



Acerca de este número

En mayo de este año lanzamos nuestro primer número, con una muy buena recepción dentro de la comunidad de la Escuela de Salud Pública, y también fuera de ella. Hoy, a poco más de tres meses, con gusto les presentamos el segundo número de *Inferencias*, el que trae tres artículos que les contaremos brevemente a continuación.

En el primer artículo, José Antonio Ruiz-Tagle, graduado en 2020 del Magíster en Bioestadística, nos cuenta sobre su experiencia personal y académica durante su paso por el programa, especialmente durante la realización de su tesis de grado. En una combinación que él mismo denomina como extraña, además de bioestadístico, Ruiz-Tagle es Cientista Político y actualmente está realizando un doctorado en esa misma disciplina. Quizás esa mezcla *sui generis* hace que su relato sea más interesante, en el que confiesa momentos duros mezclados con satisfacciones, donde lo que más releva es la mirada crítica que adquirió para ver los errores comunes en los análisis estadísticos encontrados en la literatura. En palabras de Ruiz-Tagle, “Aprender Estadística se asemeja a introducirse en el arte de la magia. Tiene la virtud de impresionar y dar un gran espectáculo, pero la gran desventaja es que mientras más trucos se aprenden, más se pierde la capacidad de asombro”.

El segundo artículo es del Doctor en Salud Pública y hasta el año 2020 académico de nuestra Escuela, Yuri Carvajal, quien actualmente es Jefe de la Unidad de Epidemiología del Hospital Carlos van Buren en Valparaíso. El Dr. Carvajal nos narra su experiencia en el cargo durante la pandemia de COVID-19, en donde mezcla en un mismo relato su visión y crítica al modelo de salud pública, mirado desde dentro de un hospital, y su quehacer específico reportando indicadores sanitarios y donde da un ejemplo de los comandos usados en Stata para lograrlo.

Finalmente, en el tercer artículo el profesor Mauricio Fuentes del Programa de Bioestadística nos presenta de una manera pedagógica los conceptos de sensibilidad, especificidad y valores predictivos de los exámenes diagnósticos, enmarcado en la pandemia de COVID-19. También nos entrega una aproximación a la relación entre positividad y prevalencia de casos, donde dicha relación cambiaría para distintos escenarios. El profesor Fuentes concluye con una reflexión acerca de la positividad y la forma en que ésta se reporta en nuestro país.

No queremos dejar pasar la ocasión para informar que nuestro Magíster en Bioestadística ha abierto su período de postulación para ingresar en marzo de 2022. Este período de postulación estará abierto hasta el 19 de noviembre de este año y se puede acceder a toda la información del proceso en el sitio web de la Escuela de Salud Pública (<http://www.saludpublica.uchile.cl/postgrado>).

Esperamos que este nuevo número sea de su interés y que disfruten su lectura, y desde ya invitamos a académicas, académicos, docentes y profesionales a colaborar con nuestro boletín y enviarnos sus escritos. Sabemos que para la gran mayoría no es sencillo hacerse un tiempo para escribir, pero trabajaremos para que el esfuerzo valga la pena.

¿Qué se aprende en el Magíster en Bioestadística?

JOSÉ ANTONIO RUIZ-TAGLE MATURANA^{1,2}

Desde el inicio del proceso de tesis, pasando por la graduación hasta hoy, me he dado el tiempo de reflexionar sobre qué aprendí en el programa. En este ensayo les contaré mi experiencia, la cual transitó por un camino muy pedregoso y con más bajos que altos. Quizás pueda serle útil a alguna persona que esté pasando por lo mismo.

Ingresé al Magíster inmediatamente después de terminar la carrera de Ciencia Política. Estoy consciente de que es una combinación extraña, sin embargo, durante el pregrado me acerqué a los “métodos cuantitativos” más que a cualquier otra asignatura. Las clases consistían en sentarnos frente a un computador con SPSS a analizar datos. Aprendí muchísimas cosas, pero me di cuenta de que no entendía lo que estaba haciendo, era muy hábil maniobrando el computador para tener el resultado, pero llegué a un momento en donde tenía tantas dudas, y a la vez tan pocas herramientas para contestarlas, que me vi en la necesidad de entrar al Magíster en Bioestadística. Además, mis intereses chocaban con los fenómenos que comúnmente se investigaban en la Ciencia Política, soy un renegado al igual que muchas personas. El primer año fue muy duro, y temía a causa de mi poco entendimiento matemático. Logré pasar el curso “Bases matemáticas para la Estadística” con un 4, teniendo que invertir en clases particulares para poder aprobarlo. En los cursos aplicados me iba relativamente bien, pero los matemáticos tuvieron una complejidad imposible de abordar desde mi posición. Me ausenté durante muchas clases porque estaba deprimido, sentía que asistir a clases era perder el tiempo. Sin embargo, nunca me quedaba de brazos cruzados, estudiaba de fuentes que me resultarían más inteligibles, trabajaba con bases de datos, simulaba comprobaciones de la teoría, etc. A pesar de que las matemáticas son una herramienta muy necesaria, personalmente no me conformaba con saber la demostración de un teorema, necesitaba entender qué significaba y por qué era tan importante. De alguna manera, percibí que los profesores estaban conscientes de eso y me tendieron la mano cuando lo necesité. Mis compañeros también me ayudaban mucho, y estaré eternamente agradecido por eso. Afortunadamente, no me ha tocado encontrar ningún estimador máximo verosímil o alguna función generadora de momentos.

