

# Pruebas falsas negativas para la infección por SARS-CoV-2: desafíos e implicaciones

Listado de autores.

- Steven Woloshin, MD,
- Neeraj Patel, BA,
- y Aaron S. Kesselheim, MD, JD, MPH

5 de junio de 2020

DOI: 10.1056 / NEJMp2015897

Este artículo fue publicado el 5 de junio de 2020 en NEJM.org.

Existe un amplio consenso de que las pruebas generalizadas de SARS-CoV-2 son esenciales para reabrir de manera segura los Estados Unidos. Una gran preocupación ha sido la disponibilidad de la prueba, pero la precisión de la prueba puede ser un problema mayor a largo plazo.

Si bien el debate se ha centrado en la precisión de las pruebas de anticuerpos, que identifican infecciones previas, las pruebas de diagnóstico, que identifican la infección actual, han recibido menos atención. Pero las pruebas de diagnóstico inexactas socavan los esfuerzos para contener la pandemia.

**Las pruebas de diagnóstico (que generalmente involucran un hisopo nasofaríngeo) pueden ser inexactas de dos maneras.** Un resultado falso positivo etiqueta erróneamente a una persona infectada, con consecuencias que incluyen cuarentena innecesaria y rastreo de contactos. Los resultados falsos negativos son más importantes porque las personas infectadas, que pueden ser asintomáticas, pueden no estar aisladas y pueden infectar a otras.

Dada la necesidad de saber qué tan bien las pruebas de diagnóstico descartan la infección, es importante revisar la evaluación de la precisión de las pruebas por parte de la Administración de Alimentos y Medicamentos (FDA) e investigadores clínicos, así como la interpretación de los resultados de las pruebas en una pandemia.

La FDA ha otorgado autorizaciones de uso de emergencia (EUA) a fabricantes de pruebas comerciales y ha emitido una guía sobre validación de pruebas.<sup>1</sup> La agencia requiere la medición del rendimiento de las pruebas analíticas y clínicas. La sensibilidad analítica indica la probabilidad de que la prueba sea positiva para material que contenga cepas de virus y la concentración mínima que la prueba puede detectar. La especificidad

analítica indica la probabilidad de que la prueba sea negativa para material que contiene patógenos distintos del virus objetivo.

**Las evaluaciones clínicas, que evalúan el rendimiento de una prueba en muestras de pacientes, varían entre los fabricantes.** La FDA prefiere el uso de "muestras clínicas naturales", pero ha permitido el uso de "muestras artificiales" producidas al agregar ARN viral o virus inactivado al material clínico sobrante. Por lo general, los estudios de rendimiento de la prueba implican que los pacientes se sometan a una prueba de índice y una prueba de "estándar de referencia" para determinar su verdadero estado. La sensibilidad clínica es la proporción de pruebas de índice positivo en pacientes que de hecho tienen la enfermedad en cuestión. La sensibilidad y su medición pueden variar con el entorno clínico. Para una persona enferma, es probable que la prueba estándar de referencia sea un diagnóstico clínico, idealmente establecido por un panel de adjudicación independiente cuyos miembros desconocen los resultados de la prueba de índice. Para SARS-CoV-2, No está claro si la sensibilidad de cualquier prueba comercial autorizada por la FDA se ha evaluado de esta manera. Según las EUA, la FDA sí permite a las empresas demostrar el rendimiento de la prueba clínica al establecer el acuerdo de la nueva prueba con una prueba autorizada de transcriptasa inversa-reacción en cadena de la polimerasa (RT-PCR) en material positivo conocido de personas sintomáticas o muestras artificiales. El uso de muestras positivas o artificiales conocidas puede conducir a sobreestimar la sensibilidad de la prueba, ya que los hisopos pueden perder material infectado en la práctica. la FDA sí permite a las empresas demostrar el rendimiento de las pruebas clínicas al establecer el acuerdo de la nueva prueba con una prueba autorizada de transcriptasa inversa-reacción en cadena de la polimerasa (RT-PCR) en material positivo conocido de personas sintomáticas o muestras artificiales. El uso de muestras positivas o artificiales conocidas puede conducir a sobreestimar la sensibilidad de la prueba, ya que los hisopos pueden perder material infectado en la práctica.<sup>1</sup>

**El diseño de un estándar de referencia para medir la sensibilidad de las pruebas de SARS-CoV-2 en personas asintomáticas es un problema no resuelto que necesita atención urgente para aumentar la confianza en los resultados de las**

**pruebas con fines de rastreo o detección de contactos.** Simplemente seguir a las personas para el desarrollo posterior de síntomas puede ser inadecuado, ya que pueden permanecer asintomáticos y ser infecciosos. La evaluación de la sensibilidad clínica en personas asintomáticas no se había informado para ninguna prueba comercial al 1 de junio de 2020.

