



FACULTAD DE FARMACIA

Grado en Farmacia

Evaluación de estrategias de diagnóstico para la identificación de infección de SARS-CoV-2 en la práctica clínica: una Revisión Sistemática y Cumplimiento de las Normas de Publicación de Estudios de Diagnóstico de acuerdo a las normas STARD.

Memoria de Trabajo Fin de Grado

Sant Joan d'Alacant

Febrero 2023

Autor: Paula Cremades Martínez
Modalidad: Revisión sistemática
Tutor/es: Blanca Juana Lumbreras Lacarra

ÍNDICE

1. RESUMEN	3
2. ANTECEDENTES	5
3. OBJETIVOS	8
4. MATERIAL Y MÉTODOS	9
5. RESULTADOS	12
6. DISCUSIÓN	17
7. CONCLUSIONES.....	21
8. BIBLIOGRAFÍA	37



1. RESUMEN

Objetivo: Nuestro objetivo fue revisar las estrategias para identificar la infección por SARS-CoV-2 antes de la disponibilidad de los resultados de las pruebas moleculares y evaluar la calidad de los informes de los estudios identificados mediante la aplicación de la guía STARD.

Material y métodos: Revisamos 3.821 artículos publicados desde el 1 de enero de 2020 hasta el 30 de abril de 2022, de los cuales 23 (0,6%) cumplieron con los criterios de inclusión: incluir al menos dos pruebas/criterios diagnósticos, estar diseñados para su uso en la práctica clínica o en un contexto de salud pública y proporcionar tasas de precisión diagnóstica. Dos investigadores realizamos de forma independiente la extracción de datos y la aplicación de los criterios STARD y las discrepancias se discutieron con un tercer autor.

Resultados: La mayoría de los estudios (16, 69,6 %) incluyeron pacientes sintomáticos con sospecha de infección, seis estudios (26,1 %) incluyeron pacientes ya diagnosticados y un estudio (4,3 %) incluyó personas con contacto estrecho con un paciente positivo para COVID. Las principales pruebas incluidas en los estudios; que incluyeron pacientes sintomáticos para mejorar el diagnóstico; fueron las pruebas de imagen y demográficas, los síntomas y el recuento de linfocitos. Los valores del área bajo la curva (AUC), oscilaron entre 53 y 97,4. Siete estudios (30,4%) validaron el modelo diagnóstico en una muestra independiente. El promedio de criterios STARD cumplidos fue de 17,6 (máximo 27 y mínimo 5).

Conclusiones: Se muestran valores altos de precisión diagnóstica cuando se considera más de una prueba diagnóstica, principalmente pruebas de imagen y demográficas, síntomas y recuento de linfocitos. Esto podría ofrecer el potencial para identificar a las personas con infección por SARS-CoV-2 con alta precisión cuando no se dispone de pruebas moleculares. Sin embargo, es urgente la validación externa de los modelos desarrollados y las evaluaciones en poblaciones lo más similares posible a aquellas en las que se aplicarán.

ABSTRACT: We aimed to review strategies for identifying SARS-CoV-2 infection before the availability of molecular test results, and to assess the reporting quality

of the studies identified through the application of the STARD guideline. We screened 3,821 articles published until 30 April 2021, of which 23 met the inclusion criteria: including at least two diagnostic variables, being designed for use in clinical practice or in a public health context and providing diagnostic accuracy rates. Data extraction and application of STARD criteria were performed independently by two researchers and discrepancies were discussed with a third author. Most of the studies (16, 69.6%) included symptomatic patients with suspected infection, six studies (26.1%) included patients already diagnosed and one study (4.3%) included individuals with close contact to a COVID-positive patient. The main variables considered in the studies, which included symptomatic patients, were imaging and demographic characteristics, symptoms, and lymphocyte count. The values for area under the receiver operating characteristic curve (AUC) ranged from 53-97.4. Seven studies (30.4%) validated the diagnostic model in an independent sample. The average number of STARD criteria fulfilled was 17.6 (maximum, 27 and minimum, 5). High diagnostic accuracy values are shown when more than one diagnostic variable is considered, mainly imaging and demographic characteristics, symptoms, and lymphocyte count. This could offer the potential to identify individuals with SARS-CoV-2 infection with high accuracy when molecular testing is not available. However, external validation for developed models and evaluations in populations as similar as possible to those in which they will be applied is urgently needed.

PALABRAS CLAVE: SARS-CoV-2, diagnóstico, STARD, COVID-19

2. ANTECEDENTES

Existe una amplia variedad de pruebas disponibles para detectar la infección actual o pasada por SARS-CoV-2, tanto para uso clínico como para ser aplicadas en salud pública. La mayoría de las pruebas para detectar la infección actual se han centrado en ensayos de la PCR cuantitativa de transcriptasa inversa (RT-qPCR). Sin embargo, estas pruebas tienen limitaciones, como el tiempo necesario para realizarlas y la presencia de resultados falsos negativos superior al 30%⁽¹⁾, particularmente cuando las personas a las que se les realizan se encuentran en las primeras etapas de infección y/o son asintomáticos. Dado que se ha demostrado que los pacientes asintomáticos representan hasta el 80% de los infectados⁽²⁾ y pueden transmitir el virus⁽³⁾, las pruebas deben proporcionar resultados rápidos con suficiente sensibilidad y especificidad para informar las medidas de prevención y aislamiento social. El uso de pruebas de diagnóstico rápido que detectan principalmente el antígeno del virus se ha implementado ampliamente en la práctica como una estrategia eficaz para proporcionar un resultado rápido en cuestión de minutos. Estas pruebas tienen problemas similares con la sensibilidad, y la limitación de que un resultado positivo debe confirmarse mediante la prueba RT-qPCR, lo que dificulta la implementación de estas estrategias diagnósticas en entornos donde estas pruebas moleculares no están disponibles⁽⁴⁾. Además, en algunas situaciones clínicas es necesario detectar la presencia de infección por SARS-CoV-2 de forma inmediata, como en pacientes con signos o síntomas de COVID-19, con el fin de iniciar el tratamiento o identificar casos asintomáticos e interrumpir la transmisión del SARS-CoV-2, cuando el riesgo comunitario o los niveles de transmisión son altos. Por lo tanto, es necesario desarrollar nuevas estrategias efectivas para el diagnóstico y la detección.

Un aspecto fundamental a considerar es que el rendimiento diagnóstico de una prueba está directamente relacionado con el contexto en el que se lleva a cabo, que debe ser tenido en cuenta cuando se interpretan los resultados, además de las características de la propia prueba⁽⁵⁾. Este contexto influirá en gran medida en la probabilidad de que la prueba sea positiva antes de realizarla (probabilidad

previa a la prueba). Por ejemplo, un resultado positivo de la prueba en personas que son contactos cercanos de casos confirmados o en personas con síntomas que sugieren COVID-19 puede dar un valor predictivo marcadamente diferente en comparación con una prueba de cribado realizada a individuos asintomáticos con el fin de realizar un viaje ⁽⁶⁾.

Una revisión sistemática anterior evaluó la precisión diagnóstica y la calidad metodológica de los estudios que evalúan las pruebas de diagnóstico rápido para la detección de antígenos en el punto de atención así como pruebas de base molecular ⁽⁷⁾. En esta revisión, los autores afirmaron que las evaluaciones prospectivas de los test rápidos para la infección por COVID-19 en entornos clínicamente relevantes se necesitaban con urgencia. Además, estos estudios no incluían suficiente información sobre variables relevantes como la sintomatología del paciente y el tiempo transcurrido desde la aparición del primer síntoma o la exposición, entre otros. Debido a esta falta de datos, no es posible saber si los resultados obtenidos pueden ser aplicados en los diferentes escenarios clínicos, es decir, en pacientes sintomáticos, levemente sintomáticos, asintomáticos, hospitalizados o para el uso en salud pública. Los autores también concluyeron que estos estudios deben cumplir las directrices de comunicación sobre los estudios de precisión de pruebas diagnósticas (STARD) para mejorar la valoración crítica y la reproducibilidad de los estudios ⁽⁸⁾. Al igual que en otros estudios clínicos, los estudios de precisión diagnóstica tienen riesgo de sesgo, principalmente debido a deficiencias metodológicas en la captación de participantes, la recopilación de datos, la ejecución o interpretación de la prueba o el análisis de los datos. Como resultado de estos sesgos, las estimaciones de la sensibilidad y la especificidad de la prueba en comparación con el estándar de referencia pueden no ser correctas, afectando de manera negativa al manejo de los pacientes y al sistema sanitario. La guía STARD se desarrolló en 2003 y se actualizó en 2015 para facilitar un informe más completo y transparente de los estudios de exactitud diagnóstica y de este modo reducir la probabilidad de sesgo. Esta guía contiene una lista de verificación de elementos que deben informarse en estos estudios.

Otra revisión más reciente que evaluó los estudios sobre la precisión diagnóstica de las pruebas basadas en RT-qPCR y otras pruebas de amplificación de ácidos nucleicos mostró que estos estudios tenían un alto riesgo de sesgo debido a un diseño del estudio incorrecto (según los objetivos). Esta revisión también destaca la necesidad de utilizar directrices estandarizadas para el desarrollo de estudios ⁽⁹⁾.

Estas revisiones anteriores se centran en la evaluación de las pruebas de diagnóstico molecular individuales aplicadas al diagnóstico o la detección de SARS-CoV-2. Dadas las limitaciones de las pruebas moleculares, generalmente se combinan varias pruebas clínicas para obtener un diagnóstico final. Además, dependiendo del entorno clínico, pueden existir diferentes procesos de diagnóstico. Por ejemplo, en la práctica clínica, para lograr una adecuada sensibilidad y especificidad diagnóstica, se recomienda una combinación de pruebas moleculares, para detectar la presencia del virus, y otras pruebas diagnósticas como las de imagen o las bioquímicas ⁽¹⁰⁾. Aunque de acuerdo con una revisión sistemática reciente ⁽¹¹⁾, la ausencia o presencia de signos y síntomas no es lo suficientemente precisa para descartar o sospechar la presencia de COVID-19, combinarlos con otras pruebas diagnósticas podría aumentar la precisión del diagnóstico.

Por lo tanto, es necesario conocer la evidencia disponible relacionada con cualquier estrategia de diagnóstico evaluada en la práctica antes de obtener un diagnóstico molecular para detectar mejor la infección por SARS-CoV-2. Dado el potencial de sesgos mostrado en anteriores revisiones sistemáticas sobre las pruebas moleculares destinadas a detectar el COVID-19, también es necesario evaluar su adhesión a la guía STARD.

3. OBJETIVOS

El objetivo del estudio fue evaluar la evidencia disponible sobre las diferentes estrategias de diagnóstico para identificar la infección por el SARS-CoV-2 antes de la disponibilidad de los resultados de las pruebas moleculares, así como la calidad de los informes de los estudios disponibles mediante la aplicación de la guía STARD.



4. MATERIAL Y MÉTODOS

La revisión sistemática se llevó a cabo de acuerdo con las recomendaciones de la declaración PRISMA (Preferred Reporting Items for Systematic Reviews and MetaAnalyses) ⁽¹²⁾.

Criterios de elegibilidad

Seleccionamos artículos de investigación originales publicados entre el 1 de enero de 2020 y el 28 de febrero de 2022. Se incluyeron estudios cuyo objetivo principal era evaluar una estrategia de diagnóstico (que incluyera al menos dos pruebas) para detectar la presencia de infección por el SARS-CoV-2 antes de obtener un diagnóstico molecular; la estrategia de diagnóstico estaba diseñada para su uso en la práctica clínica o en un contexto de salud pública, y los índices de precisión debían proporcionarse ya sea en términos de sensibilidad y especificidad, el cociente de probabilidad o la curva ROC. La búsqueda se limitó a artículos escritos en inglés o español. Se excluyeron los artículos centrados en la evaluación o validación de una sola prueba diagnóstica individual o la comparación entre diferentes pruebas diagnósticas individuales.

