando se usan modelos bayesianos). <ul style="list-style-type: none"> Código de programación. Análisis de sensibilidad. 	KCP	Si (Parcialmente)	Si (Parcialmente)	Si	Si
9. ¿Se han descrito las limitaciones y estas se han discutido suficientemente?	AMHT	Si	Si	Si	Si
<ul style="list-style-type: none"> Calidad y exhaustividad de la base de datos. Incertidumbres metodológicas, análisis de sensibilidad. Conflictos con los supuestos principales. 	KCP	Si (Parcialmente)	Si (Parcialmente)	Si	Si
Valoración global de la calidad metodológica		Moderada	Moderada	Moderada	Moderada

NA: no aplica



Estudio		Mustafá et al. 2020 (26)	Mekonnen et al. 2020 (25)	Riccó et al. 2020 (27)	Wang et al. 2020 (28)	Zhang et al. 2020 (29)
Criterio	Evaluador	Juicio				
1. ¿Se ha establecido <i>a priori</i> la pregunta? <ul style="list-style-type: none"> Descripción clara de la pregunta. Transferencia a una hipótesis estadística. Explicación de las desviaciones a partir del plan establecido originalmente. 	AMHT	Si	No (objetivo)	No (objetivo)	No	No (objetivo)
	KCP	Si (Parcialmente)	No	No	No	No
2. ¿Se ha explicado suficientemente la justificación para el uso de una comparación indirecta?	AMHT	No (No se explica claramente)	NA (No se hizo comparación indirecta)	No (No se explica claramente)	NA (No se hizo comparación indirecta)	No (No se explica claramente)
	KCP	No	NA (No se hizo comparación indirecta)	No	NA (No se hizo comparación indirecta)	No
3. ¿Se ha explicado suficientemente la selección de un comparador común en lugar de una comparación directa?	AMHT	Si	No	NA (Se realizó comparación directa)	NA (Se realizó comparación directa)	NA (Se realizó comparación directa)
	KCP	Si	No	NA (Comparación directa)	NA (Comparación directa)	NA (Comparación directa)
4. ¿Se ha realizado una revisión sistemática y exhaustiva de la literatura, y esta se ha descrito en detalle? <ul style="list-style-type: none"> ¿Para la intervención de interés primario? ¿Para el comparador común? 	AMHT	Si	No (Falta información detallada)	No (Falta información detallada)	Si	No (Falta información detallada)
	KCP	Si	Si	Si (Parcialmente)	Si	Si (Parcialmente)
5. ¿Se han usado y descrito criterios de inclusión y exclusión inicialmente definidos?	AMHT	Si	Si	Si	Si	Si
	KCP	Si	Si	Si	Si	Si
6. ¿Hubo un reporte completo de todos los datos relevantes de los estudios? <ul style="list-style-type: none"> Características de todos los estudios incluidos. Evaluación de todos los estudios incluidos. Gráficos de la red, descripción de la geometría de la red. Para todos los desenlaces relevantes, comparaciones y subgrupos: <ul style="list-style-type: none"> Resultados individuales de todos los estudios (las estimaciones del efecto y sus correspondientes intervalos de confianza). Estimaciones del efecto e intervalos de confianza a partir de los metanálisis en parejas. 	AMHT	Si	Si	Si	Si	Si
	KCP	Si (Parcialmente)	Si	Si	Si	Si
7. ¿Se han investigado los supuestos principales, y se han tratado adecuadamente los resultados de esta investigación? <ul style="list-style-type: none"> Similitud (transitividad). Homogeneidad. Consistencia (coherencia). 	AMHT	Si	Si	Si	Si	Si
	KCP	Si	Si	Si (Parcialmente)	Si (Parcialmente)	Si (Parcialmente)
8. ¿Se han usado herramientas estadísticas adecuadas, y ellas se han descrito con suficiente detalle? <ul style="list-style-type: none"> Uso de comparaciones indirectas ajustadas. Tratamiento de estudios con múltiples grupos. Detalles técnicos (especialmente, cuando se usan modelos bayesianos). <ul style="list-style-type: none"> Código de programación. Análisis de sensibilidad. 	AMHT	Si	Si	Si	Si	Si (los necesarios de acuerdo con el análisis propuesto)
	KCP	Si	Si	Si (Parcialmente)	Si	Si
9. ¿Se han descrito las limitaciones y estas se han discutido suficientemente? <ul style="list-style-type: none"> Calidad y exhaustividad de la base de datos. 	AMHT	Si	Si	Si	Si	Si
	KCP	Si	Si	Si	Si	Si



<ul style="list-style-type: none"> ▪ Incertidumbres metodológicas, análisis de sensibilidad. ▪ Conflictos con los supuestos principales. 						
Valoración global de la calidad metodológica		Moderada	Moderada	Moderada	Moderada	Moderada



Evaluación de calidad de RS con la herramienta ROBIS

Estudios incluidos

Título de la revisión: False-negative results of initial RT-PCR assays for COVID-19: A systematic review (17).
Autor principal y año de publicación: Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, et al. PLoS One. 2020;15(12):e0242958.
Nombre de la persona que aplicó la evaluación: Ana Milena Herrera T/Katherinne Cortes Palacio
Fecha de la evaluación: 25/12/2020

Fase 2: identificar preocupaciones con el proceso de revisión

AMHT	KCP
Dominio 1: criterios de elegibilidad de los estudios	
Se seleccionaron estudios observacionales de pruebas diagnósticas, cohortes e incluso series de casos. Se excluyeron estudios de concordancia, realizado en especímenes y no en pacientes, comparación de métodos de extracción de RNA, points of care entre otros. En cuanto a la población, se incluyeron pacientes sin límites de edad, sexo o ubicación geográficas, sin embargo, se excluyeron poblaciones especiales como pacientes con cáncer, diabetes, esclerosis múltiple, mujeres embarazadas o pacientes quirúrgicos. No realizan restricción por idioma o estado de publicación. Revisando las estrategias de búsquedas se limitaron al año 2019 lo cual se considera pertinente con referencia a la novedad del tema. Se excluyeron estudios que no reportaran valores numéricos completos.	
1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos? Si	1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos? Si
1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión? Si	1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión? Si
1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades? Si	1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades? Si
1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)? Si	1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)? Si
1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)? Si	1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)? Si
Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Bajo	Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 2: identificación y selección de los estudios	
Se usaron diferentes bases de datos electrónicas generales y repositorios de artículos en prepublicación o en curso específicos para el tema de la revisión. Se realizó búsqueda por referencias. Proporcionan las estrategias de búsquedas en texto suplementario. No hubo restricción de idioma si para el año 2019 lo cual es pertinente dada la novedad del tema. Dos revisores de forma independiente realizaron el tamizaje por título y resumen, resolviendo conflictos por consenso.	
2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si	2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si
2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Si	2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Si
2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Si	2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Si
2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? Si	2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? Si
2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si	2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si
Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo	Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo



Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 3: recolección de datos y evaluación de los estudios	
La extracción de los datos cualitativos y cuantitativos fue realizada por un solo investigador, pero un segundo revisó toda la información extraída. No se informa de una herramienta de extracción de datos con prueba piloto previa a su aplicación. Para riesgo de sesgo se utilizó QUADAS 2, para series de casos el instrumento del grupo Joanna Briggs para este tipo de diseño. Esta evaluación se llevó a cabo de forma pareada e independiente y se resolvieron desacuerdos por consenso.	
3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? No hay información 3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si 3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si 3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si 3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si	3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Probablemente si 3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Probablemente si 3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si 3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si 3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si
Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo	Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 4: síntesis y resultados	
Se realizó síntesis cualitativa de la evidencia. Se realizó exploración de las posibles fuentes de heterogeneidad entre los estudios.	
4.1 ¿La síntesis incluyó todos los estudios que debería? Si 4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si 4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si 4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si 4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si	4.1 ¿La síntesis incluyó todos los estudios que debería? Si 4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si 4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si 4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si 4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si 4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si
Preocupaciones respecto a la síntesis y resultados Bajo	Preocupaciones respecto a la síntesis y resultados Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones

Fase 3: Juzgar el riesgo de sesgos

Resumen de las preocupaciones identificadas durante la fase 2 de la evaluación:

Dominio	Preocupación	Justificación para la preocupación
1. Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios	Bajo	No hay preocupaciones
2. Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios	Bajo	No hay preocupaciones
3. Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios	Bajo	No hay preocupaciones



