

COVID-19: la urgente necesidad de una adecuada interpretación de las pruebas de diagnóstico de laboratorio

Covid-19: An urgent need for an adequate interpretation of laboratory diagnostic tests

Correspondencia

Fredy Canchihuamán
fredy.canchihuaman.r@upch.pe

Recibido: 11/07/2020
Aprobado: 27/08/2020

Citar como: Canchihuaman F, Gozzer E. COVID-19: la urgente necesidad de una adecuada interpretación de las pruebas de diagnóstico de laboratorio. *Acta Med Peru.* 2021;38(1):80-4. doi: <https://doi.org/10.35663/amp.2021.381.1881>

Fredy Canchihuamán^{1,a}; Ernesto Gozzer^{1,b}

- 1 Unidad de Seguridad Sanitaria Global y Diplomacia en Salud, Facultad de Salud Pública y Administración, Universidad Peruana Cayetano Heredia. Lima, Perú.
a Médico cirujano, MPH, PhD en epidemiología; b médico cirujano, MSc

Sr. Editor:

Las pruebas diagnósticas son herramientas de ayuda esenciales para la toma de decisiones en el cuidado médico de las personas a nivel individual y en el cuidado de la salud pública a nivel colectivo o poblacional; por lo tanto, deben ser utilizadas de forma complementaria a la evaluación y razonamiento clínico y de salud pública. Asumir que los resultados de las pruebas diagnósticas son una garantía final o de certeza de lo que ocurre con la salud de una persona o grupo de personas constituye un alejamiento de los conceptos fundamentales sobre su uso y aplicación.

A continuación, realizamos una breve revisión de conceptos y una propuesta de elementos prácticos para el análisis y la adecuada interpretación de los resultados de las pruebas diagnósticas de la COVID-19 en el contexto clínico y de salud pública. En el Perú y otros países de la Región esto resulta particularmente crítico por las numerosas controversias y discusión en relación con el tipo de pruebas utilizadas para el diagnóstico y las consecuencias de las decisiones resultantes de sus usos —como ha sido en el caso de las pruebas serológicas. Asimismo por la posible implementación y escalamiento de nuevas pruebas diagnósticas, incluidas aquellas para el punto de atención como la prueba de antígenos, que ameritan analizar las consideraciones de su uso^{1,2}.

Para elegir e interpretar cualquier prueba de diagnóstico resulta esencial conocer primero la indicación de uso, segundo la calidad y tercero los valores predictivos (que dependen del contexto de la persona o grupos de personas). La indicación de uso, es decir el propósito que persigue una prueba, es el primer aspecto a considerar. En otras palabras, es importante valorar la utilidad que tiene una prueba, y por tanto la razón o el motivo de la solicitud; si es para la ayuda al diagnóstico clínico de una enfermedad actual (Ej. la presencia de una infección), el monitoreo de una enfermedad o tratamiento (Ej. la evolución de la enfermedad), la detección de una enfermedad pasada (Ej. la identificación de una infección anterior), la detección de una enfermedad en un grupo de personas que no tienen síntomas (Ej. tamizaje), entre otros posibles usos. Las pruebas moleculares y las pruebas de antígeno identifican la presencia del virus detectando su material genético o sus proteínas de membrana por lo que identifican una posible infección actual; las pruebas serológicas de otro lado detectan la respuesta inmune por lo que identifican una posible infección pasada, dado que el tiempo en que pudo ocurrir la infección no siempre es posible establecer, existen numerosas reservas sobre su uso para confirmar o descartar diagnósticos clínicos (Ej. el IgG puede aparecer antes que la IgM y la IgM puede permanecer por semanas)³. Otro aspecto a considerar es la calidad diagnóstica, es decir la capacidad o probabilidad de una prueba para identificar correctamente a personas que realmente están sanas o enfermas. Uno de los parámetros para medir la calidad de una prueba de diagnóstico es la validez (grado de certeza que se tiene sobre los resultados respecto a lo que ocurre en la realidad) que se mide con dos indicadores numéricos, la sensibilidad