Estrategia de búsqueda

Se realizó una búsqueda exhaustiva y sensible basada en los términos de búsqueda desarrollados para el proyecto COVID-19 Open Access por investigadores y bibliotecarios del Instituto de Medicina Social y Preventiva de la Universidad de Berna (https://www.ispm.unibe.ch/about_us/news/new_searchable_living_evidence_in_covid_19_open_access_project/index_eng.html). La búsqueda de artículos incluyó PubMed, Embase y preprints indexados en las bases de datos bioRxiv y medRxiv.

Se aplicaron los siguientes términos de búsqueda para descubrir estudios de precisión diagnóstica: “Sensitivity and Specificity” [Mesh], “Sensitivity” or “Specificity,” “Area Under Curve” [Mesh], or “Area under curve,” “ROC Curve” [Mesh], or “ROC,” “Diagnostic Accuracy,” “Likelihood Ratio” and “Severe Acute

Respiratory Syndrome Coronavirus 2” [Supplementary Concept], “COVID- 19” [Supplementary Concept], “Coronavirus” or “Corona Virus,” “HCoV” or “nCoV,” “2019 CoV,” “Covid,” “Covid19,” “Severe Acute Respiratory Syndrome Coronavirus 2,” “SARS-CoV2,” “SARS-CoV 2,” or “SARS Coronavirus 2”.

Selección de estudios

Se incluyeron artículos originales sobre una estrategia de diagnóstico para detectar la presencia de infección por SARS-CoV-2, que incluyera la combinación de al menos dos pruebas de diagnóstico. Se realizó un cribado para excluir editoriales, cartas al editor y cualquier estudio diferente a un informe de investigación original.

Dos revisores examinamos de forma independiente el título de cada referencia y el resumen (si estaba disponible) para determinar su relevancia para la revisión. Los desacuerdos se resolvieron mediante discusión y consenso con un tercer revisor. A continuación, revisamos los artículos completos seleccionados en la primera ronda. En la segunda ronda, dos revisores aplicamos de forma independiente los criterios de selección. Se excluyeron los estudios en los que los autores no realizaron específicamente una evaluación de la precisión diagnóstica de un proceso de diagnóstico que implicara la combinación de más de una prueba.

Para la inclusión de los estudios, se estableció que la evaluación de la concordancia entre los autores (índice k) debía ser superior a 0,60.

Extracción de datos

De cada uno de los artículos incluidos se extrajeron las siguientes variables: año de publicación, país de origen, diseño del estudio, ámbito clínico, objetivo del estudio (diagnóstico o cribado), características de la población del estudio (sexo, edad, síntomas y tamaño de la muestra), pruebas incluidas en la estrategia diagnóstica, resultados de las pruebas diagnósticas y conclusiones de los autores. Esta extracción la realizamos de manera independiente dos

investigadores y luego las discrepancias se discutieron independientemente con un tercer autor.

Dos revisores aplicamos también de manera independiente los criterios de la guía STARD y las discrepancias se discutieron con un tercer autor. Los resultados de esta evaluación se resumen en el material suplementario.

También se analizó la concordancia mediante el índice kappa tanto de la extracción de las variables del estudio como de la aplicación de la guía STARD.



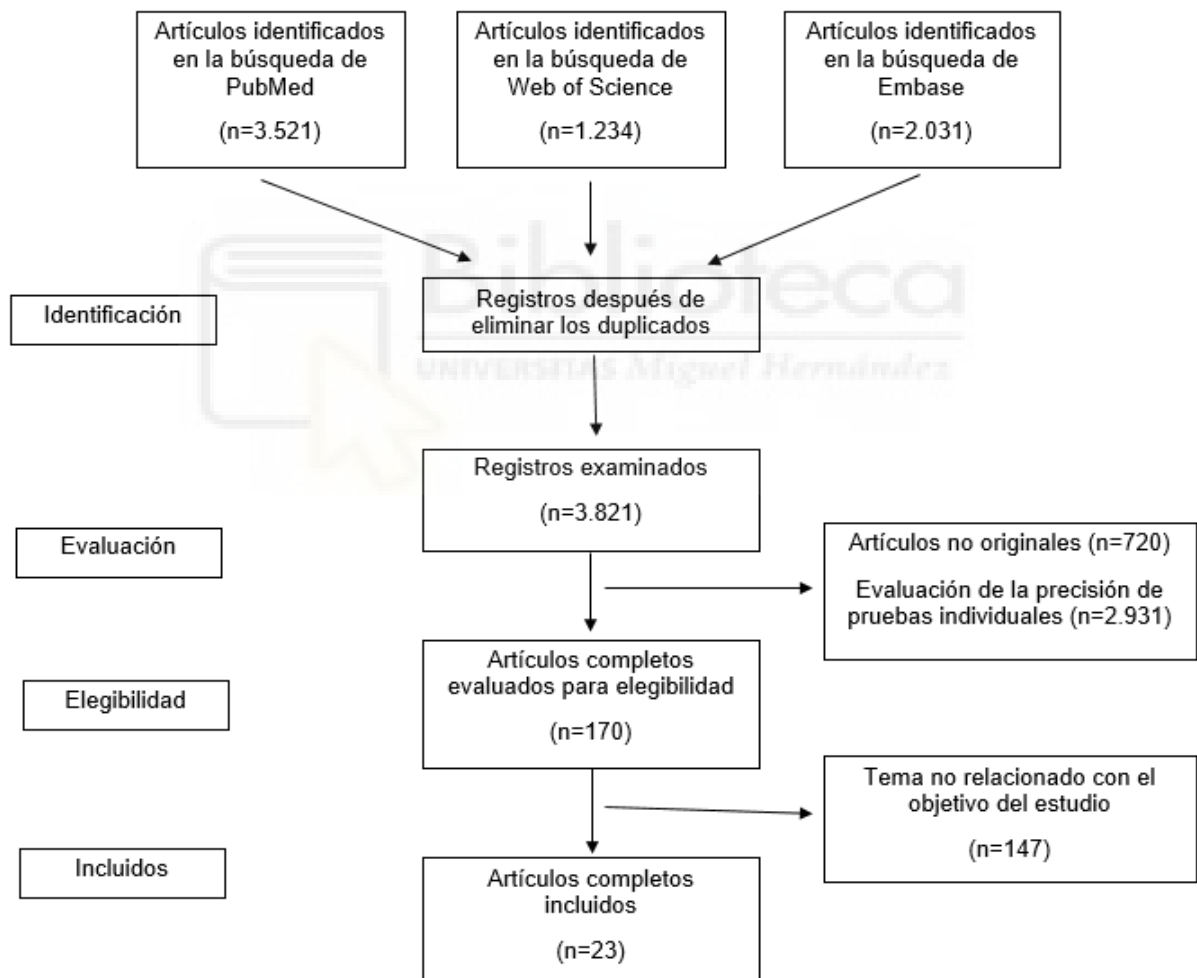
5. RESULTADOS

Este trabajo se encuentra publicado en la revista *Microbiology Spectrum* que corresponde al primer cuartil de su especialidad. El artículo completo se encuentra adjunto como anexo.

Resultados de la búsqueda

Se revisaron un total de 3.821 resúmenes. De ellos, se incluyeron 170 estudios para la revisión del texto completo, y 23⁽¹³⁻³⁵⁾ cumplieron los criterios de inclusión. (Fig. 1)

Fig. 1 Diagrama de flujo de la inclusión de artículos en la revisión sistemática según las recomendaciones PRISMA



Características de los artículos incluidos.

La tabla 1 muestra las características de los estudios incluidos, que se han clasificado según el tipo de población incluida:

- a) Pacientes con sospecha de infección por SARS-CoV-2 (16, 69'6%) ⁽¹³⁻²⁸⁾
- b) Pacientes diagnosticados con SARS-CoV-2 (6, 26'1%) ⁽²⁹⁻³⁴⁾
- c) Otros estudios (1, 4'3%) ⁽³⁵⁾

a) Estudios que incluían pacientes sintomáticos con sospecha de infección por el SARS-CoV-2

De los 16 estudios que se centraron en pacientes sintomáticos con sospecha de infección por SARS-CoV-2 ⁽¹³⁻²⁸⁾, siete incluyeron técnicas relacionadas con la inteligencia artificial (IA) ⁽¹³⁻¹⁹⁾: cinco de ellos utilizaron un estudio de cohortes, ya sea retrospectivo ^(13, 14, 18, 19) o prospectivo ⁽¹⁵⁾, y dos utilizaron un estudio transversal ^(16, 17). Los otros nueve estudios desarrollaron un algoritmo ^(20, 25, 27) o una puntuación de probabilidad ^(21-24, 26, 28) para diagnosticar infección por SARS-CoV-2. Cuatro de ellos utilizaron un estudio transversal ⁽²³⁻²⁶⁾, cuatro utilizaron un estudio de cohortes retrospectivo ^(20-22, 27), y uno utilizó un estudio de cohortes prospectivo ⁽²⁸⁾. La mayoría de los estudios incluyeron pacientes sintomáticos que ingresaron en un hospital; el estudio Trubiano JA et al. ⁽²⁵⁾ incluyó a pacientes evaluados para COVID-19 en una clínica de cribado debido a los síntomas (98,3%), el contacto con un paciente positivo en COVID-19 (17,3%), o aquellos que habían viajado al extranjero (24,6%), y Elimian KO et al. ⁽²⁷⁾ incluyeron a pacientes sintomáticos y también a los preocupados por la infección por COVID-19.

El tamaño de las muestras incluidas en estos estudios osciló entre 100 y 181.544 pacientes (90.772 en el conjunto de derivación y 90.772 en el conjunto de validación).

Las tablas 2 y 3 muestran que las principales variables incluidas en los modelos fueron el recuento de linfocitos (10/16 estudios), las características de imagen (RXC/TC) (9/16 estudios), las características demográficas (8/16 estudios) y los síntomas (7/16 estudios). Otras variables fueron menos incluidas en los modelos: pruebas rutinarias de laboratorio (4/16 estudios) y diferentes mediciones bioquímicas (LDH, 6/16 estudios; PCR, 6/16 estudios; y ferritina, 4/16 estudios). Además, uno de los estudios ⁽¹³⁾ incluyó en el modelo de diagnóstico el primer

resultado de la RT-qPCR. Los modelos mostraron diferentes niveles para el área bajo la curva de características operativas del receptor (AUC) (osciló entre 53 y 97,4).

Tres de los artículos que utilizaron IA ^(13, 16, 19) y cuatro de los estudios que desarrollaron una puntuación de algoritmo/probabilidad ^(20, 21, 27, 28) validaron el modelo desarrollado en una muestra independiente. Además, otros estudios validaron internamente el modelo mediante técnicas como la técnica bootstrap (un procedimiento estadístico que vuelve a muestrear un único conjunto de datos para crear muchas muestras simuladas) ^(22, 23, 25) o un proceso de validación cruzada k-fold (un enfoque que divide aleatoriamente el conjunto de observaciones en k grupos o pliegues, de tamaño aproximadamente igual; el primer pliegue se trata como un conjunto de validación, y el método se ajusta a los k-1 pliegues restantes) ^(14, 15, 18) (Tabla 1). De acuerdo con las conclusiones de los autores (Tabla 2), cinco de los nueve estudios que no llevaron a cabo una validación externa señalaron la importancia de una mayor validación antes de aplicar el modelo a la práctica clínica ⁽²²⁻²⁶⁾. Aljame M et al. ⁽¹⁷⁾ realizaron una validación de una puntuación anterior.

b) Estudios que incluyen a pacientes con diagnóstico de infección por SARS-CoV-2

De los 23 estudios seleccionados para la revisión, seis incluyeron pacientes con un diagnóstico confirmado por RT-qPCR en un entorno hospitalario ⁽²⁹⁻³⁴⁾. Tres de los estudios utilizaron un diseño transversal ^(30, 31, 33), uno de los estudios utilizó un diseño de cohorte retrospectivo ⁽²⁹⁾, un estudio utilizó un diseño de cohorte abierto ⁽³²⁾ y uno fue un estudio de casos y controles ⁽²³⁾. El tamaño de la muestra osciló entre 114 y 192.778 pacientes. Cuatro estudios utilizaron IA ^(30, 31, 33, 34); Planted TB et al. ⁽³³⁾ desarrollaron el modelo de un conjunto de entrenamiento y lo validaron en un conjunto independiente. Dos estudios validaron el modelo internamente mediante un proceso de validación cruzada k-fold ^(31, 32).