4.	Preocupaciones respecto a la síntesis y resultados	Bajo	No hay preocupaciones
Riesgo de sesgos en la revisión			
Describa si las conclusiones fueron sustentadas por la evidencia:			
A.	¿En la interpretación de los resultados se abordaron todas las preocupaciones identificadas en los dominios 1 a 4?	Si	
B.	¿Se consideró apropiadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión?	Si	
C.	¿Los revisores evitaron enfatizar los resultados con base en su significancia estadística?	Si	
Riesgo de sesgos en la revisión			
Bajo			
Justificación para el riesgo: Bajo riesgo. No se detectaron fallas que propiciaran una valoración diferente			

Título de la revisión: Antibody tests for identification of current and past infection with SARS-CoV-2 (18).
Autor principal y año de publicación: Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Cochrane Database Syst Rev. 2020 Jun 25;2.
Nombre de la persona que aplicó la evaluación: Ana Milena Herrera T/Katherinne Cortes Palacio
Fecha de la evaluación: 25/12/2020

Fase 2: identificar preocupaciones con el proceso de revisión

AMHT	KCP
Dominio 1: criterios de elegibilidad de los estudios	
Se especifican criterios amplios para incluir todos los grupos de pacientes y todas las variaciones de las pruebas	
1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos?	1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos?
Si	Si
1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión?	1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión?
Si	Si
1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades?	1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades?
Si	Si
1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)?	1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)?
Si	Si
1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)?	1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)?
Si	Si
Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios	Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios
Bajo	Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 2: identificación y selección de los estudios	
Se utilizan diferentes bases de datos específicas para el tema en cuestión, sin límite de idioma, sin filtros, solo restringido a 2019-2020 lo cual es apropiado por la novedad del tema. Se investigan otros tipos de fuentes como estudios de compañías productoras de las pruebas y contacto con otros grupos de investigadores.	
2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados?	2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados?
Si	Si
2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes?	2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes?
Si	Si
2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible?	2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible?
Si	Si
2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma?	2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma?



2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si	2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si
Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo	Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 3: recolección de datos y evaluación de los estudios	
La tamización por título y resumen se dio por parte de un equipo especializado en revisiones sistemáticas de la Universidad de Birmingham, dos evaluadores realizaron tamizaje de forma pareada e independiente y los desacuerdos resueltos por un tercer evaluador. Se realizó la extracción de datos y aplicación del QUADAS 2 de manera pareada e independiente por dos evaluadores quienes resolvieron desacuerdos por consenso. Se contactaron a los autores en caso de datos incompletos.	
3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si	3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si
3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si	3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si
3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Si	3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si
3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si	3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si
3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si	3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si
Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo	Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 4: síntesis y resultados	
La síntesis cualitativa de la evidencia se informó de forma clara y coherente. Se incluye descripción de la calidad metodológica de los estudios primarios. Se discute ampliamente las fuentes de heterogeneidad.	
4.1 ¿La síntesis incluyó todos los estudios que debería? Probablemente si	4.1 ¿La síntesis incluyó todos los estudios que debería? Probablemente si
4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si	4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Probablemente si
4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si	4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si
4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si	4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si
4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si	4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si
4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si	4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si
Preocupaciones respecto a la síntesis y resultados Bajo	Preocupaciones respecto a la síntesis y resultados Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones

Fase 3: Juzgar el riesgo de sesgos

Resumen de las preocupaciones identificadas durante la fase 2 de la evaluación:

Dominio	Preocupación	Justificación para la preocupación
1. Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios	Bajo	No hay preocupaciones



2.	Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios	Bajo	No hay preocupaciones
3.	Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios	Bajo	No hay preocupaciones
4.	Preocupaciones respecto a la síntesis y resultados	Bajo	No hay preocupaciones
Riesgo de sesgos en la revisión			
Describa si las conclusiones fueron sustentadas por la evidencia:			
A.	¿En la interpretación de los resultados se abordaron todas las preocupaciones identificadas en los dominios 1 a 4?		
	Si		
B.	¿Se consideró apropiadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión?		
	Si		
C.	¿Los revisores evitaron enfatizar los resultados con base en su significancia estadística?		
	Si		
Riesgo de sesgos en la revisión			
Bajo			
Justificación para el riesgo: Bajo			

Título de la revisión: Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection (19).
Autor principal y año de publicación: Dinnes J, Deeks JJ, Adriano A, Berhane S, Davenport C, Dittrich S, et al. Cochrane Database Syst Rev. 2020 Aug 26;8:CD013705.
Nombre de la persona que aplicó la evaluación: Ana Milena Herrera T/Katherine Cortes Palacio
Fecha de la evaluación: 25/12/2020

Fase 2: identificar preocupaciones con el proceso de revisión

AMHT	KCP
Dominio 1: criterios de elegibilidad de los estudios	
Los criterios de elegibilidad son claros y explícitos. Se incluyeron cualquier diseño de estudio que reportara estimaciones de exactitud diagnóstica. La población, prueba índice y de referencia son claramente establecidas y sustentadas. No hubo restricción por tipo de diseño, se excluyeron aquellos con menos de 10 participantes, pero ofrecen por qué esta consideración. Se utilizaron el repositorio específico para COVID 19 de Cochrane y de la Universidad de Bern. EMBASE y otras bases también fueron consultadas. No hubo restricción por idioma o estado de publicación.	
1.6 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos? Si	1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos? Si
1.7 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión? Si	1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión? Si
1.8 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades? Si	1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades? Si
1.9 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)? Probablemente si	1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)? Probablemente si
1.10 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)? Si	1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)? Si
Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Bajo	Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 2: identificación y selección de los estudios	
Se proporcionan las bases, términos y estrategias de búsqueda utilizadas. Utilizan otras fuentes de información, pero no realizan búsqueda de bola de nieve o bases de datos de literatura gris. La búsqueda solo se restringió al año de publicación. La selección inicial por título y abstract fue pareada e independiente y un tercer evaluador resolvió desacuerdos	



<p>2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si</p> <p>2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Si</p> <p>2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Si</p> <p>2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? Si</p> <p>2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si</p> <p>Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo</p> <p>Justificación para la preocupación: No hay preocupaciones</p>	<p>2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si</p> <p>2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Probablemente si</p> <p>2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Si</p> <p>2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? Si</p> <p>2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si</p> <p>Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo</p> <p>Justificación para la preocupación: No hay preocupaciones</p>
<p>Dominio 3: recolección de datos y evaluación de los estudios</p> <p>La tamización por título y resumen se dio por parte de un equipo especializado en revisiones sistemáticas de la Universidad de Birmingham, dos evaluadores realizaron tamizaje de forma pareada e independiente y los desacuerdos resueltos por un tercer evaluador. De igual manera se llevó la revisión por texto completo. La extracción de datos fue pareada revisada por un tercer investigador. Se utilizó QUADAS 2 para evaluación de calidad, de igual manera pareada e independiente por dos evaluadores quienes resolvieron desacuerdos por consenso.</p>	
<p>3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si</p> <p>3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si</p> <p>3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si</p> <p>3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si</p> <p>3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si</p> <p>Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo</p> <p>Justificación para la preocupación: No hay preocupaciones</p>	<p>3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si</p> <p>3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Probablemente si</p> <p>3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si</p> <p>3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si</p> <p>3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si</p> <p>Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo</p> <p>Justificación para la preocupación: No hay preocupaciones</p>
<p>Dominio 4: síntesis y resultados</p> <p>Se realizó una adecuada síntesis cualitativa de la evidencia. Se exploraron las diferentes fuentes de heterogeneidad de forma amplia.</p>	
<p>4.1 ¿La síntesis incluyó todos los estudios que debería? Si</p> <p>4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si</p> <p>4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si</p> <p>4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si</p> <p>4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si</p> <p>4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis?</p>	<p>4.1 ¿La síntesis incluyó todos los estudios que debería? Probablemente si</p> <p>4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si</p> <p>4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si</p> <p>4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Si</p> <p>4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? Si</p> <p>4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis?</p>



Si	Si
Preocupaciones respecto a la síntesis y resultados	Preocupaciones respecto a la síntesis y resultados
Bajo	Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones

Fase 3: Juzgar el riesgo de sesgos

Resumen de las preocupaciones identificadas durante la fase 2 de la evaluación:

Dominio	Preocupación	Justificación para la preocupación
5. Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios	Bajo	No hay preocupaciones
6. Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios	Bajo	No hay preocupaciones
7. Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios	Bajo	No hay preocupaciones
8. Preocupaciones respecto a la síntesis y resultados	Bajo	No hay preocupaciones
Riesgo de sesgos en la revisión		
Describa si las conclusiones fueron sustentadas por la evidencia:		
D. ¿En la interpretación de los resultados se abordaron todas las preocupaciones identificadas en los dominios 1 a 4?	Si	
E. ¿Se consideró apropiadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión?	Si	
F. ¿Los revisores evitaron enfatizar los resultados con base en su significancia estadística?	Si	
Riesgo de sesgos en la revisión		
Bajo		
Justificación para el riesgo: Bajo		

Título de la revisión: Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review (20).