Dos estudios de la provincia de Wuhan, China, suscitan preocupación por las pruebas de RT-PCR falsas negativas en pacientes con aparente enfermedad de Covid-19. En una preimpresión, Yang et al. describió 213 pacientes hospitalizados con Covid-19, de los cuales 37 estaban en estado crítico.<sup>2</sup> Recogieron 205 muestras de garganta, 490 muestras nasales y 142 muestras de esputo (mediana, 3 por paciente) y utilizaron una prueba de RT-PCR aprobada por el regulador chino. En los días 1 a 7 después del inicio de la enfermedad, el 11% del esputo, el 27% de las muestras nasales y el 40% de las muestras de garganta se consideraron falsamente negativas. Zhao y col. estudiaron 173 pacientes hospitalizados con síntomas respiratorios agudos y una TC de tórax "típica" de Covid-19, o SARS-CoV-2 detectada en al menos una muestra respiratoria. Se observó seroconversión de anticuerpos en el 93%.<sup>3</sup> Las pruebas de RT-PCR de muestras respiratorias tomadas en los días 1 a 7 de hospitalización fueron positivas al SARS-CoV-2 en al menos una muestra del 67% de los pacientes. Ninguno de los estudios informó el uso de un panel independiente, ignorante de los resultados de las pruebas de índice, para establecer un diagnóstico final de la enfermedad de Covid-19, lo que puede haber sesgado a los investigadores a sobrestimar la sensibilidad.

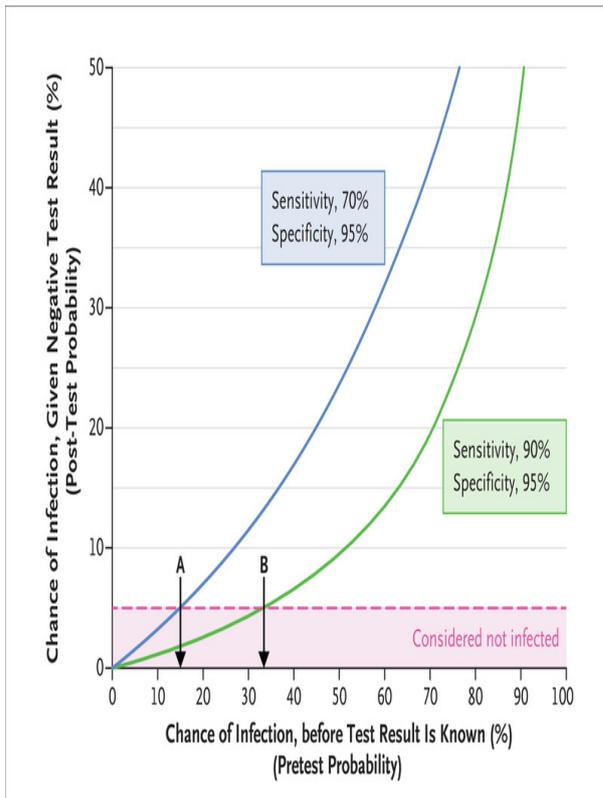
En una revisión sistemática previa a la impresión de cinco estudios (sin incluir los estudios Yang y Zhao), que incluyó a 957 pacientes ("bajo sospecha de Covid-19" o con "casos confirmados"), los falsos negativos oscilaron entre el 2 y el 29%.<sup>4</sup> Sin embargo, la certeza de la evidencia se consideró muy baja debido a la heterogeneidad de las estimaciones de sensibilidad entre los estudios, la falta de cegamiento a los resultados de las pruebas de índice en el establecimiento de diagnósticos y la falta de informe de las características clave de RT-PCR.<sup>4</sup> En conjunto, la evidencia, aunque limitada, genera preocupación sobre los resultados frecuentes falsos negativos de RT-PCR.

**Si las pruebas de diagnóstico de SARS-CoV-2 fueran perfectas, una prueba positiva significaría que alguien porta el virus y una prueba negativa que no.** Con pruebas imperfectas, un resultado negativo solo significa que una persona tiene menos probabilidades de infectarse. Para calcular la probabilidad, se puede usar el teorema de Bayes, que incorpora información sobre la persona y la precisión de la prueba (revisado recientemente<sup>5</sup>) Para una prueba negativa, hay dos entradas clave: la probabilidad previa a la prueba (una estimación, antes de la prueba, de la probabilidad de que la persona se infecte) y la sensibilidad de la prueba. La probabilidad previa a la prueba

puede depender de la prevalencia local de Covid-19, el historial de exposición al SARS-CoV-2 y los síntomas. Idealmente, la sensibilidad clínica y la especificidad de cada prueba se medirían en diversas situaciones de la vida real clínicamente relevantes (por ejemplo, fuentes de muestras variadas, tiempo y gravedad de la enfermedad).