Cuatro de los seis estudios incluían análisis de sangre completo y mediciones bioquímicas (tablas 2 y 4). Sólo dos de los estudios incluyeron características demográficas. Las variables como las características de las pruebas de imagen (RXC/TC) sólo se incluyeron en el estudio de Gatti M et al. ⁽²³⁾. Los niveles de AUC fueron superiores a los de los estudios que incluyeron pacientes con sospecha de infección por SARS-CoV-2 (entre 84,2 y 100).

c) Otros estudios

Banerjee A et al. ⁽³⁵⁾ incluyeron pacientes a los que se les tomaron muestras para realizar la técnica RT-qPCR durante una visita al hospital. Sin embargo, no incluyeron datos relacionados con las características clínicas de los pacientes. Aplicaron el aprendizaje automático y el bosque aleatorio (un algoritmo de aprendizaje supervisado que crea aleatoriamente y fusiona múltiples árboles de decisión en un “bosque”) para identificar a los pacientes positivos al SARS-CoV-2 a partir de recuentos sanguíneos completos, y mostraron un AUC de 85. No obstante, los autores afirmaron que será necesario realizar más validaciones.

Descripción del cumplimiento de los criterios STARD

La media de criterios cumplidos fue de 17,6 (máximo, 27 y mínimo, 5) sobre 34 (Tabla 5 y material suplementario Tabla 1A y 1B).

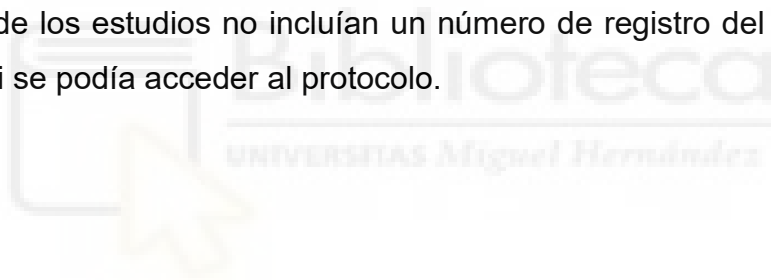
La mayoría de los artículos cumplían los dos criterios incluidos en la sección de introducción (91,3%) y detallaban los objetivos e hipótesis del estudio. Sin embargo, sólo el 52,2% de los artículos se identificó como un estudio de precisión diagnóstica, y el 69,6% incluía un resumen estructurado y los antecedentes científicos y clínicos que justificaban el estudio.

En la sección del método, menos del 60% de los estudios describieron los criterios de elegibilidad. Por ejemplo, sólo dos de los nueve estudios que incluían pacientes con sospecha de infección por SARS-CoV-2 incluyeron los días desde el inicio de los síntomas como variable, y la mitad de los estudios incluyeron una descripción de síntomas. Además, solo dos estudios incluidos en la revisión

describieron cómo los autores estimaron el tamaño de la muestra y sólo uno mencionó cómo se manejaron los resultados indeterminados de la prueba índice o del estándar de referencia. De los 23 estudios incluidos menos del 30% describió si la evaluación de la prueba índice y del estándar de referencia era ciega.

En el apartado de resultados, dadas las características de las pruebas evaluadas, ningún estudio incluyó la descripción de algún efecto adverso relacionado con la prueba índice o el estándar de referencia. Sólo cuatro (17,4%) estudios describieron el intervalo de tiempo entre la prueba índice y el estándar de referencia y cinco estudios (21,7%) incluyeron el ítem relacionado con la distribución de los diagnósticos alternativos en aquellos que no tienen la condición objetivo. Sólo el 86,9% de los estudios incluyeron la precisión del AUC.

La mayoría de los estudios no incluían un número de registro del estudio y no explicaban si se podía acceder al protocolo.



6. DISCUSION

La aparición de la pandemia COVID-19 supuso que los investigadores y las revistas hicieran un gran esfuerzo para generar pruebas científicas obteniendo resultados a menudo de forma acelerada. Esto puede haber influido en la calidad de los estudios publicados, aumentando el riesgo de sesgo, y a menudo ha proporcionado información incompleta. Asimismo, el escaso cumplimiento de las recomendaciones disponibles para la presentación de resultados en un artículo científico, como las directrices STARD (en esta revisión, el cumplimiento medio de los criterios STARD [sobre 34] fue de 17,6; rango 5-27) puede tener implicaciones para la atención de los individuos con sospecha de COVID-19 y para el sistema sanitario ⁽³⁶⁾.

Los estudios de esta revisión mostraron valores de precisión diagnóstica elevados cuando se incluyó más de una prueba, incluso superiores a los incluidos en revisiones anteriores de estudios moleculares para la detección de COVID-19 ⁽⁷⁾. Una revisión anterior ⁽³⁷⁾ mostró cómo la combinación de diferentes pruebas diagnósticas era muy recomendable para alcanzar valores adecuados de sensibilidad y especificidad diagnóstica.

Agrupamos los estudios según la población, que incluía población sintomática con sospecha de infección por SARS-CoV-2, población con diagnóstico confirmado y otros. Los modelos de diagnóstico deben desarrollarse en poblaciones lo más parecidas posible a las poblaciones en las que se aplicará la prueba en la práctica ^(38, 39); ya sea en aquellos pacientes con sospecha de infección (pacientes sintomáticos o contactos cercanos de personas con SARS-CoV-2) o en poblaciones de cribado. Alrededor del 70% de los estudios incluidos en esta revisión (16 de 23 estudios) seleccionaron pacientes sintomáticos que se presentaron en un centro hospitalario con sospecha de infección. Muchas infecciones por COVID-19 son asintomáticas pero las personas pueden transmitir el virus. Por lo tanto, la detección temprana de estos pacientes es esencial para interrumpir la vía de transmisión del virus y controlar el COVID-19. Sin embargo, dada la dificultad de detectar a estos pacientes y el hecho de que las pruebas de

rutina para los pacientes asintomáticos suelen incluir sólo una prueba (normalmente RT-qPCR), sólo uno de los estudios se realizó en una población que incluía no sólo pacientes sintomáticos, sino también a pacientes con contacto con pacientes positivos por COVID-19 conocidos o a pacientes que han viajado al extranjero, una situación habitual en la práctica clínica ⁽²⁵⁾.

Seis de los estudios de esta revisión incluían una población ya diagnosticada y un estudio no especificó las características clínicas de los pacientes incluidos. Estos estudios (30% de los estudios incluidos), por lo tanto, no proporcionaron resultados que puedan ayudar en la identificación de los casos que van a tratar en la práctica clínica. En una revisión de estudios moleculares ⁽⁷⁾, la mitad de los artículos incluían pacientes con un diagnóstico confirmado de infección por COVID-19.

Para que un modelo de diagnóstico se aplique en la práctica, debe validarse en una muestra independiente de aquella en la que se ha desarrollado ⁽⁴⁰⁾. Sólo siete de los 16 estudios que incluían pacientes sintomáticos con sospecha de infección ^(13, 16, 19-21, 27, 28) y uno de los dos estudios que incluían pacientes ya diagnosticados ⁽³⁴⁾ llevó a cabo dicha validación externa. Los autores manifestaron la necesidad de una mayor validación del modelo sólo en cinco de los nueve estudios que no incluyeron una validación externa ⁽²²⁻²⁶⁾. Ocho estudios ^(14, 15, 18, 22, 23, 25, 31, 32) llevaron a cabo una validación interna mediante la aplicación de técnicas como el bootstrap o la validación cruzada k-fold. Estas técnicas proporcionaron estimaciones razonablemente válidas del modelo de diagnóstico, pero no debe pasarse por alto la validación externa con tamaños de muestra considerables ⁽⁴¹⁾.

Los resultados de las pruebas de imagen y los recuentos de linfocitos se incluyeron con mayor frecuencia en modelos de diagnóstico de los pacientes sintomáticos con sospecha de infección. Un estudio anterior ⁽⁴²⁾ también demostró el valor diagnóstico del recuento de linfocitos. Hasta la fecha, el diagnóstico por imagen ha sido una prueba esencial en la práctica clínica para el diagnóstico de COVID ⁽⁴³⁾, dado que la presencia de neumonía bilateral es una

indicación de riesgo para el paciente. Sin embargo, los estudios no incluyen ninguna información sobre el uso de los hallazgos en las pruebas de imagen. Aunque los hallazgos de la radiografía de tórax y del TAC de tórax pueden ser útiles, también son inespecíficos, pueden cambiar con la progresión de la enfermedad o se superponen con otras neumonías víricas ⁽⁴⁴⁾. Otras variables como las características sociodemográficas, la presencia de comorbilidades o los días desde el inicio de los síntomas sólo se incluyeron en menos de la mitad de los estudios. Una revisión anterior que evaluaba los síntomas asociados a la infección no identificó ninguna combinación útil de síntomas que pudiera asociarse de forma fiable a la presencia de la infección ⁽¹¹⁾. Uno de los estudios ⁽¹³⁾ incluyó la primera RT-qPCR realizada en el modelo, estableciendo el resultado final de la RT-qPCR como el estándar de oro. Una revisión anterior ⁽⁴⁵⁾ indicó la frecuencia de resultados falsos negativos de la primera RT-qPCR; su inclusión en el modelo puede conducir a un alto riesgo de sesgo de verificación diagnóstica si sólo algunos de los participantes que recibieron esta primera RT-qPCR con un resultado concreto recibieran la prueba de referencia estándar ⁽⁴⁶⁾.

Diez de los 23 estudios incluidos aplicaron métodos de inteligencia artificial, nueve estudios centrados en pacientes sintomáticos ⁽²⁰⁻²⁸⁾, cuatro estudios en pacientes ya diagnosticados ^(30, 31, 33, 34) y en un estudio las características clínicas de los pacientes incluidos no se especifican ⁽³⁵⁾. Aunque se está intentando elaborar unas directrices para la descripción de los estudios diagnósticos basados en IA ⁽⁴⁷⁾, dada la novedad de la aplicación de este procedimiento, todavía no se han desarrollado completamente. Esto también puede haber contribuido a que estos estudios no se ajusten a los criterios de la guía STARD.

La media de cumplimiento de los criterios de la guía STARD fue de 17,6 sobre 34 (rango 5-27). Dadas las características de los estudios, tanto el modelo de diagnóstico como el estándar de referencia se llevaron a cabo en un corto espacio de tiempo, lo que limitó los errores derivados de cambios en el estado de la enfermedad de los pacientes. Sin embargo, esto también condujo a un

menor cumplimiento de los criterios relacionados con la evaluación independiente de las dos pruebas. Dentro de la sección de resultados, una gran mayoría de los estudios no describía el diagrama de flujo de pacientes, ni la distribución de la gravedad de la enfermedad y otros diagnósticos alternativos en aquellos pacientes con y sin la enfermedad en estudio. Esto puede limitar la utilidad del estudio para los clínicos que intentan aplicar los resultados en la práctica. Sólo el 35% de los estudios evaluados incluía una descripción del proceso de selección de pacientes (consecutivo, aleatorio u otro) y más del 60% de ellos no describían los criterios de inclusión de los pacientes. Aunque este sesgo también se ha descrito en la evaluación de pruebas moleculares para COVID-19 ⁽⁴⁸⁾, en esta revisión esta descripción representa una de las revisiones anteriores, en las que el 98% de los estudios presentaban un alto riesgo de sesgo de selección ⁽⁴⁹⁾. Los artículos no mostraban la determinación del tamaño de la muestra, ni describían cómo tratar los resultados indeterminados o ausentes. Por último, casi todos los estudios no mostraron el registro previo del estudio, así como si el protocolo era accesible. Esta es una cuestión importante, ya que nos permitiría establecer si la estrategia probada se propuso antes de ver los resultados. Si los investigadores sólo informan sobre las combinaciones de pruebas que obtienen valores altos de precisión diagnóstica, corren el riesgo de desarrollar modelos que se ajustan a los datos (overfitting) pero que tienen un valor clínico limitado. Esto es especialmente relevante cuando se une a la ausencia de validación en una muestra independiente.