Autor principal y año de publicación: Jarrom D, Elston L, Washington J, Prettyjohns M, Cann K, Myles S, et al. BMJ Evidence-Based Med. 2020;0(0): bmjebm-2020-111511.

Nombre de la persona que aplicó la evaluación: Ana Milena Herrera T/Katherine Cortes Palacio

Fecha de la evaluación: 26/12/2020

Fase 2: identificar preocupaciones con el proceso de revisión

AMHT	KCP
Dominio 1: criterios de elegibilidad de los estudios	
Los criterios de elegibilidad son claros y explícitos.	
1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos?	1.1 ¿La revisión obedeció a objetivos y criterios de elegibilidad preestablecidos?
Si	Si
1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión?	1.2 ¿Los criterios de elegibilidad fueron apropiados para la pregunta de la revisión?
Si	Si
1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades?	1.3 ¿Los criterios de elegibilidad fueron planteados sin ambigüedades?
Si	Si
1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)?	1.4 ¿Fueron apropiadas todas las restricciones en los criterios de elegibilidad basadas en las características de los estudios (p.ej., fecha, tamaño de la muestra, calidad del estudio, desenlaces medidos)?
Si	Si
1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)?	1.5 ¿Fue apropiada cualquier restricción en los criterios de elegibilidad basada en las fuentes de información (p.ej., estado o formato de publicación, idioma, disponibilidad de los datos)?
Probablemente si	No



Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Bajo	Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios Alto
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: A pesar de que los criterios de elegibilidad son claros y acorde a la PICO, solo se incluyeron artículos en inglés.
Dominio 2: identificación y selección de los estudios	
Se usaron un amplio rango de bases de datos, incluyendo de literatura gris. Se siguieron los pasos indicados para la selección y tamización de los artículos.	
2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si	2.1 ¿La búsqueda incluyó un rango apropiado de bases de datos/fuentes electrónicas para reportes publicados y no publicados? Si
2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Si	2.2 ¿Se usaron métodos adicionales a la búsqueda en bases de datos para identificar reportes relevantes? Si
2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Si	2.3 ¿Los términos y la estructura de la estrategia de búsqueda probablemente recuperaron tantos estudios elegibles como fue posible? Probablemente si
2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? Probablemente no	2.4 ¿Fueron apropiadas las restricciones basadas en fecha, formato de publicación o idioma? No
2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si	2.5 ¿Se hicieron esfuerzos para minimizar el error en la selección de los estudios? Si
Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Bajo	Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios Alto
Justificación para la preocupación: Se limitaron al idioma inglés	Justificación para la preocupación: Limitación por idioma de publicación.
Dominio 3: recolección de datos y evaluación de los estudios	
La recolección de datos fue pareada e independiente, al igual que la aplicación del QUADAS 2. Se ofrece descripción de las características de cada estudio incluido.	
3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si	3.1 ¿Se hicieron esfuerzos para minimizar el error en la recolección de los datos? Si
3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si	3.2 ¿Estuvieron disponibles suficientes características de los estudios, para permitirles a los autores de la revisión y a los lectores interpretar los resultados? Si
3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Si	3.3 ¿Se recolectaron todos los resultados relevantes de los estudios, para su uso en la síntesis? Probablemente si
3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si	3.4 ¿El riesgo de sesgos (o calidad metodológica) fue evaluado formalmente usando criterios apropiados? Si
3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si	3.5 ¿Se hicieron esfuerzos para minimizar el error en la evaluación del riesgo de sesgos? Si
Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo	Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios Bajo
Justificación para la preocupación: No hay preocupaciones	Justificación para la preocupación: No hay preocupaciones
Dominio 4: síntesis y resultados	
Se realizó síntesis cualitativa de la evidencia. Se realizó exploración de las posibles fuentes de heterogeneidad entre los estudios.	
4.1 ¿La síntesis incluyó todos los estudios que debería? Si	4.1 ¿La síntesis incluyó todos los estudios que debería? Si
4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si	4.2 ¿Todos los análisis predefinidos fueron reportados o las desviaciones en este sentido fueron explicadas? Si
4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si	4.3 ¿La síntesis fue apropiada dada la naturaleza y similitud en las preguntas de investigación, el diseño de los estudios y los desenlaces a través de los estudios incluidos? Si



<p>4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? Probablemente si</p> <p>4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? No hay información</p> <p>4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si</p>	<p>4.4 ¿La variación entre los estudios (heterogeneidad) fue mínima o se abordó en la síntesis? No</p> <p>4.5 ¿Los resultados fueron sólidos (p.ej., como se demostró mediante un gráfico de embudo o un análisis de sensibilidad)? No hay información</p> <p>4.6 ¿Los sesgos en los estudios primarios fueron mínimos o se abordaron en la síntesis? Si</p>
<p>Preocupaciones respecto a la síntesis y resultados Poco claro</p> <p>Justificación para la preocupación: No hay detalles acerca de la evaluación de sensibilidad</p>	<p>Preocupaciones respecto a la síntesis y resultados Alto</p> <p>Justificación para la preocupación: Finalmente ofrecen un reporte cualitativo de los resultados. No realizan metanálisis, pero no informan si fue por las fuentes de heterogeneidad, tema que no fue abordado.</p>

Fase 3: Juzgar el riesgo de sesgos

Resumen de las preocupaciones identificadas durante la fase 2 de la evaluación:

Dominio	Preocupación	Justificación para la preocupación
1. Preocupaciones respecto a la especificación de los criterios de elegibilidad de los estudios	Bajo/Alto	No hay preocupaciones/ Justificación para la preocupación: A pesar de que los criterios de elegibilidad son claros y acorde a la PICO, solo se incluyeron artículos en inglés.
2. Preocupaciones respecto a los métodos usados para identificar o seleccionar los estudios	Bajo/Alto	Se limitaron al idioma inglés/ Limitación por idioma de publicación.
3. Preocupaciones respecto a los métodos usados para recolectar los datos y evaluar los estudios	Bajo/Bajo	No hay preocupaciones
4. Preocupaciones respecto a la síntesis y resultados	Poco claro/Alto	No hay detalles acerca de la evaluación de sensibilidad/ Finalmente ofrecen un reporte cualitativo de los resultados. No realizan metanálisis, pero no informan si fue por las fuentes de heterogeneidad, tema que no fue abordado.
Riesgo de sesgos en la revisión		
Describa si las conclusiones fueron sustentadas por la evidencia:		
<p>A. ¿En la interpretación de los resultados se abordaron todas las preocupaciones identificadas en los dominios 1 a 4? Si</p> <p>B. ¿Se consideró apropiadamente la relevancia de los estudios identificados para la pregunta de investigación de la revisión? Probablemente si</p> <p>C. ¿Los revisores evitaron enfatizar los resultados con base en su significancia estadística? Si</p>		
<p>Riesgo de sesgos en la revisión Poco claro</p> <p>Justificación para el riesgo: Existen fallas metodológicas que son subsanadas en el análisis cualitativo. Aunque se limitaron los estudios solo al idioma inglés, es una preocupación que no justifica su exclusión dada que la mayoría de los estudios de alta calidad metodológica sobre COVID19 están publicados en este idioma.</p>		

[illegible]



Metodos																
Protocolo y registro	6	Se indica dónde se puede acceder al protocolo de revisión (por ejemplo, dirección web) y proporcione el número de registro del ensayo si está disponible	Si	No	No	No	No	No	No	No	Si	No	Si	No	No	Si
Criterio de elegibilidad	7	Se especifican las características del estudio (participantes, entorno, prueba índice, estándares de referencia, condiciones objetivo y diseño del estudio) y características del informe (por ejemplo, años considerados, idioma, estado de publicación) utilizados como criterios de elegibilidad y fundamento.	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si
Fuentes de información	8	Se describen todas las fuentes de información (p. Ej., Bases de datos con fechas de cobertura, contacto con los autores del estudio para identificar estudios adicionales) en la búsqueda y la fecha de la última búsqueda.	Si	Parcial	Si	Si	Si	Si	Parcial	Si	Si	Si	Si	Si	Si	Si
Búsqueda	9	Se presentan estrategias de búsqueda completas para todas las bases de datos electrónicas y otras fuentes buscadas, incluidos los límites utilizados para que puedan repetirse	Si	Parcial	Si	Si	Parcial	Si	No	Si	Si	Si	No	Si	No	No
Selección de estudios	10	Se indica el proceso de selección de estudios (es decir, selección, elegibilidad, si está incluido en la revisión sistemática v. si	Si	Parcial	Si	Si	Parcial	Si	No	Si	Si	Si	Si	Si	Si	Si