Suponga que una prueba de RT-PCR fue perfectamente específica (siempre negativa en personas no infectadas con SARS-CoV-2) y que la probabilidad de prueba previa para alguien que, por ejemplo, se sintió enfermo después del contacto cercano con alguien con Covid-19 fue del 20%. Si la sensibilidad de la prueba fuera del 95% (el 95% de las personas infectadas dan positivo), la probabilidad de infección posterior a la prueba con una prueba negativa sería del 1%, lo que podría ser lo suficientemente bajo como para considerar a alguien no infectado y puede brindarles seguridad en las visitas altas -los parientes de riesgo. La probabilidad posterior a la prueba se mantendría por debajo del 5% incluso si la probabilidad previa a la prueba fuera tan alta como 50%, una estimación más razonable para alguien con exposición reciente y síntomas tempranos en un área de "punto caliente".

**Pero la sensibilidad para muchas pruebas disponibles parece ser sustancialmente menor:** los estudios citados anteriormente sugieren que **el 70% es probablemente una estimación razonable**. En este nivel de sensibilidad, con una probabilidad previa del 50%, la probabilidad posterior a la prueba con una prueba negativa sería del 23%, demasiado alta para asumir con seguridad que alguien no está infectado.



Probabilidad de infección por SARS-CoV-2, dado un resultado

negativo de la prueba, de acuerdo con la probabilidad previa a la prueba.

El gráfico muestra cómo la probabilidad de infección posterior a la prueba varía con la probabilidad previa a la prueba para pruebas con sensibilidad baja (70%) y alta (95%). La línea horizontal indica un umbral de probabilidad por debajo del cual sería razonable actuar como si la persona no estuviera infectada (por ejemplo, permitir que la persona visite a una abuela mayor). Donde se debe establecer este umbral, aquí, 5%, es un juicio de valor y variará con el contexto (por ejemplo, menor para las personas que visitan a un pariente de alto riesgo). El umbral destaca por qué se necesitan pruebas de diagnóstico muy sensibles. Con un resultado negativo en la prueba de baja sensibilidad, el umbral se supera cuando la probabilidad previa a la prueba excede el 15%, pero con una prueba de alta sensibilidad, se puede tener una probabilidad previa de hasta el 33% y aún así, suponiendo que el umbral del 5% , se considera seguro estar en contacto con otros.

El gráfico también destaca por qué los esfuerzos para reducir la probabilidad previa a la prueba (por ejemplo, por distanciamiento social, posiblemente usando máscaras) son importantes. Si la probabilidad previa a la prueba es demasiado alta (por encima del 50%, por ejemplo), la prueba pierde su valor porque los resultados negativos no pueden reducir la probabilidad de infección lo suficiente como para alcanzar el umbral.

**Sacamos varias conclusiones:**

**Primero**, las pruebas de diagnóstico ayudarán a abrir el país de manera segura, pero solo si las pruebas son altamente sensibles y validadas bajo condiciones realistas contra un estándar de referencia clínicamente significativo.

**En segundo lugar**, la FDA debe garantizar que los fabricantes proporcionen detalles de la sensibilidad y especificidad clínica de las pruebas en el momento de la autorización de comercialización. Las pruebas sin dicha información tendrán menos relevancia para el cuidado del paciente.

**Tercero**, medir la sensibilidad de la prueba en personas asintomáticas es una prioridad urgente. También será importante desarrollar métodos (por ejemplo, reglas de predicción) para estimar la probabilidad de infección previa a la prueba (para personas asintomáticas y sintomáticas) para permitir el cálculo de las probabilidades posteriores a la prueba después de resultados positivos o negativos.

**Cuarto**, los resultados negativos, incluso en una prueba altamente sensible, no pueden descartar infección si la probabilidad previa a la prueba es alta, por lo que los médicos no deben confiar en resultados negativos inesperados (es decir, asumir que un resultado negativo es un "falso negativo" en una persona con síntomas típicos y conocidos exposición). Es posible que realizar varias pruebas simultáneas o repetidas pueda superar la sensibilidad limitada de una prueba individual; sin embargo, tales estrategias necesitan validación.

**Finalmente**, se deben desarrollar umbrales para descartar infección para una variedad de situaciones clínicas. Dado que definir estos umbrales es un juicio de valor, el aporte público será crucial.