Las limitaciones de este estudio están relacionadas principalmente con la búsqueda bibliográfica. Aunque la búsqueda se realizó en diferentes bases de datos (PubMed, Embase y las bases de datos bioRxiv y medRxiv), dada la amplia difusión de la investigación durante la pandemia, es posible que hayamos pasado por alto algunos estudios relevantes. Dada la urgencia por desarrollar nuevas estrategias de diagnóstico en la pandemia de COVID-19, puede haber habido un sesgo relacionado con la publicación de sólo aquellos resultados que muestran un hallazgo significativo.

7. CONCLUSIONES

Según esta revisión sistemática, la inclusión de más de una prueba diagnóstica en el proceso de diagnóstico de la infección por COVID-19 muestra valores de precisión diagnóstica elevados. Las características de las pruebas de imagen, los síntomas de los pacientes, sus características demográficas y el recuento de linfocitos fueron las variables más frecuentemente incluidas en los modelos de diagnóstico. Sin embargo, hay algunos aspectos a mejorar; los modelos desarrollados deben ser validados externamente antes de llegar a conclusiones sobre su utilidad en la práctica. Además, es importante tener en cuenta que la prueba debe evaluarse en poblaciones lo más parecidas posible a aquellas en las que se aplicará en la práctica. Aplicar las recomendaciones existentes para el desarrollo de pruebas diagnósticas, incluyendo el seguimiento de la guía STARD o similares en el caso de los estudios de IA, mejorará la información de los estudios de evaluación diagnóstica.

Tabla 1. Descripción de las características principales de los estudios incluidos en la revisión sistemática (características del estudio, ámbito y variables de la población incluida).

Referencia	Año	País	Diseño del estudio	Objetivo	Ámbito
Pacientes sintomáticos con sospecha de infección por COVID-19					
Guo X (13)	2020	China	Estudio de cohorte retrospectivo	Desarrollar un modelo predictivo integrado de funciones múltiples basado en un bosque aleatorio para diferenciar el COVID-19 de la gripe estacional y la neumonía causada por otros virus respiratorios comunes.	Dos hospitales
Langer T (14)	2020	Italia	Estudio de cohorte retrospectivo	Desarrollar un modelo de aprendizaje automático para predecir los resultados de RT-PCR para SARS-COV-2 utilizando datos clínicos, radiológicos y de laboratorio de rutina básicos disponibles	Servicio de urgencias en Niguarda Ca' Granda, Milán, Italia.
Hermans JJR (15)	2020	Países bajos	Estudio de cohorte prospectivo	Construir un modelo predictivo de aprendizaje automático basado en TC de tórax y datos adicionales para mejorar la precisión diagnóstica de la TC de tórax (validación cruzada de 10 veces)	El hospital Franciscus Gasthuis & Vlietland en Róterdam y Schiedam, Países bajos, que tiene un centro de trauma de nivel 2 con 48000 visitas anuales en el Departamento de Emergencias.

Yang HS (16)	2020	EEUU	Estudio transversal	Desarrollar un modelo de aprendizaje automático que integre la edad, el género, la raza y los análisis de sangre de laboratorio de rutina, que estén fácilmente disponibles con un tiempo de respuesta corto.	New York Presbyterian Hospital/Weill Cornell Medicine (NYPH/WCM) y New York Presbyterian Hospital/Lower Manhattan Hospital (NYPH/LMH) del 11 de marzo al 29 de abril.
AlJame M (17)	2021	Brasil y Italia	Estudio transversal	Proponer un modelo de predicción de aprendizaje automático para diagnosticar con precisión el COVID-19 a partir de datos clínicos y/o de laboratorio rutinarios.	Hospital Israelita Albert Einstein de São Paulo (Brasil) y Hospital San Raffaele de Milán (Italia)
Tschoellitsch T (18)	2021	Austria	Estudio de cohorte retrospectivo	Evaluar si el aprendizaje automático podía excluir la infección por RT-qPCR de SARS-CoV-2 utilizando los valores de laboratorio disponibles de forma rutinaria. (Validación interna mediante validación cruzada quíntuple).	Hospital Universitario Kepler de Linz (Austria)
Du R (19)	2021	China	Estudio de cohorte retrospectivo	Aplicar el aprendizaje automático a la tarea de detección de COVID-19 mediante marcadores básicos de laboratorio y explorar el papel complementario de las radiografías de tórax.	24 hospitales de Hong Kong
Kurstjens S (20)	2020	Países bajos	Estudio de cohorte retrospectivo	Desarrollar un algoritmo para evaluar rápidamente el riesgo de infección por SARS-CoV-2 de un individuo en el servicio de urgencias.	Urgencias de tres (cohorte de descubrimiento) y cuatro (cohorte de validación) hospitales diferentes.
Tordjman M (21)	2020	Francia	Estudio de cohorte retrospectivo	Determinar una puntuación de probabilidad previa a la prueba para la infección por SARS-CoV-2.	Cuatro hospitales universitarios diferentes entre el 10 de marzo y el 30 de abril de 2020.
Gupta-Wright A (22)	2021	Reino Unido	Estudio de cohorte retrospectivo	Desarrollar y validar internamente (bootstrap) una puntuación de riesgo de diagnóstico para predecir el riesgo de COVID-19 (incluido COVID-19 con RT-PCR negativo) entre las admisiones médicas.	Dos hospitales dentro de un NHS Trust agudo en Londres, Reino Unido
Vieceli T (23)	2020	Brasil	Estudio transversal	Desarrollar una herramienta predictiva útil para el diagnóstico de COVID-19 basada en datos clínicos, de laboratorio y de imagen previos a la confirmación de la prueba de RT-PCR.	Hospital de Clínicas de Porto Alegre
Pardo Lledias J (24)	2020	España	Estudio transversal	Definir un puntaje con diferentes probabilidades clínicas y analizar la utilidad de la repetición de frotis nasofaríngeos en base a estos.	Dos salas de Medicina Interna COVID en el Hospital Universitario Marqués de Valdecilla, Santander, España (hospital de tercer nivel que atiende a una población de unos 350.000 habitantes), de marzo a abril de 2020.

Trubiano JA (25)	2020	Australia	Estudio transversal	Informar sobre los predictores clínicos y epidemiológicos de la COVID-19 a partir de una base de datos prospectiva de derivación única y presentar una herramienta de decisión clínica de la COVID-19 lista para el punto de atención (validación interna con bootstrap).	Una clínica de cribado de evaluación rápida de COVID-19 en Austin Health el 11 de marzo de 2020.
Ahmed S (26)	2021	Pakistan	Estudio transversal	Validar el rendimiento de Corona-Score en una cohorte de pacientes paquistaníes que solicitan atención por sospecha de infección.	Sección de Patología Química en colaboración con la Sección de Patología Molecular, Departamento de Patología y Medicina de Laboratorio, Universidad Aga Khan (AKU), Karachi.
Elimian KO (27)	2021	Estados de Nigeria y Territorio de la Capital Federal	Estudio de cohorte retrospectivo	Desarrollar y validar la capacidad predictiva de los signos y síntomas clínicos sobre el resultado positivo en las pruebas de COVID-19.	Base de datos del Sistema de vigilancia, gestión y análisis de la respuesta a los brotes epidémicos (SORMAS) del 27 de febrero al 27 de agosto de 2020.
McRae AD (28)	2021	Canada	Estudio de cohorte prospectivo	Desarrollar y validar una puntuación de riesgo clínico que pueda cuantificar con precisión la probabilidad de infección por SARS-CoV- 2 en pacientes que acuden a un servicio de urgencias sin necesidad de pruebas de laboratorio.	32 servicios de urgencias de ocho provincias canadienses (registro canadiense de la Red de Respuesta Rápida de los Servicios de Urgencias COVID-19 (CCEDRRN)).
Infección confirmada por COVID 19					
Gatti M (29)	2020	Italia	Estudio de cohorte retrospectivo	Describir los hallazgos de la CXR y las características clínicas y de laboratorio asociadas con la CXR positiva y negativa.	Departamento de emergencia de dos hospitales del norte de Italia
Góreke V (30)	2021	Israel	Estudio transversal	Obtener un nuevo grupo de características basado en los hallazgos de laboratorio utilizando el algoritmo desarrollado y detectar COVID-19 con un nuevo clasificador híbrido basado en aprendizaje profundo que utiliza el nuevo grupo de características como parámetros de entrada.	Hospital Israelita Albert Einstein
Arpaci I (31)	2021	China	Estudio transversal	Predicir los casos COVID-19 positivos o negativos basándose en 14 características clínicas mediante algoritmos de clasificación de aprendizaje automático. (Validación interna mediante validación cruzada de 10 veces).	Hospital de Taizhou de la provincia china de Zhejiang
Marateb HR (32)	2021	Iran	Estudio de cohorte abierto	Diseñar e implementar un método de diagnóstico COVID-19 fiable para proporcionar el riesgo de infección utilizando datos demográficos, síntomas y signos, marcadores sanguíneos y antecedentes familiares de enfermedades para tener una excelente concordancia con los resultados obtenidos por la RT-PqCR y el CT-scan. (Validación interna mediante validación cruzada de 10 veces)	Khorshid Covid Cohort es un estudio de vigilancia hospitalaria para investigar pacientes con neumonía COVID-19 y no COVID desde febrero de 2020.

Yousif AY (33)	2022	Irak	Estudio transversal	Encontrar un método de diagnóstico para el virus COVID-19 en Irak mediante algoritmos de aprendizaje automático basados en las características de los análisis de sangre de los pacientes iraquíes intentando mejorar la precisión de la clasificación seleccionando la adecuada para la predicción precoz del COVID-19.	Muestras recogidas en numerosos laboratorios privados de Irak/Bagdad
Plante TB (34)	2020	EEUU e Israel	Casos y controles	Describir el desarrollo de un modelo de aprendizaje automático para descartar COVID-19 utilizando solo pruebas de laboratorio recolectadas de forma rutinaria.	Tres conjuntos de datos de pacientes en el departamento de emergencias (ED): A) Premier Healthcare Database (155 hospitales) B) Cedars-Sinai Medical center, un centro médico académico de 886 camas en Los Ángeles C) Beth Israel Deaconess Medical Center, un centro médico académico de 673 camas en Boston.
Otros					
Banerjee A (35)	2020	Brasil	Estudio transversal	Para utilizar el aprendizaje automático, una red neuronal artificial, un bosque aleatorio y una prueba estadística simple para identificar a los pacientes con SARS-CoV-2 positivo a partir de hemogramas completos sin conocimiento de los síntomas o el historial de los individuos.	Hospital Israelita Albert Einstein, en São Paulo, Brazil.