		corresponde, incluido en el metanálisis).													
Proceso de recopilación de datos	11	Se describen los métodos de extracción de datos de los informes (por ejemplo, formularios piloto, de forma independiente, por duplicado) y cualquier proceso para obtener y confirmar datos de los investigadores.	Parcial	Si	Si	Si	Parcial	Si	No	Si	Si	Si	No	Si	Si
Definiciones para la extracción de datos	12	Se proporcionan definiciones utilizadas en la extracción de datos y clasificaciones de condiciones objetivo, pruebas índice, estándares de referencia y otras características (p. Ej., Diseño del estudio, entorno clínico).	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Parcial
Riesgo de sesgo y aplicabilidad	13	Se describen los métodos utilizados para evaluar el riesgo de sesgo en estudios individuales y las preocupaciones con respecto a la aplicabilidad a la pregunta de revisión.	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	No	Si	Si
Medidas de precisión diagnóstica	14	Se indican las principales medidas de precisión diagnóstica informadas (p. Ej., Sensibilidad, especificidad) y establezca la unidad de evaluación (p. Ej., Por paciente frente a por lesión).	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Parcial
Síntesis de resultados	15	Se describen los métodos de manejo de los datos, combinando los resultados de los estudios y describiendo la variabilidad entre estudios. Esto podría incluir, pero no se limita a (1) manejo de múltiples definiciones de la condición objetivo, (2) manejo de múltiples umbrales de positividad	Si	Si	Si	Si	Si	No	Si	Si	Si	Si	Parcial	Si	No



		de la prueba, (3) manejo de múltiples lectores de prueba de índice, (4) manejo de resultados de prueba indeterminados, (5) agrupación y comparación de pruebas, y (6) manejo de diferentes estándares de referencia.													
Metanálisis	16	Se informan los métodos estadísticos utilizados para los metanálisis, si se realizan.	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si
Análisis adicionales	17	Se describen los métodos de los análisis adicionales (p. Ej., Análisis de sensibilidad o de subgrupos, metarregresión) si se realizaron, indicando cuáles fueron preespecificados.	Si	Si	Si	Si	Si	No	Si	Si	Si	Si	No	No	No
Resultados															
Selección de estudios	18	Se proporciona el número de estudios examinados, evaluados para la elegibilidad, incluidos en la revisión e incluidos en el metanálisis si corresponde, con los motivos de las exclusiones en cada etapa, idealmente con un diagrama de flujo.	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si
Características del estudio	19	Para cada estudio incluido, proporcionan citas y presentan las características clave, incluidas (1) características de los participantes (presentación, pruebas previas), (2) entorno clínico, (3) diseño del estudio, (4) definición de la condición objetivo, (5) prueba índice, (6) estándar de referencia,	Si	Si	Si	Si	Parcial	Si	Si	Si	Si	Si	Si	Si	Si

[illegible]



Resumen	24	Se resumen los principales hallazgos, incluida la solidez de la evidencia	Si	Parcial	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si
Limitaciones	25	Se discuten las limitaciones de los estudios incluidos (p. Ej., Riesgo de sesgo y preocupaciones con respecto a la aplicabilidad) y del proceso de revisión (p. Ej., Recuperación incompleta de la investigación identificada).	Si	Si	Si	Si	No	Si	No	Si	Si	Parcial	Si	Si	Si
Conclusiones	26	Se proporcionan una interpretación general de los resultados en el contexto de otra evidencia. Discuta las implicaciones para la investigación y la práctica clínica futuras (por ejemplo, el uso previsto y la función clínica de la prueba índice).	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Si	Parcial
Otras															
Fondos	27	Para la revisión sistemática, se describen las fuentes de financiamiento y otro tipo de apoyo y el papel de los financiadores.	Si	Si	Si	Si	No	Si	Si	Si	Si	Si	No	Si	Si
Total			26,5	23,5	26	26	19	21,5	17,5	26,5	25,5	26	19	22,5	20,5

DTA: Exactitud de la prueba de diagnóstico



Anexo 7. Perfiles de evidencia GRADE

Pruebas moleculares (NAT)

Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas moleculares que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Böger et al. 2021. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19 (21).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: RT-PCR (sangre); Comparador: RT-PCR (RdRp, RdRp-P2, ORF1ab – en diferentes tipos de muestras)

Sensibilidad agrupada: 0.073 (IC 95%:0.041-0.117) // Especificidad agrupada: RT-PCR

Verdaderos positivos	3 (92)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	7 (4 a 12)	15 (8 a 23)	37 (21 a 59)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								93 (88 a 96)	185 (177 a 192)	463 (441 a 479)		

Intervención: RT-PCR (esputo); Comparador: RT-PCR (RdRp, RdRp-P2, ORF1ab – en diferentes tipos de muestras)

Sensibilidad agrupada: 0.972 (IC 95%:0.903-0.997) // Especificidad agrupada: RT-PCR

Verdaderos positivos	2 (91)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	97 (90 a 100)	194 (181 a 199)	486 (452 a 499)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								3 (0 a 10)	6 (1 a 19)	14 (1 a 48)		



Intervención: RT-PCR (saliva); Comparador: RT-PCR (RdRp, RdRp-P2, ORF1ab – en diferentes tipos de muestras)

Sensibilidad agrupada: 0.733 (IC 95%:0.681-0.780) // Especificidad agrupada: RT-PCR

Verdaderos positivos	4 (287)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	73 (68 a 78)	147 (136 a 156)	367 (341 a 390)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								27 (22 a 32)	53 (44 a 64)	133 (110 a 159)		

Verdaderos positivos	2 (38)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	62 (55 a 70))	125 (109 a 139)	312 (273 a 348)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								38 (30 a 45)	75 (61 a 91)	188 (152 a 227)		

Explicaciones

- a. 26% de los estudios seleccionados no incluyeron descripción de los métodos de selección de los participantes. Se incluyó estudio de casos y controles y un 46% incluyeron pacientes previamente diagnosticados.
- b. Los estudios incluidos difieren en términos de tamaño, riesgo de sesgo y validez externa. Se encuentran altas tasas de heterogeneidad entre los estudios, probablemente debido a diferencias en los métodos, las características de los pacientes y las muestras utilizadas. El estándar de referencia es variable y no es detallada la diferencia entre RT-PCR como prueba índice vs como prueba de referencia para la definición de los casos.
- c. No se presenta evaluación de sesgo de publicación. No se identifican factores de confusión por limitada información de los estudios primarios.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas moleculares que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Dinnes et al. 2020. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection (Review) (19).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: Pruebas rápidas moleculares (ID NOW; Xpert Xpress); Comparador: RT-PCR

Rango de sensibilidad agrupada: 0.68 - 1 // Rango de especificidad: 0.92 - 1

Verdaderos positivos	11 (2255)	Observacionales Cohortes	Muy serio ^a	Serio	No es serio	Serio	No es serio	68 a 100	136 a 200	340 a 500	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								0 a 32	0 a 64	0 a 160		
Verdaderos negativos	11 (2255))	Observacionales Cohortes	Muy serio ^a	Serio	No es serio	Serio	No es serio	828 a 900	736 a 800	460 a 500	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								0 a 72	0 a 64	0 a 40		

Explicaciones

- a. Los autores consideraron que el riesgo de sesgo era alto en el 50% de los estudios por cómo seleccionaron las muestras y un 72% porque consideraron que una RT-PCR negativa era suficiente para confirmar la ausencia de infección por COVID-19. La falta de detalles proporcionados por los artículos significó que no se pudiera evaluar claramente si había riesgo de sesgo a través de la realización de la prueba índice (61%), o de la forma en que se realizó y analizó el estudio (56%). Se consideró que existían grandes preocupaciones sobre la aplicabilidad de la evidencia relacionada con los participantes, con la prueba índice y con el estándar de referencia.
- b. Los estudios no utilizaron los métodos más fiables o no proporcionaron suficiente información para juzgar su metodología. Esto puede afectar las estimaciones sobre la precisión de la prueba, pero no es posible identificar en qué medida.
- c. Variabilidad del espectro clínico, no fue posible determinar los días de evolución de la enfermedad y toma de la muestra en la mayoría de los estudios, amplitud en las estimaciones individuales de los estudios.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas moleculares que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Mustafa et al. 2020. Nucleic acid amplification tests on respiratory samples for the diagnosis of coronavirus infections: a systematic review and metaanalysis (26).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: NAAT (nucleic acid amplification test); Comparador: RT-PCR y criterios clínicos