Tabla 1. Probabilidades pre-prueba y post-prueba para pruebas diagnósticas en relación a valores hipotéticos de sensibilidad y especificidad

Contexto del caso	Prueba A*		Prueba B**		Prueba C***	
	Sensibilidad= 90% Especificidad= 99%		Sensibilidad= 80% Especificidad= 97%		Sensibilidad= 50% Especificidad= 90%	
	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)	Probabilidad post-prueba de COVID-19 dado un resultado de la prueba (%)
Probabilidad pre-prueba de COVID-19 (juicio clínico y/o prevalencia)	positivo	negativo	positivo	negativo	positivo	negativo
1%	47.6	0.1	21.2	0.2	4.8	0.6
5%	82.6	0.5	58.4	1.1	20.8	2.8
10%	90.9	1.1	74.8	2.2	35.7	5.8
20%	95.7	2.5	87.0	4.9	55.6	12.2
30%	97.5	4.1	92.0	8.1	68.2	19.2
40%	98.4	6.3	94.7	12.1	76.9	27.0
50%	98.9	9.2	96.4	17.1	83.3	35.7
60%	99.3	13.2	97.6	23.6	88.2	45.5
70%	99.5	19.1	98.4	32.5	92.1	56.5
80%	99.7	28.8	99.1	45.2	95.2	69.0
90%	99.9	47.6	99.6	65.0	97.8	83.3
95%	99.9	65.7	99.8	79.6	99.0	91.4
99%	100.0	90.9	100.0	95.3	99.8	98.2

Los estimados fueron realizados utilizando la herramienta web para el cálculo de probabilidades post prueba de la Universidad de California, San Francisco (UCSF) disponible en <https://sample-size.net/post-probability-calculator-test-new/>

En la estimación de las probabilidades post-prueba de la COVID-19 se redondearon los números al decimal o entero más cercano, diferencias en los cálculos pueden explicarse debido a estas aproximaciones.

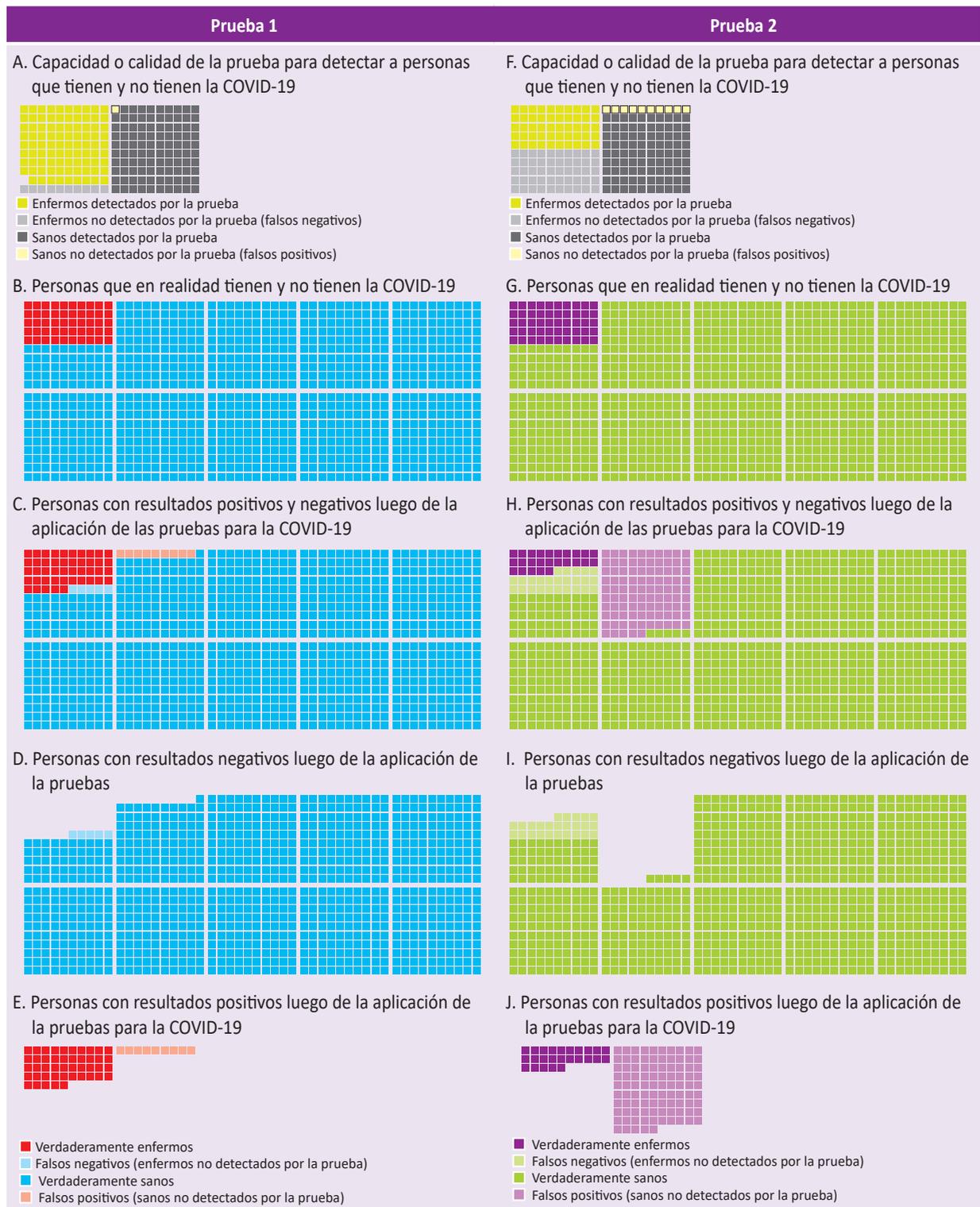
En la interpretación del resultado final (probabilidad post-prueba) influyen el contexto del caso (probabilidad pre-prueba) y la calidad (sensibilidad y especificidad) de una prueba.