Tabla 1. (Continuación)

Referencia	Tamaño muestral	Sexo	Edad	Criterios de inclusión	Criterios de exclusión
Pacientes sintomáticos con sospecha de infección por COVID-19					
Guo X (13)	105 pacientes: a) cohorte descubrimiento 50 pacientes, y b) cohorte validación 55 pacientes	Cohorte de descubrimiento: hombres 60% Cohorte de validación: hombres 60%	Rango de edad Cohorte de descubrimiento: 25-31 años (23.1-62.5); Cohorte de validación: 31-39 años (25.0-52.3)	Pacientes sospechosos de COVID-19 que se presentaron en el hospital	NA
Langer T (14)	199	Hombres 63.8%	Media (rango): 65 [46–78]:	Pacientes que presentan síntomas compatibles con COVID-19 (fiebre, dolor de garganta, tos, disnea, dolor torácico, cefalea, síncope, astenia, artralgia, diarrea, náuseas y vómitos)	Edad < 12 años y ausencia de evaluación de la fórmula leucocitaria (definida como porcentajes de los cinco tipos de leucocitos: neutrófilos, linfocitos, eosinófilos, basófilos y monocitos) en el servicio de urgencias
Hermans JJR (15)	319 pacientes	Hombres RT-PCR+: 55.6%; RT-PCR-: 44.6%	Media (IQR) RT-PCR+: 59 (50-68); RT-PCR-: 62 (44-75)	Pacientes consecutivos que visitaron el servicio de urgencias entre el	(a) Infección por COVID-19 previamente confirmada; (b)

				27 de marzo y el 20 de abril de 2020, (a) edad \geq 18 años; (b) sospecha de infección por COVID-19 en combinación con al menos uno de los siguientes: (1) nuevos síntomas respiratorios que persisten durante \leq 2 semanas y presentes durante las últimas 24h, (2) saturación \leq 94% y/o frecuencia respiratoria \geq 20/min y/o molestias abdominales, y (3) alta sospecha clínica incluso en ausencia de síntomas; y (c) RT-PCR y TC de tórax realizadas dentro de las 24h una después de la otra.	inestabilidad definida como una saturación periférica de oxígeno $<$ 92% a pesar de 5 l de oxígeno y/o una presión arterial sistólica $<$ 90 mmHg; c) presentación principal por traumatismo de alta energía, trombólisis o síndrome coronario agudo; (d) embarazo; y (e) primer resultado de RT-PCR no interpretable.
Yang HS (16)	1898 pacientes: Conjunto de entrenamiento: 1402 pacientes positivos y 1954 negativos; Conjunto de validación: 496 pacientes positivos y 968 negativos	NYPH/WCM: 46,4% NYPH/LMH: 43,31%	Media (DE): NYPH/WCM: 56,44 (19,46) y NYPH/LMH: 56,20 (20,81)	Pacientes a los que se realizó una prueba de RT-PCR de COVID-19 en el Hospital.	Pacientes $<$ 18 años, pacientes que tuvieron resultados de RT-PCR indeterminados y pacientes sin resultados de laboratorio en los dos días anteriores a la realización de la prueba de RT-PCR
AlJame M (17)	5923	NA	NA	Pacientes ingresados en hospitales del 28 de marzo al 3 de abril de 2020 en Brasil, de finales de febrero de 2020 a mediados de marzo de 2020 en Italia.	NA
Tschoellitsch T (18)	1357	Hombres: 49,9%	Media 56,3 años (sd 26,6)	Pacientes con una prueba del SRAS realizada entre el 1 de marzo y el 30 de abril	NA
Du R (19)	Cohorte primaria: 5230 pacientes; Cohorte de validación: 605 pacientes	Hombres: 56%	NA	Pacientes con sospecha clínica de infección por COVID-19 que acudieron al servicio de urgencias desde el inicio del brote de COVID-19 y se sometieron a una prueba RT-PCR para detectar el SARS-CoV-2. Los criterios de inclusión fueron: (i) tener radiografías frontales de tórax en la fecha de la prueba RT-qPCR, (ii) tener pruebas de	Pacientes menores de 16 años.

				laboratorio realizadas, concretamente hemograma con o sin recuento diferencial, proteína C reactiva y lactato deshidrogenasa en la fecha de la prueba RT-qPCR.	
Kurstjens S (20)	967 pacientes: a) Cohorte descubrimiento: 375 pacientes, y b) Cohorte validación: 592 pacientes.	Cohorte de descubrimiento: hombres negativos a COVID-19 43%; Hombres positivos a COVID-19 64,1%. Cohorte de validación: COVID-19 negativo 53.3% hombres; COVID-19 positivo 63.8% hombres.	Media (DE). Cohorte descubrimiento: COVID-19 negativo 62 +-16; COVID-19 positivo 70 +-12 Cohorte de validación: COVID-19 negativo 63 +-17; COVID-19 positivo 69 +-12	Pacientes que acuden al servicio de urgencias con síntomas respiratorios o sospecha de infección por COVID-19 debido a molestias gastrointestinales (1-2 % de esta cohorte) y posterior SARS-CoV-2 RT-PCR	Pacientes de otros servicios y pacientes sin síntomas respiratorios ni sospecha de COVID-19. Para la población de validación, se excluyeron los pacientes con valores faltantes o muestras hemolíticas
Tordjman M (21)	605: Cohorte de descubrimiento: 200 pacientes (100 pacientes y 100 controles). Cohorte de validación 405 pacientes.	Cohorte descubrimiento: hombres: pacientes 65% y controles 45%. Cohorte de validación: 55% hombres	Cohorte de descubrimiento: media de pacientes 65 (sd 24) y media de controles 60 (sd 31). Cohorte de validación: media 65	Cohorte de descubrimiento: pacientes ambulatorios con resultados de RT-PCR y tomografía computarizada evaluados por sospecha de infección por SARS-Cov-2 entre el 10 de marzo y el 8 de abril de 2020. Cohorte de validación: pacientes ambulatorios consecutivos con sospecha de infección por SARS-Cov-2 con ambos RT-Resultados de PCR y CT-scan disponibles	Cohorte de descubrimiento: 1) diagnóstico aún en investigación; 2) falta de análisis de sangre que incluya un conteo completo de glóbulos blancos y electrolitos séricos; 3) ausencia de características clínicas reportadas
Gupta-Wright A (22)	4008 pacientes	Hombres: 58.3%	Media de 65 (IQR 57-76)	Todos los pacientes ingresados en salas médicas entre el 2 de marzo y el 3 de mayo de 2020 con una prueba de RT-PCR (la decisión de realizar la prueba se basó en una sospecha clínica).	NA
Vieceli T (23)	100	Hombres 43%	58 (40-69.5)	Los primeros pacientes consecutivos de 18 años o más ingresados en el hospital por sospecha de COVID-19 del 17 de marzo al 10 de abril de 2020.	Pacientes dados de alta en las 24 h siguientes al ingreso

Pardo Lledias J (24)	145	Hombres 59.9%	Media (SD). 66.9 (17.8)	Pacientes ingresados por sospecha de infección por SARS-CoV-2	NA
Trubiano JA (25)	2935	Hombres 36,5%	Media, 39 (IQR: 29, 53)	Pacientes evaluados para COVID-19 en una clínica de cribado debido a síntomas (98,3%), contacto con paciente positivo conocido para COVID-19 (17,3%) o viaje al extranjero (24,6%).	NA
Ahmed S (26)	60 pacientes: 30 RT-qPCR negativos y 30 RT-PCR positivos.	Hombres: 41,7%.	RT-qPCR positiva: media 33,1 años (DE, 6,5); RT-PqCR negativa: media 60,5 (DE, 16,5)	Los casos sospechosos de COVID-19 con síntomas respiratorios se recuperaron de las historias clínicas electrónicas.	NA
Elimian KO (27)	181.544 pacientes: conjunto de derivación: 90.722; conjunto de validación: 90,693.	Varones: niños, 62,2%; adultos, 64%, y ancianos, 63,7%.	Niños (≤ 17 años): 37%; Adultos: 61,8%; Ancianos: 1.8%.	Pacientes sometidos a pruebas de detección del SRAS-CoV-2 (de acuerdo con la definición de caso estándar del NCDC COVID-19 para casos sospechosos) con al menos un síntoma registrado positivamente. También se incluyeron las personas preocupadas por la infección por COVID-19 en el momento de su presentación en los centros de pruebas.	No se dispone del resultado de la PCR; no hay datos sobre edad, sexo.
McRae AD (28)	27665	Hombres: 49,5%.	Cohorte de derivación: media 57 años (IQR 38-73); cohorte de validación: media 56 años (IQR 37-73)	Pacientes consecutivos mayores de 18 años a los que se recogió una muestra biológica para la prueba de amplificación de ácidos nucleicos (NAAT) en su visita al servicio de urgencias o, si estaban ingresados, en las 24 horas siguientes a su llegada al servicio de urgencias por factores de riesgo conocidos de infección por SARS-CoV-2, trabajo como profesional sanitario, vida institucional, contactos personales o domésticos cercanos con	Pacientes que tuvieron una NAAT positiva para el SARS-CoV-2 en los 14 días anteriores a su visita al servicio de urgencias, pacientes con parada cardíaca antes de su llegada al servicio de urgencias y aquellos en los que faltaban datos de resultados.

				infección por SARS-CoV-2 y síntomas como tos, anosmia o disgeusia, fiebre, mialgias y signos vitales.	
Infección confirmada por COVID 19					
Gatti M (29)	260 pacientes	Hombres 61%	62.8 ± 15.8 años	Pacientes consecutivos admitidos en el departamento de urgencias de dos hospitales del norte de Italia entre el 1 y el 31 de marzo de 2020 con infección por COVID-19 confirmada con RT-PCR y los que se sometieron a CXR durante las primeras 24h tras el hisopado.	NA
Górecki V (30)	600 pacientes: 520 diagnosticados de COVID-19 y 80 individuos sanos.	NA	NA	NA	NA
Arpaci I (31)	114	Hombres: 59,6%	Media, 39,6 (DE 18,8) (intervalo 1-80)	NA	NA
Marateb HR (32)	Conjunto de datos 1: 634 pacientes con neumonía COVID-19 y 118 con neumonía no COVID-19. Conjunto de datos 2: 634 pacientes con COVID-19 y 634 y 598 con neumonía sin COVID-19; Conjunto de datos 3: individuos sanos de otros estudios/registros.	Conjunto de datos 1: Hombres: 47,5% (neumonía no COVID-19) y 61,4% (COVID-19) Conjunto de datos 2: 45,9%. Conjunto de datos 3 no proporcionado.	Conjunto de datos 1: media, 61,7 (DE, 18,3) (neumonía no COVID-19) y 57,0 (15,4) (pacientes con COVID-19) Conjunto de datos 2: mayores de 60 años; 89 (14,0%) Conjunto de datos 3: no facilitado	NA	NA
Yousif AY (33)	300 (213 no infectados, 87 infectados)	NA	NA	NA	NA

Plante TB (34)	192,778: conjunto de entrenamiento (12,183 pacientes) y conjunto de validación (172,754 pacientes); análisis de sensibilidad (7842 pacientes)	Edad media decil 50 (IQR 30-60 años)	Hombres: 40.5%	Adultos de 20 o más años en un servicio de urgencias incluido durante uno de los periodos de tiempo pandémico o pre-pandémico.	La falta de un resultado de laboratorio incluido en el modelo el día de la presentación al ED o si alguno de los resultados de laboratorio se informó con unidades inapropiadas o tipo de muestra incorrecto.
Otros					
Banerjee A (35)	598	NA	NA	Pacientes a los que se les recogieron muestras para realizar la RT-PCR de SARS-CoV-2 y pruebas de laboratorio adicionales durante una visita al hospital	Pacientes en unidad semi-intensiva y UCI

Tabla 2. Descripción de las variables incluidas en el estudio, los resultados obtenidos y las principales conclusiones del estudio.