Sensibilidad agrupada: 0.904 (95% CI: 0.837 a 0.945) // Especificidad agrupada: 0.989 (95% CI: 0.98 a 0.994)

Verdaderos positivos	29 (8742)	Observacionales: Cohortes retrospectiva y casos y controles	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	90 (84 a 95)	181 (167 a 189)	452 (419 a 473)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								10 (5 a 16)	19 (11 a 33)	48 (27 a 81)		
Verdaderos negativos	29 (8742)	Observacionales: Cohortes retrospectiva y casos y controles	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	890 (882 a 895)	791 (784 a 795)	495 (490 a 497)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								10 (5 a 18)	9 (5 a 16)	5 (3 a 10)		

Explicaciones

- a. Todos los artículos incluidos en el metanálisis (29) tienen alto riesgo de sesgo según valoración QUADAS-2 reportado por la revisión sistemática. Predomina alto riesgo de sesgo en dominios 1 y 2 con preocupaciones de aplicabilidad para estos dominios.
- b. Variabilidad en los métodos de ejecución de la prueba índice evaluada
- c. Sesgo de publicación no evaluado



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas moleculares que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Ricco et al. 2020. RT-qPCR assays based on saliva rather than on nasopharyngeal swabs are possible but should be interpreted with caution: results from a systematic review and metanalysis (27).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: RT-PCR (saliva); Comparador: RT-PCR (hisopado nasofaríngeo)

Sensibilidad agrupada: 0.834 (95% CI: 0.731 a 0.904) // Especificidad agrupada: 0.977 (95% CI: 0.938 a 0.992)

Verdaderos positivos	11 (1118)	Observacionales: Cohortes y casos y controles	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	83 (73 a 90)	167 (146 a 181)	417 (366 a 452)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								17 (10 a 27)	33 (19 a 54)	83 (48 a 134)		
Verdaderos negativos	11 (1118)	Observacionales: Cohortes y casos y controles	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	879 (844 a 893)	782 (750 a 794)	489 (469 a 496)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								21 (7 a 56)	18 (6 a 50)	11 (4 a 31)		

Explicaciones

- a. No hay evaluación de riesgo de sesgo reportada en la revisión sistemática.
- b. Amplitud del rango para la sensibilidad reportada por los estudios primarios.
- c. Sesgo de publicación detectado por asimetría de funnel plot y análisis de regresión.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas moleculares que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Jarrom et al. 2020. Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review (20).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: RT-LAMP /RT PCR /Amplificación isotérmica /Hibridación de ADN y detección electroquímica); Comparador: No Reporta

Rango de sensibilidad agrupada: 0.747 - 1 // Rango de especificidad: 0.887 - 1

Verdaderos positivos	5 (972)	Observacionales: Cohortes, series de casos retrospectivas.	Serio ^a	Serio ^b	Serio ^c	No es serio	No es serio	75 a 100	149 a 200	374 a 500	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								0 a 25	0 a 51	0 a 126		
Verdaderos negativos	5 (972)	Observacionales: Cohortes, series de casos retrospectivas.	Serio ^a	Serio ^b	Serio ^c	No es serio	No es serio	798 a 900	710 a 800	444 a 500	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								0 a 102)	0 a 90	0 a 56		

Explicaciones

- a. No se especifica evaluación de riesgo de sesgo para estudios cuya prueba índice fuera molecular.
- b. Variabilidad de los estudios, inclusión de estudios sin comparador, variabilidad en las pruebas de referencia y criterios diagnósticos para determinar la enfermedad



Pruebas inmunológicas (detección de anticuerpos y detección de antígenos virales)

Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Böger et al, 2020. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19 (21).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgM (suero, sangre, plasma); Comparador: RT-PCR

Sensibilidad agrupada: 0.770 (IC 95%: 0.745-0.795) // Especificidad agrupada: 0.933 (IC 95%: 0.886-0.965)

Verdaderos positivos	5 (864)	Observacionales: Cohortes retrospectiva y casos y controles	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	77 (75 a 80)	154 (149 a 159)	385 (373 a 398)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								23 (20 a 25)	46 (41 a 51)	115 (102 a 127)		
Verdaderos negativos	5 (864)	Observacionales: Cohortes	Muy serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	840 (797 a 868)	746 (709 a 772)	467 (443 a 483)	⊕○○○	CRÍTICO



La salud
es de todos

Minsalud

Falsos positivos		retrospectiva y casos y controles						60 (32 a 103)	54 (28 a 91)	33 (17 a 57)	MUY BAJA	
------------------	--	-----------------------------------	--	--	--	--	--	---------------	--------------	--------------	----------	--

Intervención: IgG (suero, sangre, plasma); Comparador: RT-PCR

Sensibilidad agrupada: 0.694 (IC 95%: 0.666-0.721) // Especificidad agrupada: 0.694 (IC 95%: 0.666-0.721)

Verdaderos positivos	5 (864)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	69 (67 a 72)	139 (133 a 144)	347 (333 a 361)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								31 (28 a 33)	61 (56 a 67)	153 (139 a 167)		
Verdaderos negativos	5 (864)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	625 (599 a 649)	555 (533 a 577)	347 (333 a 361)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								275 (251 a 301)	245 (223 a 267)	153 (139 a 167)		

Intervención: IgG/IgM – dual (suero, sangre, plasma); Comparador: RT-PCR

Sensibilidad agrupada: 0.845 (IC 95%: 0.822-0.866) // Especificidad agrupada: 0.916 (IC 95%: 0.860-0.954)

Verdaderos positivos	4 (841)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	85 (82 a 87)	169 (164 a 173)	423 (411 a 433)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								15 (13 a 18)	31 (27 a 36)	77 (67 a 89)		



Verdaderos negativos	5 (841)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	824 (774 a 859)	733 (688 a 763)	458 (430 a 477)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								76 (41 a 126)	67 (37 a 112)	42 (23 a 70)		

Explicaciones

- a. 26% de los estudios seleccionados no incluyeron descripción de los métodos de selección de los participantes. Se incluyó estudio de casos y controles y un 46% incluyeron pacientes previamente diagnosticados. La varolación del riesgo de sesgo no discrimina entre los estudios relacionados a pruebas serológicas como aquellos de pruebas moleculares, se incluyen estudios que evalúan ambos tipos de pruebas.
- b. Los estudios incluidos difieren en términos de tamaño, riesgo de sesgo y validez externa. Se encuentran altas tasas de heterogeneidad entre los estudios, probablemente debido a diferencias en los métodos, las características de los pacientes y las muestras utilizadas. El estándar de referencia es variable y no es detallada la diferencia entre RT-PCR como prueba índice vs como prueba de referencia para la definición de los casos.
- c. No se presenta evaluación de sesgo de publicación. No se identifican factores de confusión por limitada información de los estudios primarios.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Deeks et al. 2020. Antibody tests for identification of current and past infection with SARS-CoV-2 (Review)(18).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM (día 8 -14); Comparador: RT PCR

Sensibilidad agrupada: 0.722 (IC 95% 0.635 - 0.795) // Especificidad agrupada: 0.98 (IC 95% 0.972 - 0.994)

Verdaderos positivos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	72 (64 a 80)	144 (127 a 159)	361 (318 a 398)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								28 (20 a 36)	56 (41 a 73)	139 (102 a 182)		
Verdaderos negativos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	882 (875 a 895)	784 (778 a 795)	490 (486 a 497)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (5 a 25)	16 (5 a 22)	10 (3 a 14)		



Intervención: IgG/IgM (día 15 -21); Comparador: RT PCR

Sensibilidad agrupada: 0.914 (IC 95% 0.87 - 0.944) // Especificidad agrupada: 0.98 (IC 95% 0.972 - 0.994)

Verdaderos positivos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	91 (87 a 94)	183 (174 a 189)	457 (435 a 472)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								9 (6 a 13)	17 (11 a 26)	43 (28 a 65)		
Verdaderos negativos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	882 (875 a 895)	784 (778 a 795)	490 (486 a 497)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (5 a 25)	16 (5 a 22)	10 (3 a 14)		

Intervención: IgG/IgM (día 22 -35); Comparador: RT PCR

Sensibilidad agrupada: 0.96 (IC 95% 0.906 - 0.983) // Especificidad agrupada: 0.98 (IC 95% 0.972 - 0.994)

Verdaderos positivos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	96 (91 a 98)	192 (181 a 197)	480 (453 a 492)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								4 (2 a 9)	8 (3 a 19)	20 (8 a 47)		



Verdaderos negativos	9	Observacionales: cohortes	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	882 (875 a 895)	784 (778 a 795)	490 (486 a 497)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (5 a 25)	16 (5 a 22)	10 (3 a 14)		