En el caso de un resultado positivo de la prueba A y B (pruebas con altas especificidades), para una probabilidad pre-prueba de 5% (probabilidad baja), la probabilidad final post-prueba de COVID-19 es de 82.6% y 58.4% respectivamente. Esto es diferente en el caso de un resultado positivo de la prueba C (prueba con especificidad no tan alta), que para una probabilidad pre-prueba de 5%, la probabilidad final post-prueba de COVID-19 es de 20.8%. La realización de una prueba adicional puede disminuir la posibilidad de un falso positivo.

En el caso de un resultado negativo de la prueba A y B (pruebas con altas sensibilidades), para una probabilidad pre-prueba de 90% (probabilidad alta), la probabilidad final post-prueba de COVID-19 es de 47.6% y 65.0% respectivamente. En el caso de un resultado negativo de la prueba C (prueba con sensibilidad moderada), para una probabilidad pre-prueba de 90% (probabilidad alta), la probabilidad final post-prueba de COVID-19 es de 83.3%. La realización de una prueba adicional puede disminuir la posibilidad de un falso negativo.

y la especificidad diagnóstica. Para establecer estos valores se comparan sus resultados contra los del “estándar de oro”, una prueba que señala cuales son los verdaderos diagnósticos. En ocasiones pueden no existir un estándar de oro aceptado y las pruebas de referencia y comparación empleadas ser imperfectas. Estos valores determinan los márgenes de error, cuanto más baja la sensibilidad y especificidad diagnóstica mayor será la proporción de resultados incorrectos. Es importante advertir la posibilidad de resultados negativos en personas que tienen una enfermedad o condición (falsos negativos) y resultados positivos en personas que no la tienen (falsos positivos). En ambos casos una segunda prueba puede mejorar la calidad de los resultados, aunque esto dependerá del resultado y tipo de prueba^[4]. Una particularidad de la sensibilidad y especificidad diagnóstica de una prueba es que pueden variar dependiendo del tiempo en que esta se realiza en relación con el inicio de una enfermedad. En el caso de las pruebas moleculares o de antígenos para la COVID-19,