Referencia	Tamaño muestral	Sexo	Edad	Criterios de inclusión	Criterios de exclusión
Pacientes sintomáticos con sospecha de infección por COVID-19					
Guo X (13)	105 pacientes: a) cohorte descubrimiento 50 pacientes, y b) cohorte validación 55 pacientes	Cohorte de descubrimiento: hombres 60% Cohorte de validación: hombres 60%	Rango de edad Cohorte de descubrimiento: 25-31 años (23.1-62.5); Cohorte de validación: 31-39 años (25.0-52.3)	Pacientes sospechosos de COVID-19 que se presentaron en el hospital	NA
Langer T (14)	199	Hombres 63.8%	Media (rango): 65 [46–78]:	Pacientes que presentan síntomas compatibles con COVID-19 (fiebre, tos, disnea, dolor torácico, cefalea, síncope, astenia, artralgia, diarrea, náuseas y vómitos)	Edad < 12 años y ausencia de evaluación de la fórmula leucocitaria (definida como porcentajes de los cinco tipos de leucocitos: neutrófilos, linfocitos, eosinófilos, basófilos y monocitos) en el servicio de urgencias
Hermans JJR (15)	319 pacientes	Hombres RT-PCR+: 55.6%; RT-PCR-: 44.6%	Media (IQR) RT-PCR+: 59 (50-68); RT-PCR-: 62 (44-75)	Pacientes consecutivos que visitaron el servicio de urgencias entre el 27 de marzo y el 20 de abril de 2020, (a) edad ≥ 18 años; (b) sospecha de infección por COVID-	(a) Infección por COVID-19 previamente confirmada; (b) inestabilidad definida como una saturación periférica de oxígeno < 92% a pesar de 5 l de

				19 en combinación con al menos uno de los siguientes: (1) nuevos síntomas respiratorios que persisten durante ≤ 2 semanas y presentes durante las últimas 24h, (2) saturación $\leq 94\%$ y/o frecuencia respiratoria $\geq 20/\text{min}$ y/o molestias abdominales, y (3) alta sospecha clínica incluso en ausencia de síntomas; y (c) RT-PCR y TC de tórax realizadas dentro de las 24h una después de la otra.	oxígeno y/o una presión arterial sistólica < 90 mmHg; c) presentación principal por traumatismo de alta energía, trombólisis o síndrome coronario agudo; (d) embarazo; y (e) primer resultado de RT-PCR no interpretable.
Yang HS (16)	1898 pacientes: Conjunto de entrenamiento: 1402 pacientes positivos y 1954 negativos; Conjunto de validación: 496 pacientes positivos y 968 negativos	NYPH/WCM: 46,4% NYPH/LMH: 43,31%	Media (DE): NYPH/WCM: 56,44 (19,46) y NYPH/LMH: 56,20 (20,81)	Pacientes a los que se realizó una prueba de RT-PCR de COVID-19 en el Hospital.	Pacientes < 18 años, pacientes que tuvieron resultados de RT-PCR indeterminados y pacientes sin resultados de laboratorio en los dos días anteriores a la realización de la prueba de RT-PCR
Allame M (17)	5923	NA	NA	Pacientes ingresados en hospitales del 28 de marzo al 3 de abril de 2020 en Brasil, de finales de febrero de 2020 a mediados de marzo de 2020 en Italia.	NA
Tschoellitsch T (18)	1357	Hombres: 49,9%	Media 56,3 años (sd 26,6)	Pacientes con una prueba del SRAS realizada entre el 1 de marzo y el 30 de abril	NA
Du R (19)	Cohorte primaria: 5230 pacientes; Cohorte de validación: 605 pacientes	Hombres: 56%	NA	Pacientes con sospecha clínica de infección por COVID-19 que acudieron al servicio de urgencias desde el inicio del brote de COVID-19 y se sometieron a una prueba RT-PCR para detectar el SARS-CoV-2. Los criterios de inclusión fueron: (i) tener radiografías frontales de tórax en la fecha de la prueba RT-qPCR, (ii) tener pruebas de laboratorio realizadas, concretamente hemograma con o sin recuento	Pacientes menores de 16 años.

				diferencial, proteína C reactiva y lactato deshidrogenasa en la fecha de la prueba RT-qPCR.	
Kurstjens S (20)	967 pacientes: a) Cohorte descubrimiento: 375 pacientes, y b) Cohorte validación: 592 pacientes.	Cohorte de descubrimiento: hombres negativos a COVID-19 43%; Hombres positivos a COVID-19 64,1%. Cohorte de validación: COVID-19 negativo 53.3% hombres; COVID-19 positivo 63.8% hombres.	Media (DE). Cohorte descubrimiento: COVID-19 negativo 62 +-16; COVID-19 positivo 70 +-12 Cohorte de validación: COVID-19 negativo 63 +-17; COVID-19 positivo 69 +-12	Pacientes que acuden al servicio de urgencias con síntomas respiratorios o sospecha de infección por COVID-19 debido a molestias gastrointestinales (1-2 % de esta cohorte) y posterior SARS-CoV-2 RT-PCR	Pacientes de otros servicios y pacientes sin síntomas respiratorios ni sospecha de COVID-19. Para la población de validación, se excluyeron los pacientes con valores faltantes o muestras hemolíticas
Tordjman M (21)	605: Cohorte de descubrimiento: 200 pacientes (100 pacientes y 100 controles). Cohorte de validación 405 pacientes.	Cohorte descubrimiento: hombres: pacientes 65% y controles 45%. Cohorte de validación: 55% hombres	Cohorte de descubrimiento: media de pacientes 65 (sd 24) y media de controles 60 (sd 31). Cohorte de validación: media 65	Cohorte de descubrimiento: pacientes ambulatorios con resultados de RT-PCR y tomografía computarizada evaluados por sospecha de infección por SARS-Cov-2 entre el 10 de marzo y el 8 de abril de 2020. Cohorte de validación: pacientes ambulatorios consecutivos con sospecha de infección por SARS-Cov-2 con ambos RT-Resultados de PCR y CT-scan disponibles	Cohorte de descubrimiento: 1) diagnóstico aún en investigación; 2) falta de análisis de sangre que incluya un conteo completo de glóbulos blancos y electrolitos séricos; 3) ausencia de características clínicas reportadas
Gupta-Wright A (22)	4008 pacientes	Hombres: 58.3%	Media de 65 (IQR 57–76)	Todos los pacientes ingresados en salas médicas entre el 2 de marzo y el 3 de mayo de 2020 con una prueba de RT-PCR (la decisión de realizar la prueba se basó en una sospecha clínica).	NA
Vieceli T (23)	100	Hombres 43%	58 (40–69.5)	Los primeros pacientes consecutivos de 18 años o más ingresados en el hospital por sospecha de COVID-19 del 17 de marzo al 10 de abril de 2020.	Pacientes dados de alta en las 24 h siguientes al ingreso
Pardo Lledias J (24)	145	Hombres 59.9%	Media (SD). 66.9 (17.8)	Pacientes ingresados por sospecha de infección por SARS-CoV-2	NA

Trubiano JA (25)	2935	Hombres 36,5%	Media, 39 (IQR: 29, 53)	Pacientes evaluados para COVID-19 en una clínica de cribado debido a síntomas (98,3%), contacto con paciente positivo conocido para COVID-19 (17,3%) o viaje al extranjero (24,6%).	NA
Ahmed S (26)	60 pacientes: 30 RT-qPCR negativos y 30 RT-PCR positivos.	Hombres: 41,7%.	RT-qPCR positiva: media 33,1 años (DE, 6,5); RT-PqCR negativa: media 60,5 (DE, 16,5)	Los casos sospechosos de COVID-19 con síntomas respiratorios se recuperaron de las historias clínicas electrónicas.	NA
Elimian KO (27)	181.544 pacientes: conjunto de derivación: 90.722; conjunto de validación: 90,693.	Varones: niños, 62,2%; adultos, 64%, y ancianos, 63,7%.	Niños (≤ 17 años): 37%; Adultos: 61,8%; Ancianos: 1.8%.	Pacientes sometidos a pruebas de detección del SRAS-CoV-2 (de acuerdo con la definición de caso estándar del NCDC COVID-19 para casos sospechosos) con al menos un síntoma registrado positivamente. También se incluyeron las personas preocupadas por la infección por COVID-19 en el momento de su presentación en los centros de pruebas.	No se dispone del resultado de la PCR; no hay datos sobre edad, sexo.
McRae AD (28)	27665	Hombres: 49,5%.	Cohorte de derivación: media 57 años (IQR 38-73); cohorte de validación: media 56 años (IQR 37-73)	Pacientes consecutivos mayores de 18 años a los que se recogió una muestra biológica para la prueba de amplificación de ácidos nucleicos (NAAT) en su visita al servicio de urgencias o, si estaban ingresados, en las 24 horas siguientes a su llegada al servicio de urgencias por factores de riesgo conocidos de infección por SARS-CoV-2, trabajo como profesional sanitario, vida institucional, contactos personales o domésticos cercanos con infección por SARS-CoV-2 y síntomas como tos, anosmia o disgeusia, fiebre,	Pacientes que tuvieron una NAAT positiva para el SARS-CoV-2 en los 14 días anteriores a su visita al servicio de urgencias, pacientes con parada cardíaca antes de su llegada al servicio de urgencias y aquellos en los que faltaban datos de resultados.

				mialgias y signos vitales.	
Infección confirmada por COVID 19					
Gatti M (29)	260 pacientes	Hombres 61%	62.8 ± 15.8 años	Pacientes consecutivos admitidos en el departamento de urgencias de dos hospitales del norte de Italia entre el 1 y el 31 de marzo de 2020 con infección por COVID-19 confirmada con RT-PCR y los que se sometieron a CXR durante las primeras 24h tras el hisopado.	NA
Góreke V (30)	600 pacientes: 520 diagnosticados de COVID-19 y 80 individuos sanos.	NA	NA	NA	NA
Arpaci I (31)	114	Hombres: 59,6%	Media, 39,6 (DE 18,8) (intervalo 1-80)	NA	NA
Marateb HR (32)	Conjunto de datos 1: 634 pacientes con neumonía COVID-19 y 118 con neumonía no COVID-19. Conjunto de datos 2: 634 pacientes con COVID-19 y 634 y 598 con neumonía sin COVID-19; Conjunto de datos 3: individuos sanos de otros estudios/registros.	Conjunto de datos 1: Hombres: 47,5% (neumonía no COVID-19) y 61,4% (COVID-19) Conjunto de datos 2: 45,9%. Conjunto de datos 3 no proporcionado.	Conjunto de datos 1: media, 61,7 (DE, 18,3) (neumonía no COVID-19) y 57,0 (15,4) (pacientes con COVID-19) Conjunto de datos 2: mayores de 60 años; 89 (14,0%) Conjunto de datos 3: no facilitado	NA	NA
Yousif AY (33)	300 (213 no infectados, 87 infectados)	NA	NA	NA	NA
Plante TB (34)	192,778: conjunto de entrenamiento (12,183 pacientes) y conjunto de validación (172,754 pacientes); análisis de	Edad media decil 50 (IQR 30-60 años)	Hombres: 40.5%	Adultos de 20 o más años en un servicio de urgencias incluido durante uno de los periodos de tiempo pandémico o pre-pandémico.	La falta de un resultado de laboratorio incluido en el modelo el día de la presentación al ED o si alguno de los resultados de laboratorio se

	sensibilidad (7842 pacientes)				informó con unidades inapropiadas o tipo de muestra incorrecto.
Otros					
Banerjee A (35)	598	NA	NA	Pacientes a los que se les recogieron muestras para realizar la RT-PCR de SARS-CoV-2 y pruebas de laboratorio adicionales durante una visita al hospital	Pacientes en unidad semi-intensiva y UCI

Tabla 3. Variables finalmente incluidas en el modelo para diagnosticar la infección en pacientes sintomáticos con sospecha de infección por COVID 19.