Explicaciones

- a. La valoración del riesgo de sesgo informó el uso de diseños de múltiples grupos, la inclusión de solo casos de COVID-19, la falta de cegamiento de la prueba índice y el estándar de referencia, verificación diferencial, y la falta de claridad sobre el número de participantes, las características y las exclusiones de los estudios. La mayoría de los estudios solo incluyeron personas hospitalizadas debido a una infección por COVID-19 presunta o confirmada. No hubo estudios exclusivamente en participantes asintomáticos.
- b. Al examinar los forest plot se observa amplia variabilidad de los estudios primarios en cuanto a las estimaciones de sensibilidad y especificidad de las pruebas serológicas.
- c. Evidencia directa de reporte selectivo, no se informaron marcas de pruebas serológicas basadas en LFIA.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Guo et al. 2020. Seropositivity rate and diagnostic accuracy of serological tests in 2019-nCoV cases: a pooled analysis of individual studies (22).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgM; Comparador: RT-PCR

Sensibilidad agrupada: 0.71 (IC 95%: 0.69-0.73) // Especificidad agrupada: 0.97 (IC 95%: 0.96-0.97)

Verdaderos positivos	16 (3831)	Observacionales: No definen tipo	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	71 (69 a 73)	142 (138 a 146)	355 (345 a 365)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos	No especifican según prueba índice aplicada							29 (27 a 31)	58 (54 a 62)	145 (135 a 155)		
Verdaderos negativos	16 (3831)	Observacionales: No definen tipo	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	873 (864 a 873)	776 (768 a 776)	485 (480 a 485)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos	No especifican según prueba índice aplicada							27 (27 a 36)	24 (24 a 32)	15 (15 a 20)		

Intervención: IgG; Comparador: RT-PCR



Sensibilidad agrupada: 0.76 (IC 95%: 0.75-0.78) // Especificidad agrupada: 0.97 (IC 95%: 0.96-0.97)

Verdaderos positivos	16 (3831)	Observacionales: No definen tipo	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	76 (75 a 78)	152 (150 a 156)	380 (375 a 390)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos	No especifican según prueba índice aplicada							24 (22 a 25)	48 (44 a 50)	120 (110 a 125)		
Verdaderos negativos	16 (3831)	Observacionales: No definen tipo	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	873 (864 a 873)	776 (768 a 776)	485 (480 a 485)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos	No especifican según prueba índice aplicada							27 (27 a 36)	24 (24 a 32)	15 (15 a 20)		

Intervención: IgG/IgM – dual; Comparador: RT-PCR

Sensibilidad agrupada: 0.82 (IC 95%: 0.81-0.84) // Especificidad agrupada: 0.97 (IC 95%: 0.96-0.97)

Verdaderos positivos	16 (3831)	Observacionales: Cohortes retrospectiva y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	82 (81 a 84)	164 (162 a 168)	410 (405 a 420)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos	No especifican según prueba índice aplicada							18 (16 a 19)	36 (32 a 38)	90 (80 a 95)		
Verdaderos negativos	16 (3831)	Observacionales: Cohortes	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	873 (864 a 873)	776 (768 a 776)	485 (480 a 485)	⊕○○○	CRÍTICO



Falsos positivos	No especifican según prueba índice aplicada	retrospectiva y casos y controles						27 (27 a 36)	24 (24 a 32)	15 (15 a 20)	MUY BAJA	
------------------	---	-----------------------------------	--	--	--	--	--	--------------	--------------	--------------	----------	--

Explicaciones

- a. A pesar de que se reporta que la mayoría de los estudios son de aceptable calidad metodológica, la revisión sistemática no ofrece una descripción detallada de cada dominio de la herramienta QUADAS-2 que aplicaron para cada estudio primario. No se especifica el tipo de diseño de los estudios incluidos.
- b. Variabilidad de las estimaciones individuales de sensibilidad y especificidad entre los estudios primarios
- c. Aunque estadísticamente se reporte ausencia de sesgo de publicación para cada prueba índice, la búsqueda estuvo limitada a artículos en idioma inglés lo que hace sospechar la presencia de este sesgo.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Jarrom et al. 2020. Effectiveness of tests to detect the presence of SARS-CoV-2 virus, and antibodies to SARS-CoV-2, to inform COVID-19 diagnosis: a rapid systematic review (20).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: Pruebas de anticuerpos, incluye: IgM, IgG, IgG/IgM. (Por técnicas de ELISA; CLIA; LFIA; Oro coloidal); Comparador: RT PCR

Rango de sensibilidad: 0.184 – 0.961 // Rango de especificidad: 0.889 – 1.0

Verdaderos positivos	12 (1207)	Observacionales: cohortes, estudios prospectivos	Serio ^b	Serio ^b	No es serio	Serio ^b	Sospecha fuerte ^c	18 a 96	37 a 192	92 a 481	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								4 a 82	8 a 163	19 a 408		
Verdaderos negativos	12 (1207)		Serio ^b	Serio ^b	No es serio	Serio ^b	Sospecha fuerte ^c	798 a 900	710 a 800	444 a 500	⊕○○○	CRÍTICO



Falsos positivos		Observacionales: cohortes, estudios prospectivos						0 a 102	0 a 90	0 a 56	MUY BAJA	
------------------	--	--	--	--	--	--	--	---------	--------	--------	----------	--

Explicaciones

- a. Dentro de la evaluación QUADAS 2, el método de selección de pacientes se consideró poco claro en el 36% de los estudios y alto en otro 52%. Hubo un riesgo de sesgo alto y poco claro con respecto a cómo se realizó o interpretó la prueba índice en el 72% y el 12% de los estudios, respectivamente.
- b. Variabilidad en cuanto a la población, diseño de los estudios y características de la prueba índice. Rango amplio de los valores de sensibilidad.
- c. Búsqueda limitada por idioma.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnostica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Kontou et al. 2020. Antibody Tests in Detecting SARS-CoV-2 Infection: A Meta-Analysis (23).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM (ELISA); Comparador: RT-PCR

Sensibilidad agrupada: 0.935 (IC 95%: 0.900 - 0.971) // Especificidad agrupada: 0.987 (IC 95%: 0.973 - 1.000)

Verdaderos positivos	5 (1244)	Observacionales: casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No detectado	94 (90 a 97)	187 (180 a 194)	468 (450 a 486)	⊕⊕○○ BAJA	CRÍTICO
Falsos negativos								6 (3 a 10)	13 (6 a 20)	32 (14 a 50)		
Verdaderos negativos	5 (1244)	Observacionales: casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No detectado	888 (876 a 900)	790 (778 a 800)	494 (487 a 500)	⊕⊕○○	CRÍTICO



Falsos positivos								12 (0 a 24)	10 (0 a 22)	6 (0 a 13)	BAJA	
------------------	--	--	--	--	--	--	--	-------------	-------------	------------	------	--

Intervención: IgG/IgM (LFIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.828 (IC 95%: 0.770 - 0.886) // Especificidad agrupada: 0.994 (IC 95%: 0.984 - 0.998)

Verdaderos positivos	2 (1824)	Observacionales: casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No detectado	83 (77 a 89)	166 (154 a 177)	414 (385 a 443)	⊕⊕○○ BAJA	CRÍTICO
Falsos negativos								17 (11 a 23)	34 (23 a 46)	86 (57 a 115)		
Verdaderos negativos	2 (1824)	Observacionales: casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No detectado	895 (886 a 898)	795 (787 a 798)	497 (492 a 499)	⊕⊕○○ BAJA	CRÍTICO
Falsos positivos								5 (2 a 14)	5 (2 a 13)	3 (1 a 8)		

Intervención: IgG/IgM (CLIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.907 (IC 95%: 0.753 – 1.0) // Especificidad agrupada: 0.981 (IC 95%: 0.944 – 1.0)

Verdaderos positivos	2 (790)	Observacionales: casos y controles	Serio ^a	No es serio	No es serio	Serio ^b	No detectado	91 (75 a 100)	181 (151 a 200)	454 (377 a 500)	⊕⊕○○	CRÍTICO
----------------------	---------	---------------------------------------	--------------------	-------------	-------------	--------------------	--------------	---------------	-----------------	-----------------	------	---------



Falsos negativos								9 (0 a 25)	19 (0 a 49)	46 (0 a 123)	BAJA	
Verdaderos negativos	2 (790)	Observacionales: casos y controles	Serio ^a	No es serio	No es serio	Serio ^b	No detectado	883 (850 a 900)	785 (755 a 800)	491 (472 a 500)	⊕⊕○○ BAJA	CRÍTICO
Falsos positivos								17 (0 a 50)	15 (0 a 45)	9 (0 a 28)		