la sensibilidad diagnóstica es máxima durante la primera semana desde el inicio de los síntomas para luego disminuir^[1]; en el caso de las pruebas de anticuerpos, la sensibilidad diagnóstica es máxima durante la tercera semana^[5]. Existen otros parámetros diferentes y complementarios de una prueba, como la sensibilidad y la especificidad analítica, que hacen referencia junto a otros parámetros a su calidad como método en un contexto experimental de laboratorio^[6]. La sensibilidad analítica es la capacidad de una prueba para identificar cambios mínimos en cantidades o concentraciones de un elemento (patógeno, marcador biológico o sustancia) presentes en las muestras; la especificidad analítica es la capacidad para identificar correctamente a un elemento y no identificar a otros de similares características^[4,6]. Así por ejemplo, una prueba con baja especificidad analítica será aquella que detecte a un patógeno diferente al que se busca (reacción cruzada)^[6]. Al elegir e interpretar una prueba, es necesario considerar



La figura ilustra el siguiente escenario hipotético: en una población de 1000 personas, en las que el 5% tienen la COVID-19, se realizan dos pruebas de diagnóstico de laboratorio para la COVID-19. Se asume que la calidad diagnóstica de las pruebas tiene los siguientes valores:
 A. Prueba 1: Sensibilidad=90% y especificidad=99%
 B. Prueba 2: Sensibilidad=50% y especificidad=90%
 De acuerdo a la prevalencia, en total 50 personas tienen la COVID-19 y 950 no tienen la COVID-19 (B y G). Luego de la aplicación de las pruebas en esta población se obtienen los siguientes resultados:
 C y H. Personas con resultados positivos y negativos luego de la aplicación de las pruebas 1 y 2 (con diferentes valores de sensibilidad y especificidad) para la COVID-19. La prueba 1 tiene la capacidad para detectar a 45 de 50 personas que tienen la COVID-19 y 941 de 950 que no tienen la COVID-19. La prueba 2 tiene la capacidad para detectar a 25 de 50 personas que tienen la COVID-19 y 855 de 950 que no tienen la COVID-19.

D. Entre las personas con resultados negativos, luego de la aplicación de la prueba 1, la probabilidad de que esta prueba sea realmente negativa es 941 de 946, **valor predictivo negativo de 99.5%**.
 I. Entre las personas con resultados negativos, luego de la aplicación de la prueba 2, la probabilidad de que esta prueba sea realmente negativa es 855 de 880, **valor predictivo negativo de 97.2%**.
 E. Entre las personas con resultados positivos, luego de la aplicación de la prueba 1, la probabilidad de que esta prueba sea realmente positiva es 45 de 54, **valor predictivo positivo de 83.3%**.
 J. Entre las personas con resultados positivos, luego de la aplicación de la prueba 2, la probabilidad de que esta prueba sea realmente positiva es 25 de 120, **valor predictivo positivo de 20.8%**.
 La cantidad de resultados falsos positivos o negativos podrían reducirse con la aplicación de pruebas adicionales. **Pruebas realizadas de forma consecutiva (una después de otra) o simultáneas (varias en paralelo)**. Para estimar la magnitud de los resultados en una población de 1 millón de personas, estos valores se pueden multiplicar por 1,000.

Figura 1. Esquema ilustrativo de caso hipotético para una población de 1000 personas a quienes se les realiza pruebas de laboratorio

dos indicadores adicionales que dependen de la presunción diagnóstica y/o de la prevalencia de la enfermedad, los valores predictivos positivos y los valores predictivos negativos, los que indican cuál es la probabilidad que los resultados positivos o negativos obtenidos —cuando se sospecha o no se conoce si tienen una enfermedad— sean realmente correctos, es decir que una persona verdaderamente tenga o no la enfermedad.