Referencias	AUC (IC 95%)	RT-qPCR	Demografía (edad, sexo)	Características clínicas					Imagen		Análisis de sangre completo			
				Comorbilidades	Síntomas	Exposición o viaje internacional	ECG	Días desde el inicio	CT	CXR	RBC	LYP	LYC	WBC
Guo X (13)	93 (79-100)	X							X			X		
Langer T (14)	90		X	X	X		X			X	X		X	
Hermans JJR (15)	91.4 (87.9-94.9)				X			X		X			X	
Yang HS (16)	85.4 (82.9-87.8)		X								X	X	X	X
Aljame M (17)	Precisión 99.5%		X								X	X	X	X
Tschoellitsch T (18)	0.74										X			
Du R (19)	89.9-95.8		X							X	X	X	X	X
Kurstjens S (20)	91		X						X	X			X	X
Tordjman M (21)	91.8 (89.1-94.6)												X	X
Gupta-Wright A (22)	83 (82-85)		X		X					X				
Vieceli T (23)	82.7 (75-90)									X			X	
Pardo Lledias J (24)	89 (83.1-94.6)				X			X						
Trubiano J (25)	83.6		X		X	X								
Ahmed S (26)	59 (0.45-0.74)									X		X	X	
Elimian KO (27)	Pérdida de olfato en niños: 0,56 (0,55- 0,56) Fiebre o tos en adultos: 0,57 (0,56-0,58). Dificultad para respirar en ancianos: 0.53 (0.48-0.58).				X									

Mc Rae AD (28)	Sensibilidad 97,4 (96,4-98,3) Especificidad: 95,9 (95,5-96)		X	X	X	X	X	X	X	X	X	X	X	X
----------------	--	--	---	---	---	---	---	---	---	---	---	---	---	---

Tabla 3. (Continuación)

Referencias	Análisis bioquímicos							
	Pruebas rutinarias de laboratorio	CRP	LDH	Ferritina	CK	Gasometría arterial/venosa	Colinesterasa	Tiempo de protrombina
Guo X (13)								
Langer T (14)								
Hermans JJR (15)				X	X			
Yang HS (16)	X							
Aljame M (17)	X							
Tschoellitsch T (18)	X	X					X	X
Du R (19)		X	X					
Kurstjens S (20)		X	X	X				
Tordjman M (21)								
Gupta-Wright A (22)		X				X		
Vieceli T (23)			X					
Pardo Lledias J (24)			X					
Trubiano J (25)								
Ahmed S (26)		X	X	X				
Elimian KO (27)								
Mc Rae AD (28)	X	X	X	X	X	X		

Tabla 4. Variables finalmente incluidas en el modelo para diagnosticar la infección por SARS-COV-2 en pacientes diagnosticados de COVID-19.

Referencias	AUC	Demografía (edad, sexo, raza)	Profesión	Contacto con pacientes de COVID-19	Características clínicas			Imagen		Análisis de sangre completo			
					Comorbilidades	Síntomas	Días desde el inicio	CT	CXR	RBC	LYP	LYC	WBC
Gatti M (29)	87.8						X		X				
Plante TB (34)	91 (90-92)	X					X						
Goreke V (30)	100									X	X	X	X
Arpaci I (31)	Precisión 84.2									X	X	X	X
Marateb HR (32)	87-96	X	X	X	X	X							X
Yousif AY (33)	Precisión 87									X	X	X	X

Tabla 4. (Continuación)

Referencias	Análisis bioquímicos					
	Pruebas rutinarias de laboratorio	CRP	LDH	Ferritina	Gasometría arterial/venosa	Procalcitonina
Gatti M (29)		X	X			
Plante TB (34)	X					
Goreke V (30)	X	X				
Arpaci I (31)						X
Marateb HR (32)	X				X	X
Yousif AY (33)	X				X	

Tabla 5. Descripción del cumplimiento de los ítems incluidos en la lista de comprobación de las directrices Standards for Reporting Diagnostic Accuracy Studies (STARD).

			Y	(%)
Título o resumen				
	1	Identificación como estudio de precisión diagnóstica utilizando al menos una medida de precisión (como sensibilidad, especificidad, valores predictivos o AUC).	12	52.2
Resumen				
	2	Resumen estructurado del diseño, los métodos, los resultados y las conclusiones del estudio (para una orientación específica, véase STARD for Abstracts).	16	69.6
Introducción				
	3	Antecedentes científicos y clínicos, incluido el uso previsto y el papel clínico de la prueba índice.	21	91.3
	4	Objetivos e hipótesis del estudio.	21	91.3
Métodos				
Diseño del estudio				
	5	Si la recogida de datos se planificó antes de realizar la prueba índice y el estándar de referencia (estudio prospectivo) o después (estudio retrospectivo).	18	78.3
Participantes				
	6	Criterios de elegibilidad.	13	56.5
	7	En qué se basó la identificación de los participantes potencialmente elegibles (síntomas, resultados de pruebas anteriores, inclusión en el registro, etc.).	17	73.9
	8	Dónde y cuándo se identificó a los participantes potencialmente elegibles (entorno, lugar y fechas).	19	82.6
	9	Si los participantes formaban una serie consecutiva, aleatoria o de conveniencia.	12	52.2
Métodos de ensayo				
	10 a	Prueba de índice, con suficiente detalle para permitir su reproducción.	17	73.9
	10 b	Norma de referencia, con suficiente detalle para permitir su reproducción.	14	60.9
	11	Justificación de la elección del estándar de referencia (si existen alternativas).	10	43.5
	12 a	Definición y justificación de los puntos de corte de positividad de la prueba o de las categorías de resultados de la prueba índice, distinguiendo los preespecificados de los exploratorios.	14	60.9
	12 b	Definición y justificación de los umbrales de positividad de las pruebas o de las categorías de resultados del patrón de referencia, distinguiendo los preespecificados de los exploratorios.	9	39.1
	13 a	Si la información clínica y los resultados del estándar de referencia estaban a disposición de los ejecutores/lectores de la prueba índice.	1	4.35
	13 b	Si la información clínica y los resultados de las pruebas índice estaban a disposición de los evaluadores de la norma de referencia.	6	26
Análisis				
	14	Métodos para estimar o comparar medidas de precisión diagnóstica.	20	86.9
	15	Cómo se trataron los resultados indeterminados de la prueba índice o del patrón de referencia.	1	4.35
	16	Cómo se trataron los datos que faltaban sobre la prueba índice y el patrón de referencia.	5	21.7
	17	Cualquier análisis de la variabilidad de la precisión diagnóstica, distinguiendo los preespecificados de los exploratorios.	13	56.5
	18	Tamaño previsto de la muestra y cómo se ha determinado.	2	8.69
Resultados				
Participantes				
	19	Flujo de participantes utilizando un diagrama.	8	34.8
	20	Características demográficas y clínicas basales de los participantes.	18	78.3
	21 a	Distribución de la gravedad de la enfermedad en las personas que la padecen.	10	43.5
	21 b	Distribución de diagnósticos alternativos en las personas sin la enfermedad diana.	5	21.7
	22	Intervalo de tiempo y cualquier intervención clínica entre la prueba índice y el estándar de referencia.	4	17.4
Resultados test				
	23	Tabulación cruzada de los resultados de la prueba índice (o su distribución) por los resultados del estándar de referencia.	19	82.6
	24	Estimaciones de la exactitud diagnóstica y su precisión (como intervalos de confianza del 95%).	20	86.9
	25	Cualquier acontecimiento adverso derivado de la realización de la prueba índice o del estándar de referencia	0	0
Discusión				
	26	Limitaciones del estudio, incluidas las fuentes de posible sesgo, incertidumbre estadística y generalizabilidad.	21	91.3
	27	Implicaciones para la práctica, incluido el uso previsto y el papel clínico de la prueba índice.	19	82.6
Otra información				
	28	Nº de registro y nombre del registro.	2	8.69
	29	Dónde puede consultarse el protocolo completo del estudio.	2	8.69
	30	Fuentes de financiación y otras ayudas; papel de los financiadores.	15	65.2

8. BIBLIOGRAFÍA

1. Mina MJ, Andersen KG. 2021. COVID-19 testing: One size does not fit all. *Science* 371:126–127. <https://doi.org/10.1126/science.abe9187>.
2. Henderson CE, Rezai S, Jackman JM. 2021. Universal screening for novel coronavirus disease 2019 (COVID-19) for asymptomatic parturients: May not be beneficial at this time. *J Med Virol* 93:1198–1200. <https://doi.org/10.1002/jmv.26507>.
3. Yanes-Lane M, Winters N, Fregonese F, Bastos M, Perlman-Arrow S, Campbell JR, Menzies D. 2020. Proportion of asymptomatic infection among COVID-19 positive persons and their transmission potential: a systematic review and meta-analysis. *PLoS One* 15:e0241536. <https://doi.org/10.1371/journal.pone.0241536>.
4. Quaresima V, Naldini MM, Cirillo DM. 2020. The prospects for the SARS-CoV-2 pandemic in Africa. *EMBO Mol Med* 12:e12488. <https://doi.org/10.15252/emmm.202012488>.
5. Vila Muntadas M, Agustí Sunyer I, Agustí Garcia-Navarro A. 2021. COVID19 diagnostic tests: Importance of the clinical context. *Med Clin (Engl Ed)* 157:185–190. <https://doi.org/10.1016/j.medcle.2021.03.008>.
6. Watson J, Whiting PF, Brush JE. 2020. Interpreting a covid-19 test result. *BMJ* 369:m1808. <https://doi.org/10.1136/bmj.m1808>.
7. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, Dittrich S, Emperador D, Takwoingi Y, Cunningham J, Beese S, Domen J, Dretzke J, Ferrante di Ruffano L, Harris IM, Price MJ, Taylor-Phillips S, Hooft L, Leeflang MMG, McInnes MDF, Spijker R, Van den Bruel A, Cochrane COVID-19 Diagnostic Test Accuracy Group. 2021. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* 24:CD013705. <https://doi.org/10.1002/14651858.CD013705.pub2>.
8. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HCW, Bossuyt PMM. 2016. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open* 6:e012799. <https://doi.org/10.1136/bmjopen-2016-012799>.
9. Axell-House DB, Lavingia R, Rafferty M, Clark E, Amirian ES, Chiao EY. 2020. The estimation of diagnostic accuracy of tests for COVID-19: a scoping review. *J Infect* 81:681–697. <https://doi.org/10.1016/j.jinf.2020.08.043>.
10. Sayan M, Sarigul Yildirim F, Sanlidag T, Uzun B, Uzun Ozsahin D, Ozsahin I. 2020. Capacity Evaluation of Diagnostic Tests For COVID-19 Using Multicriteria Decision-Making Techniques. *Comput Math Methods Med* 2020: 1560250. <https://doi.org/10.1155/2020/1560250>.
11. Struyf T, Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Leeflang MM, Cochrane COVID-19 Diagnostic Test Accuracy Group. 2021. Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19. *Cochrane Database Syst Rev* 2CD013665.
12. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. *Open Me* 3:e123.
13. Guo X, Li Y, Li H, Li X, Chang X, Bai X, Song Z, Li J, Li K. 2020. An improved multivariate model that distinguishes COVID-19 from seasonal flu and other respiratory diseases. *Aging (Albany NY)* 12:19938–19944. <https://doi.org/10.18632/aging.104132>.
14. Langer T, Favarato M, Giudici R, Bassi G, Garberi R, Villa F, Gay H, Zeduri A, Bragagnolo S, Molteni A, Beretta A, Corradin M, Moreno M, Vismara C, Perno CF, Buscema M, Grossi E,