Explicaciones

- a. La mayoría de los estudios fueron clasificados como bajo riesgo de sesgo según la revisión sistemática a través de herramienta QUADAS-2, sin embargo, aunque no se precise el diseño específicamente de los estudios primarios, da a entender que son estudios de casos y controles.
- b. Variabilidad de las estimaciones de sensibilidad y especificidad individuales para detección de anticuerpos por método de ELISA y LFAI. Para CLIA no se presentó dicha variabilidad, sin embargo, se presenta un amplio intervalo de confianza para la estimación de sensibilidad.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Lisboa et al. 2020. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis (24).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM (ELISA); Comparador: RT-PCR o carga viral

Sensibilidad agrupada: 0.843 (IC 95%: 0.756 - 0.909) // Especificidad agrupada: 0.976 (IC 95%: 0.932 – 0.994)

Verdaderos positivos	6 (1875)	Observacionales: casos y controles	Muy serio ^a	No es serio	No es serio	No es serio	No evaluado	84 (76 a 91)	169 (151 a 182)	422 (378 a 455)	⊕⊕○○ BAJA	CRÍTICO
Falsos negativos								16 (9 a 24)	31 (18 a 49)	78 (45 a 122)		
Verdaderos negativos	6 (1875)	Observacionales: casos y controles	Muy serio ^a	No es serio	No es serio	No es serio	No evaluado	878 (839 a 895)	781 (746 a 795)	488 (466 a 497)	⊕⊕○○	CRÍTICO



Falsos positivos								22 (5 a 61)	19 (5 a 54)	12 (3 a 34)	BAJA	
------------------	--	--	--	--	--	--	--	-------------	-------------	-------------	------	--

Intervención: IgG/IgM (LFIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.66 (IC 95%: 0.493 - 0.793) // Especificidad agrupada: 0.966 (IC 95%: 0.943 - 0.982)

Verdaderos positivos	11 (5543)	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No evaluado	66 (49 a 79)	132 (99 a 159)	330 (247 a 397)	⊕⊕○○ BAJA	CRÍTICO
Falsos negativos								34 (21 a 51)	68 (41 a 101)	170 (103 a 253)		
Verdaderos negativos	11 (5543)	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	No es serio	No evaluado	869 (849 a 884)	773 (754 a 786)	483 (472 a 491)	⊕⊕○○ BAJA	CRÍTICO
Falsos positivos								31 (16 a 51)	27 (14 a 46)	17 (9 a 28)		

Intervención: IgG/IgM (CLIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.978 (IC 95%: 0.46.2– 1.0) // Especificidad agrupada: No estimable

Verdaderos positivos	2 (858)		Serio ^a	Serio ^b	No es serio	Serio	No evaluado	98 (46 a 100)	196 (92 a 200)	489 (231 a 500)	⊕○○○	CRÍTICO
----------------------	---------	--	--------------------	--------------------	-------------	-------	-------------	---------------	----------------	-----------------	------	---------



Falsos negativos		Observacionales: Cohortes y casos y controles						2 (0 a 54)	4 (0 a 108)	11 (0 a 269)	MUY BAJA BAJA	
------------------	--	--	--	--	--	--	--	------------	-------------	--------------	------------------	--

Explicaciones

- a. Para el dominio de selección de pacientes, se observó un riesgo alto o poco claro de sesgo en el 98% de las evaluaciones de QUADAS-2, en su mayoría relacionadas con un diseño de casos y controles (para ELISA todos fueron de este tipo de diseño) y sin utilizar un muestreo consecutivo o aleatorio. Para prueba índice el riesgo fue alto o poco claro porque no estaba claro si la prueba serológica se interpretó ciega al estándar de referencia o si los valores de corte para clasificar los resultados como positivo, negativo o indeterminado fueron preespecificados. También existió riesgo poco claro referente al estándar de referencia, y preocupaciones sobre aplicabilidad de prueba índice.
- b. Se observa variabilidad del reporte de sensibilidad con algunas estimaciones que tienen intervalos de confianza amplios, especialmente para LFIA y CLIA.
- c. La estimación de sensibilidad en de IgG/IgM por CLIA, presenta un IC bastante amplio, sólo fueron incluidos dos estudios.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Mekonne et al. 2020. Diagnostic accuracy of serological tests and kinetics of severe acute respiratory syndrome coronavirus 2 antibody: A systematic review and meta-analysis (25).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM (ELISA); Comparador: RT-PCR

Sensibilidad agrupada: 0.85 (IC 95%: 0.77 - 0.93) // Especificidad agrupada: 1.0 (IC 95%: 0.92 – 1.0)

Verdaderos positivos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	85 (77 a 93)	170 (154 a 186)	425 (385 a 465)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								15 (7 a 23)	30 (14 a 46)	75 (35 a 115)		
Verdaderos negativos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	900 (828 a 900)	800 (736 a 800)	500 (460 a 500)	⊕○○○	CRÍTICO



Falsos positivos								0 (0 a 72)	0 (0 a 64)	0 (0 a 40)	MUY BAJA	
------------------	--	--	--	--	--	--	--	------------	------------	------------	----------	--

Intervención: IgG/IgM (LFIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.77 (IC 95%: 0.69 - 0.84) // Especificidad agrupada: 0.98 (IC 95%: 0.97 - 0.99)

Verdaderos positivos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	77 (69 a 84)	154 (138 a 168)	385 (345 a 420)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								23 (16 a 31)	46 (32 a 62)	115 (80 a 155)		
Verdaderos negativos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	882 (873 a 891)	784 (776 a 792)	490 (485 a 495)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (9 a 27)	16 (8 a 24)	10 (5 a 15)		

Intervención: IgG/IgM (CLIA); Comparador: RT-PCR

Sensibilidad agrupada: 0.92 (IC 95%: 0.85 – 0.96 // Especificidad agrupada: 0.98 (IC 95%: 0.78 – 0.999)

Verdaderos positivos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	92 (85 a 96)	184 (170 a 192)	460 (425 a 480)	⊕○○○	CRÍTICO
----------------------	-------------	-------------	--------------------	--------------------	-------------	-------------	------------------------------	--------------	-----------------	-----------------	------	---------



Falsos negativos								8 (4 a 15)	16 (8 a 30)	40 (20 a 75)	MUY BAJA BAJA	
Verdaderos negativos	No reportan	No reportan	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	882 (702 a 899)	784 (624 a 799)	490 (390 a 500)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (1 a 198)	16 (1 a 176)	10 (0 a 110)		

Explicaciones

- a. La evaluación de la calidad metodológica indicó que la mayoría de los artículos no proporcionaron información clara sobre el cegamiento al leer las pruebas índice y estándar. Además, no se proporcionó información sobre la estrategia de muestreo (aleatoria, consecutiva o de casos y controles). No reportan diseño metodológico de los estudios primarios
- b. No es claro establecer características de la población, no se definen fuentes de heterogeneidad.
- c. No se presenta evaluación de sesgo de publicación. Búsqueda limitada a idioma inglés, no especifican búsqueda manual por referencias.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Wang et al. 2020. Meta-analysis of diagnostic performance of serology tests for COVID-19: impact of assay design and post-symptom-onset intervals (28).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM – anticuerpos totales; Comparador: NAATs

Sensibilidad agrupada: 0.78 (IC 95%: 0.7 - 0.85) // Especificidad agrupada: 0.97 (IC 95%: 0.93 – 0.99)

Verdaderos positivos	No reportan	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	78 (70 a 85)	156 (140 a 170)	390 (350 a 425)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								22 (15 a 30)	44 (30 a 60)	110 (75 a 150)		
Verdaderos negativos	No reportan		Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	873 (837 a 891)	776 (744 a 792)	485 (465 a 495)	⊕○○○	CRÍTICO



Falsos positivos		Observacionales: Cohortes y casos y controles						27 (9 a 63)	24 (8 a 56)	15 (5 a 35)	MUY BAJA	
------------------	--	--	--	--	--	--	--	-------------	-------------	-------------	----------	--

Intervención: IgM; Comparador: NAATs

Sensibilidad agrupada: 0.69 (IC 95%: 0.59 - 0.78) // Especificidad agrupada: 0.95 (IC 95%: 0.91 - 0.98)

Verdaderos positivos	No reportan	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	69 (59 a 78)	138 (118 a 156)	345 (295 a 390)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								31 (22 a 41)	62 (44 a 82)	155 (110 a 205)		
Verdaderos negativos	No reportan	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	Serio	Sospecha fuerte ^c	855 (819 a 882)	760 (728 a 784)	475 (455 a 490)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								45 (18 a 81)	40 (16 a 72)	25 (10 a 45)		

Intervención: IgG; Comparador: NAATs

Sensibilidad agrupada: 0.76 (IC 95%: 0.65 - 0.86) // Especificidad agrupada: 0.98 (IC 95%: 0.96 - 0.99)

Verdaderos positivos	No reportan	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	Serio ^b	Sospecha fuerte ^c	76 (65 a 86)	152 (130 a 172)	380 (325 a 430)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								24 (14 a 35)	48 (28 a 70)	120 (70 a 175)		



Verdaderos negativos	No reportan	Observacionales: Cohortes y casos y controles	Serio ^a	Serio ^b	No es serio	Serio ^b	Sospecha fuerte ^c	882 (864 a 891)	784 (768 a 792)	490 (480 a 495)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								18 (9 a 36)	16 (8 a 32)	10 (5 a 20)		

Explicaciones

- a. Alto riesgo de sesgo para selección de participantes en más del 75% de los estudios.
- b. La especificidad podría estar sobreestimada porque la mayoría de los grupos control utilizó muestras de donantes sanos antes de 2020, lo que evita posibles reactividades cruzadas. La inscripción de los participantes se limitó a un período breve y sin un criterio de exclusión claro.
- c. Búsqueda limitada por idioma. Sin evaluación de sesgo de publicación.