En la práctica clínica, un método pragmático para interpretar las pruebas, que integra ambos conceptos (calidad y valores predictivos) con la historia clínica y epidemiológica, es el uso de probabilidades explícitas denominadas probabilidad pre-prueba y post-prueba^[7,8]. El método consiste en asignar una probabilidad pre-prueba de la enfermedad para una persona, “juicio clínico” inicial^[7,9], la que luego se modifica al realizar la prueba de diagnóstico (que tiene una calidad determinada), resultando en una probabilidad final post-prueba, “juicio clínico final”. En el caso de la COVID-19, se puede considerar los siguientes elementos en la estimación o selección de la probabilidad pre-prueba: 1) la prevalencia de la enfermedad o las tasas de infección nacional, regional y local en la población de donde proviene la persona 2) la historia de exposición individual a riesgos de contagio; 3) el cuadro clínico; y 4) la probabilidad de un diagnóstico diferencial^[7,8]. Una fortaleza de este método es que al utilizar datos numéricos la interpretación es más simple y explícita; una limitación es que requiere de la asignación de valores que podrían no ser conocidos y producir errores (sesgos) en las estimaciones y por consiguiente en la interpretación. Plantear escenarios hipotéticos en la realización de las pruebas diagnósticas, ilustran la importancia de estos conceptos (Tabla 1) y reflejan la magnitud de posibles resultados positivos y negativos incorrectos y la necesidad de una interpretación cautelosa (Figura 1).

En la práctica de salud pública, los profesionales de este campo pueden utilizar los mismos conceptos, esta vez a nivel poblacional, “juicio de salud pública”. En este campo, las pruebas se pueden utilizar en estudios epidemiológicos para conocer la magnitud, realizar tendencias y proyecciones de una enfermedad; para la vigilancia; el rastreo de contactos; la terminación de un periodo de aislamiento; el diseño de algoritmos; y en la implementación de programas de tamizaje e intervenciones de salud pública^[9-11]. En la decisión de la implementación de un programa de tamizaje, realizado en personas que no tienen síntomas, aunque esta es una forma apropiada de detectar tempranamente casos, los beneficios deben considerarse en relación a los riesgos potenciales para la salud, de los cuales no están exentos (como en situaciones en las que brindar resultados positivos a personas que no tienen la enfermedad, por la aplicación de un determinado tipo de pruebas, resulten en procedimientos, intervenciones o tratamientos que impliquen riesgos)^[12]. Las condiciones que justifican la realización de un tamizaje para una enfermedad han sido descritas extensamente —es un problema de salud pública, sus etapas evolutivas permiten una detección temprana, existen tratamientos o intervenciones efectivas que pueden modificar su historia natural o prevenir su propagación, existe efectividad en los tratamientos o intervenciones cuando se inician en etapas tempranas, y existen pruebas adecuadas^[12].

Si la prevalencia de una enfermedad es baja en una población y las pruebas no son suficientemente sensibles o específicas, la cantidad de falsos positivos o negativos será significativa con una serie de implicancias que se deben tener en cuenta (Figura 1).

Estos conceptos generales son aplicables para el uso de todo tipo de pruebas diagnósticas, moleculares, de antígenos, serológicas en el caso de la COVID-19 así como para pruebas realizadas en otras enfermedades (Ej. el antígeno prostático específico (PSA) para el diagnóstico de cáncer de próstata o la mamografía para el cáncer de mama). En el caso de las pruebas serológicas rápidas, debido a la indicación (detección de infección pasada), la calidad (variable, con baja sensibilidad en las dos primeras semanas del inicio de síntomas y especificidad no suficientemente alta) y los valores predictivos (bajos cuando son aplicados en personas o grupos de personas con poco riesgo) es que se ha desaconsejado su uso tanto en la práctica clínica como de salud pública con condicionadas excepciones^[5], inclusive desde el principio de la pandemia.