- Fumagalli R. 2020. Development of machine learning models to predict RT-PCR results for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in patients with influenza like symptoms using only basic clinical data. *Scand J Trauma Resusc Emerg Med* 28:113. <https://doi.org/10.1186/s13049-020-00808-8>.
15. Hermans JJR, Groen J, Zwets E, Boxma-De Klerk BM, Van Werkhoven JM, Ong DSY, Hanselaar WEJJ, Waals-Prinzen L, Brown V. 2020. Chest CT for triage during COVID-19 on the emergency department: myth or truth? *Emerg Radiol* 27:641–651. <https://doi.org/10.1007/s10140-020-01821-1>.
16. Yang HS, Hou Y, Vasovic LV, Steel PAD, Chadburn A, Racine-Brzostek SE, Velu P, Cushing MM, Loda M, Kaushal R, Zhao Z, Wang F. 2020. Routine Laboratory Blood Tests Predict SARS-CoV-2 Infection Using Machine Learning. *Clin Chem* 66:1396–1404. <https://doi.org/10.1093/clinchem/hvaa200>.
17. AlJame M, Imtiaz A, Ahmad I, Mohammed A. 2021. Deep forest model for diagnosing COVID-19 from routine blood tests. *Sci Rep* 11:16682. <https://doi.org/10.1038/s41598-021-95957-w>.
18. Tschoellitsch T, Dünser M, Böck C, Schwarzbauer K, Meier J. 2021. Machine Learning Prediction of SARS-CoV-2 Polymerase Chain Reaction Results with Routine Blood Tests. *Lab Med* 52:146–149. <https://doi.org/10.1093/labmed/lmaa111>.
19. Du R, Tsougenis ED, Ho JWK, Chan JKY, Chiu KWH, Fang BXH, Ng MY, Leung S-T, Lo CSY, Wong H-YF, Lam H-YS, Chiu L-FJ, So TY, Wong KT, Wong YCI, Yu K, Yeung Y-C, Chik T, Pang JWK, Wai AK-C, Kuo MD, Lam TPW, Khong P-L, Cheung N-T, Vardhanabhuti V. 2021. Machine learning application for the prediction of SARS-CoV-2 infection using blood tests and chest radiograph. *Sci Rep* 11:14250. <https://doi.org/10.1038/s41598-021-93719-2>.
20. Kurstjens S, van der Horst A, Herpers R, Geerits MWL, Kluiters-de Hingh YCM, Göttgens E-L, Blaauw MJT, Thelen MHM, Elisen MGLM, Kusters R. 2020. Rapid identification of SARS-CoV-2-infected patients at the emergency department using routine testing. *Clin Chem Lab Med* 58:1587–1593. <https://doi.org/10.1515/ccim-2020-0593>.
21. Tordjman M, Mekki A, Mali RD, Saab I, Chassagnon G, Guillo E, Burns R, Eshagh D, Beaune S, Madelin G, Bessis S, Feydy A, Mihoubi F, Doumenc B, Mouthon L, Carlier R-Y, Drapé J-L, Revel M-P. 2020. Pre-test probability for SARS-Cov-2-related infection score: The PARIS score. *PLoS One* 15:e0243342. <https://doi.org/10.1371/journal.pone.0243342>.
22. Gupta-Wright A, Macleod CK, Barrett J, Filson SA, Corrah T, Parris V, Sandhu G, Harris M, Tennant R, Vaid N, Takata J, Duraisingham S, Gandy N, Chana H, Whittington A, McGregor A, Papineni P. 2021. False-negative RT-PCR for COVID-19 and a diagnostic risk score: a retrospective cohort study among patients admitted to hospital. *BMJ Open* 11:e047110. <https://doi.org/10.1136/bmjopen-2020-047110>.
23. Vieceli T, Oliveira Filho C. M d, Berger M, Saadi MP, Salvador PA, Anizelli LB, Crivelaro P. C d F, Butzke M, Zappellini R. d S, Seligman BGDS, Seligman R. 2020. A predictive score for COVID-19 diagnosis using clinical, laboratory and chest image data. *Braz J Infect Dis* 24:343–348. <https://doi.org/10.1016/j.bjid.2020.06.009>.
24. Pardo Lledias J, Ayarza L, González-García P, Salmón González Z, Calvo Montes J, Gozalo Marguello M, Hernández Hernández JL, Olmos Martínez JM, Servicio de Medicina Interna. Hospital Universitario Marqués de Valdecilla-IDIVAL. Santander. Cantabria. 2020. Repetition of microbiological tests in suspect of SARS-CoV-2 infection: utility of a score based on clinical probability. *Rev Esp Quimioter* 33:410–414. <https://doi.org/10.37201/req/080.2020>.
25. Trubiano JA, Vogrin S, Smibert OC, Marhoon N, Alexander AA, Chua KYL, James FL, Jones NRL, Grigg SE, Xu CLH, Moini N, Stanley SR, Birrell MT, Rose MT, Gordon CL, Kwong JC, Holmes NE. 2020. COVID-MATCH65-A prospectively derived clinical decision rule for severe

acute respiratory syndrome coronavirus 2. PLoS One 15:e0243414. <https://doi.org/10.1371/journal.pone.0243414>.

26. Ahmed S, Naeem Effendi MU, Ahmed ZA, Siddiqui I, Jafri L. 2021. A cross sectional evaluation of the corona-score for swift identification of SARSCoV-2 infection at a tertiary care hospital in Pakistan. *Ann Med Surg (Lond)* 67:102537. <https://doi.org/10.1016/j.amsu.2021.102537>.

27. Elimian KO, Aderinola O, Gibson J, Myles P, Ochu CL, King C, Okwor T, Gaudenzi G, Olayinka A, Zaiyad HG, Ohonsi C, Ebhodaghe B, Dan-Nwafor C, Nwachukwu W, Abdus-Salam IA, Akande OW, Falodun O, Arinze C, Ezeokafor C, Jafiya A, Ojimba A, Aremu JT, Joseph E, Bowale A, Mutiu B, Saka B, Jinadu A, Hamza K, Ibeh C, Bello S, Asuzu M, Mba N, Oladejo J, Ilori E, Alfvén T, Igumbor E, Ihekweazu C. 2021. Assessing the capacity of symptom scores to predict COVID-19 positivity in Nigeria: a national derivation and validation cohort study. *BMJ Open* 11:e049699. <https://doi.org/10.1136/bmjopen-2021-049699>.

28. McRae AD, Hohl CM, Rosychuk R, Vatanpour S, Ghaderi G, Archambault PM, Brooks SC, Cheng I, Davis P, Hayward J, Lang E, Ohle R, Rowe B, Welsford M, Yadav K, Morrison LJ, Perry J, Canadian COVID-19 Emergency Department Rapid Response Network (CCEDRRN) investigators for the Network of Canadian Emergency Researchers and the Canadian Critical Care Trials Group. 2021. CCEDRRN COVID-19 Infection Score (CCIS): development and validation in a Canadian cohort of a clinical risk score to predict SARS-CoV-2 infection in patients presenting to the emergency department with suspected COVID-19. *BMJ Open* 11:e055832. <https://doi.org/10.1136/bmjopen-2021-055832>.

29. Gatti M, Calandri M, Barba M, Biondo A, Geninatti C, Gentile S, Greco M, Morrone V, Piatti C, Santonocito A, Varello S, Bergamasco L, Cavallo R, Di Stefano R, Riccardini F, Boccuzzi A, Limerutti G, Veltri A, Fonio P, Faletti R. 2020. Baseline chest X-ray in coronavirus disease 19 (COVID-19) patients: association with clinical and laboratory data. *Radiol Med* 125:1271–1279. <https://doi.org/10.1007/s11547-020-01272-1>.

30. Göreke V, Sarı V, Kockanat S. 2021. A novel classifier architecture based on deep neural network for COVID-19 detection using laboratory findings. *Appl Soft Comput* 106:107329. <https://doi.org/10.1016/j.asoc.2021.107329>.

31. Arpaci I, Huang S, Al-Emran M, Al-Kabi MN, Peng M. 2021. Predicting the COVID-19 infection with fourteen clinical features using machine learning classification algorithms. *Multimed Tools Appl* 80:11943–11957. <https://doi.org/10.1007/s11042-020-10340-7>.

32. Marateb HR, Ziaie Nezhad F, Mohebian MR, Sami R, Haghjooy Javanmard S, Dehghan Niri F, Akafzadeh-Savari M, Mansourian M, Mañanas MA, Wolkewitz M, Binder H. 2021. Automatic classification between COVID-19 and non-COVID-19 pneumonia using symptoms, comorbidities, and laboratory findings: The Khorshid COVID Cohort Study. *Front Med (Lausanne)* 8:768467. <https://doi.org/10.3389/fmed.2021.768467>.

33. Yousif AY, Younis SM, Hussein SA, Al-Saidi NMG. 2022. An intelligent computing for diagnosing COVID-19 using available blood tests. *ICIC International* 18:57–72.

34. Plante TB, Blau AM, Berg AN, Weinberg AS, Jun IC, Tapson VF, Kanigan TS, Adib AB. 2020. Development and external validation of a machine-learning tool to rule out COVID-19 among adults in the emergency department using routine blood tests: a large, multicenter, real-world study. *J Med Internet Res* 22:e24048. <https://doi.org/10.2196/24048>.

35. Banerjee A, Ray S, Vorselaars B, Kitson J, Mamalakis M, Weeks S, Baker M, Mackenzie LS. 2020. Use of machine learning and artificial intelligence to predict SARS-CoV-2 infection from full blood counts in a population. *Int Immuno-pharmacol* 86:106705. <https://doi.org/10.1016/j.intimp.2020.106705>.

36. Glasziou PP, Sanders S, Hoffmann T. 2020. Waste in covid-19 research. *BMJ* 369:m1847.
37. Böger B, Fachi MM, Vilhena RO, Cobre AF, Tonin FS, Pontarolo R. 2021. Systematic review with meta-analysis of the accuracy of diagnostic tests for COVID-19. *Am J Infect Control* 49:21–29. <https://doi.org/10.1016/j.ajic.2020.07.011>.
38. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M, Winget M, Yasui Y. 2001. Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst* 93:1054–1061. <https://doi.org/10.1093/jnci/93.14.1054>
39. Lumbreras B, Parker LA, Porta M, Pollán M, Ioannidis JPA, Hernández-Aguado I. 2009. Overinterpretation of clinical applicability in molecular diagnostic research. *Clin Chem* 55:786–794. <https://doi.org/10.1373/clinchem.2008.121517>.
40. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. 2015. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *J Clin Epidemiol* 68:279–289. <https://doi.org/10.1016/j.jclinepi.2014.06.018>.
41. Bleeker SE, Moll HA, Steyerberg EW, Donders ART, Derksen-Lubsen G, Grobbee DE, Moons KGM. 2003. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 56:826–832. [https://doi.org/10.1016/S0895-4356\(03\)00207-5](https://doi.org/10.1016/S0895-4356(03)00207-5).
42. Guneyasu F, Yurumez Y, Guclu E, Koroglu M, Karacan A, Guner NG, Durgun Y. 2020. The diagnostic process of covid-19 in the emergency department: laboratory and imaging methods. *Rev Assoc Med Bras* 66:58–64. <https://doi.org/10.1590/1806-9282.66.s2.58>.
43. Gillman AG, Lunardo F, Prinable J, Belous G, Nicolson A, Min H. 2021. Automated COVID-19 diagnosis and prognosis with medical imaging and who is publishing: a systematic review. *Phys Eng Sci Med* 17:1–17.
44. Ye Z, Zhang Y, Wang Y, Huang Z, Song B. 2020. Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *Eur Radiol* 30:4381–4389. <https://doi.org/10.1007/s00330-020-06801-0>.
45. Arevalo-Rodriguez I, Buitrago-Garcia D, Simancas-Racines D, Zambrano-Achig P, Del Campo R, Ciapponi A, Sued O, Martinez-García L, Rutjes AW, Low N, Bossuyt PM, Perez-Molina JA, Zamora J. 2020. False-negative results of initial RT-PCR assays for COVID-19: A systematic review. *PLoS One* 15:e0242958. <https://doi.org/10.1371/journal.pone.0242958>.
46. O'Sullivan JW, Banerjee A, Heneghan C, Pluddemann A. 2018. Verification bias. *BMJ Evid Based Med* 23:54–55. <https://doi.org/10.1136/bmjebm-2018-110919>.
47. Sounderajah V, Ashrafian H, Golub RM, Shetty S, De Fauw J, Hooft L, Moons K, Collins G, Moher D, Bossuyt PM, Darzi A, Karthikesalingam A, Denniston AK, Mateen BA, Ting D, Treanor D, King D, Greaves F, Godwin J, Pearson-Stuttard J, Harling L, McInnes M, Rifai N, Tomasev N, Normahani P, Whiting P, Aggarwal R, Vollmer S, Markar SR, Panch T, Liu X, STARD-AI Steering Committee. 2021. Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol. *BMJ Open* 11:e047709. <https://doi.org/10.1136/bmjopen-2020-047709>.
48. Bossuyt PM. 2020. Testing COVID-19 tests faces methodological challenges. *J Clin Epidemiol* 126:172–176. <https://doi.org/10.1016/j.jclinepi.2020.06.037>.
49. Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui L-P, Johnston JC, Lan Z, Law S, MacLean E, Trajman A, Menzies D, Benedetti A, Ahmad Khan F. 2020. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. *BMJ* 370:m2516. <https://doi.org/10.1136/bmj.m2516>