Autor(es): KC

Pregunta: ¿Cuál es la exactitud diagnóstica de las pruebas basadas en anticuerpos que detectan la presencia del SARS-CoV-2 para informar el diagnóstico de COVID-19?

Bibliografía: Zhang et al, 2020. Diagnostic efficacy of anti-SARS-CoV-2 IgG/IgM test for COVID-19: A meta-analysis (29).

Evaluación de la certeza								Efecto por 1000 pacientes analizados			Certeza	Importancia
Resultado	No. de estudios (N° participantes)	Diseño de estudio	Riesgo de sesgo	Inconsistencia	Evidencia indirecta	Imprecisión	Sesgo de publicación	Probabilidad pre-test de 10%	probabilidad pre-test de 20%	probabilidad pre-test de 50%		

Intervención: IgG/IgM (ELISA); Comparador: RT PCR

Sensibilidad agrupada: 0.69 (IC 95%: 0.50 - 0.85) // Especificidad agrupada: 1.00 (IC 95%: 1.00 - 1.00)

Verdaderos positivos	No reportan	Observacionales: Estudios retrospectivos	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	69 (50 a 85)	138 (100 a 170)	345 (250 a 425)	⊕○○○ MUY BAJA	CRÍTICO
Falsos negativos								31 (15 a 50)	62 (30 a 100)	155 (75 a 250)		
Verdaderos negativos	No reportan	Observacionales: Estudios retrospectivos	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	900 (900 a 900)	800 (800 a 800)	500 (500 a 500)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								0 (0 a 0)	0 (0 a 0)	0 (0 a 0)		



Intervención: IgG/IgM (CLIA); Comparador: RT PCR

Sensibilidad agrupada: 0.96 (IC 95%: 0.91 - 0.98) // Especificidad agrupada: 1.00 (IC 95%: 1.00 - 1.00)

Verdaderos positivos	No reportan	Observacionales: Estudios retrospectivos	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	96 (91 a 98)	192 (182 a 196)	480 (455 a 490)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								4 (2 a 9)	8 (4 a 18)	20 (10 a 45)		
Verdaderos negativos	No reportan	Observacionales: Estudios retrospectivos	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	900 (900 a 900)	800 (800 a 800)	500 (500 a 500)	⊕○○○ MUY BAJA	CRÍTICO
Falsos positivos								0 (0 a 0)	0 (0 a 0)	0 (0 a 0)		

Intervención: IgG/IgM (GICA); Comparador: RT PCR

Sensibilidad agrupada: 0.84 (IC 95% 0.78 - 0.90) // Especificidad agrupada: 0.95 (IC 95% 0.93 - 0.98)

Verdaderos positivos	No reportan	Observacionales: Estudios retrospectivos y prospectivo	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	84 (78 a 90)	168 (156 a 180)	420 (390 a 450)	⊕○○○ MUY BAJA BAJA	CRÍTICO
Falsos negativos								16 (10 a 22)	32 (20 a 44)	80 (50 a 110)		
Verdaderos negativos	No reportan	Observacionales: Estudios	Serio ^a	Serio ^b	No es serio	No es serio	Sospecha fuerte ^c	855 (837 a 882)	760 (744 a 784)	475 (465 a 490)	⊕○○○	CRÍTICO



Falsos positivos		retrospectivos y prospectivo						45 (18 a 63)	40 (16 a 56)	25 (10 a 35)	MUY BAJA	
------------------	--	------------------------------	--	--	--	--	--	--------------	--------------	--------------	----------	--

Explicaciones

- a. El 95,45% (21 de 22) para la selección de pacientes, el 36,36% (8 de 22) para la prueba índice, el 13,64% (3 de 22) para el flujo y el tiempo, y el 4,55% (1 de 22) para el estándar de referencia mostraron un alto riesgo de sesgo. No se evitaron estudios de casos y controles y exclusiones inapropiadas. No se incluyeron pacientes "difíciles de diagnosticar" lo que reduce el espectro de acción de las pruebas.
- b. Se encontró heterogeneidad estadística. En forest plot se observa variabilidad de las estimaciones, sin solapamiento de los IC.
- c. A pesar de no encontrar sesgo de publicación por test de Deeks, las búsquedas son limitadas por idioma.



La salud
es de todos

Minsalud



MINSALUD



www.minsalud.gov.co



Carrera 13 No. 32-76, piso 1
Bogotá, D.C., Colombia



@MinSaludCol



Instituto de Evaluación
Tecnológica en Salud®



www.iets.org.co



Carrera 49 a No. 91 - 91
Bogotá, D.C., Colombia



(+571) 3770100



contacto@iets.org.co



@ietscolombia



[ietscolombia](https://www.instagram.com/ietscolombia)



Análisis parcial de productos Componente 1 del Contrato 9677-MECOV19-1009-2020, siguiendo lo establecido en el párrafo primero de la cláusula primera de dicho contrato, mediante la herramienta AMSTAR para revisiones sistemáticas.

¿La pregunta de investigación y los criterios de inclusión de la revisión incluyeron los componentes de la estructura PICO o de otra estructura específica según el objetivo?

Si. La estructura PICO presentada es acorde con la pregunta de investigación y el objetivo.

¿Contiene un apartado indicando que los métodos de revisión fueron establecidos antes de la realización de la misma?

No. Indicar en una parte de los métodos esto.

¿Se sustentan los diseños seleccionados para incluir en la revisión?

No. Se sugiere sustentar brevemente por qué solo se utilizaron revisiones sistemáticas.

¿Se utiliza una estrategia de búsqueda exhaustiva?

Si. Se incluyen las bases recomendadas para revisiones sistemáticas, el uso de términos controlados y libres y una descripción detallada de las estrategias utilizadas en el Anexo 1.

¿La selección de los estudios se realiza por duplicado?

Si. Se sugiere incluir si se realizó un piloto para este paso.

¿Se realiza la extracción por duplicado?

Si. La revisión involucró dos investigadores de con revisión independiente de la información consignada en la herramienta de extracción. Se sugiere incluir si se realizó un piloto para este paso.

¿Se presenta un listado de estudios excluidos y se justifica la razón?

Si. en el Anexo 4.

¿Se describen los estudios incluidos en detalle adecuado?

Si. Mediante el Anexo 3.

¿Se utiliza una técnica adecuada de evaluación de riesgo de sesgos en los estudios incluidos?

Si. Se utilizan diferentes herramientas disponibles según el tipo de estudio.

¿Se reportan las fuentes de financiamiento de los estudios incluidos?

No. Pero dado el tipo de documentación recolectada al no ser estudios primarios este apartado es opcional.

¿Se considera la evaluación de riesgo de sesgos de estudios individuales al interpretar o discutir los resultados de la revisión?

Si. Se consideran las particularidades metodológicas de los documentos al presentar los resultados.



¿Se presenta una explicación y se discute la heterogeneidad observada en los resultados?

Si. Se tiene en consideración las diferencias observadas en los resultados y los estudios en la discusión.

Si se realiza una síntesis cuantitativa, ¿Se lleva a cabo una adecuada indagación de los sesgos de publicación, y se discute su probable impacto en los resultados de la revisión?

No aplica. Se realiza una síntesis narrativa de los resultados.

¿Se menciona la fuente de financiamiento y fuentes de conflicto de interés para realizar la revisión?

Si, los autores informan no tener conflicto de interés y su independencia editorial, y explícitamente se menciona que esta revisión es producto del Contrato No. 9677-MECOV19-1009-2020 por solicitud del Fondo de gestión del riesgo de desastres.