Para una apropiada elección e interpretación de las pruebas diagnósticas, especialmente en patologías nuevas como la COVID-19, es indispensable que el objetivo de su indicación clínica y de salud pública sean claras, que exista una información precisa sobre la calidad de las mismas, y que el contexto guíe las decisiones tanto a nivel individual como poblacional. El resultado de una prueba no define un diagnóstico si es que no se considera sus características propias y se hace una unidad clínica y epidemiológica. Finalmente, hacemos un llamado a la publicación y difusión —fácilmente accesible— sobre la calidad de las pruebas. Y a su interpretación cautelosa, utilizando criterios como los planteados previamente y distinguiendo las particularidades correspondientes a la salud individual y la salud pública o colectiva.

Contribuciones de autoría: FC y EG concibieron la idea, revisaron, redactaron y aprobaron la versión final.

Potenciales conflictos de interés: Ninguno por declarar.

Fuentes de financiamiento: Autofinanciado.

ORCID

Fredy Canchihuamán, <https://orcid.org/0000-0001-5165-2357>
Ernesto Gozzer, <https://orcid.org/0000-0002-5597-8669>

REFERENCIAS BIBLIOGRAFICAS

- Guglielmi G. Fast coronavirus tests: what they can and can't do. *Nature*. 2020;585(7826):496–8. doi:10.1038/d41586-020-02661-2
- Mina MJ, Parker R, Larremore DB. Rethinking Covid-19 Test Sensitivity — A Strategy for Containment. *New England Journal of Medicine*. 2020;383(22):e120. doi:10.1056/NEJMp2025631
- Van Elslande J, Houben E, Depypere M, Brackenier A, Desmet S, André E, et al. Diagnostic performance of 7 rapid IgG/IgM antibody tests and the Euroimmun IgA/IgG ELISA in COVID-19 patients. *Clin*

- Microbiol Infect [Internet]. 2020 [cited 2020 Jun 29]; doi:10.1016/j.cmi.2020.05.023
4. Woloshin S, Patel N, Kesselheim AS. False Negative Tests for SARS-CoV-2 Infection — Challenges and Implications. *New England Journal of Medicine*. 2020;0(0):null. doi:10.1056/NEJMp2015897
 5. Deeks JJ, Dinnes J, Takwoingi Y, Davenport C, Spijker R, Taylor-Phillips S, et al. Antibody tests for identification of current and past infection with SARS-CoV-2. *Cochrane Database of Systematic Reviews* [Internet]. 2020 [cited 2020 Oct 19];(6). doi:10.1002/14651858.CD013652
 6. Saah AJ, Hoover DR. “Sensitivity” and “specificity” reconsidered: the meaning of these terms in analytical and diagnostic settings. *Ann Intern Med*. 1997;126(1):91–4. doi:10.7326/0003-4819-126-1-199701010-00026
 7. Watson J, Whiting PF, Brush JE. Interpreting a covid-19 test result. *BMJ* [Internet]. 2020 [cited 2020 Jun 29];369. doi:10.1136/bmj.m1808
 8. Good CB, Hernandez I, Smith K. Interpreting COVID-19 Test Results: a Bayesian Approach. *J Gen Intern Med*. 2020;35(8):2490–1. doi:10.1007/s11606-020-05918-8
 9. Hanson KE, Caliendo AM, Arias CA, Englund JA, Lee MJ, Loeb M, et al. Infectious Diseases Society of America Guidelines on the Diagnosis of COVID-19. *Clin Infect Dis* [Internet]. [cited 2020 Jun 29]; doi:10.1093/cid/ciaa760
 10. Cheng MP, Papenburg J, Desjardins M, Kanjilal S, Quach C, Libman M, et al. Diagnostic Testing for Severe Acute Respiratory Syndrome–Related Coronavirus 2. *Annals of Internal Medicine*. 2020;172(11):726–34. doi:10.7326/M20-1301
 11. Mahase E. Covid-19: the problems with case counting. *BMJ* [Internet]. 2020 [cited 2020 Oct 19];370. doi:10.1136/bmj.m3374
 12. Koepsell TD, Weiss NS. *Epidemiologic Methods: Studying the Occurrence of Illness*. New York: Oxford University Press; 2003. Chapter 18, Screening; p. 442-463.