Durante el programa fallé en todo lo que pude fallar: reprobé los hitos de integración 1 y 2, cambié mi

proyecto de tesis y tuve que alargar un semestre para poder terminar mi investigación. Las ganas de abandonar nunca me faltaron, sin embargo, el incentivo más fuerte que tenía para no rendirme era que de mi generación empezamos nueve y solo quedábamos cuatro. En el último semestre que me quedaba para terminar mi tesis, empezaron las clases de Doctorado en Políticas Públicas, las cuales eran en modalidad telemática. A partir de ese momento, comencé a reflexionar sobre todo lo que había aprendido en el Magíster en Bioestadística.

Resulta interesante pensar que, a pesar de que hay muchas maneras de aproximarse a un fenómeno, no puede quedar fuera de la ecuación el “¿qué me gustaría aprender?”. Particularmente, el programa de Magíster en Bioestadística no se trata sólo de abordar un fenómeno utilizando ciertos métodos, es estar al tanto de que existe una gama de posibilidades que podrían ser viables, por lo que debemos justificar cuidadosamente nuestras elecciones. A raíz de esto, se vuelve necesario darse un tiempo para pensar en cómo haríamos el estudio si tuviésemos recursos ilimitados. Por ejemplo, mi tesis buscaba estudiar la relación entre el consumo de sustancias con las conductas sexuales de riesgo. Utilicé como indicador el número de parejas sexuales en el último año. Para hacerlo utilicé datos del Instituto Nacional de la Juventud (INJUV), y apliqué un modelo de regresión binomial negativo generalizado (NB-P). El artículo resultante de este proceso aún se encuentra en proceso de edición. El diseño de muestreo de la encuesta no era óptimo para cumplir con los objetivos de mi investigación, porque correspondía a una muestra especialmente diseñada para estimar a través de intervalos de confianza. El problema es que ese tipo de diseños muestrales no considera dos elementos esenciales para obtener los anhelados p valores: potencia estadística y tamaño del efecto. Por otro lado, las encuestas rara vez preguntan lo que uno exactamente necesita, y nos tendremos que conformar con acercarnos lo más posible al fenómeno con la información disponible. Evidentemente, tener recursos hace posible recolectar nuestra propia muestra, donde las estimaciones tendrán una precisión deseada y establecida *a priori*, las mediciones serán más apegadas al fenómeno de interés y podríamos evitar algunos sesgos. Al estar consciente de esto, no pude evitar preguntarme por qué había tantos estudios que hacían lo mismo que yo. Esto me resultaba preocupante porque parecía que los

¹Magíster en Bioestadística (Universidad de Chile).

²Candidato a Doctor en Políticas Públicas (Universidad Mayor).

artículos científicos no respetaban al pie de la letra lo que dictaminaba la teoría estadística. A raíz de esto, pude concluir dos cosas: 1) es mejor estar consciente del error y no ignorarlo, y 2) mientras más cosas sean controladas en el diseño, más simples son los métodos para el análisis.

Elegí el modelo de regresión binomial negativa porque la primera vez que vi una variable que tenía una desviación estándar mayor a la media me llamó profundamente la atención. En ese momento no sabía que uno de los supuestos de la distribución de Poisson era la igualdad entre la media y la varianza. De alguna manera, los datos me estaban diciendo que había un comportamiento heterogéneo en torno a esa variable. Me pareció interesante, yo pensaba que era un problema, y que simplemente se debía a personas mal intencionadas que buscaban sesgar los resultados. Después entendí que no es sencillo esclarecer las causas de esta heterogeneidad, y que frente a este escenario, es mejor considerarlo una característica. Para poder aplicar el método, tuve que dedicar muchas horas de estudio a los modelos de tipo conteo, teniendo que responder preguntas como ¿por qué quiero usar una binomial negativa y no una Poisson?, ¿cómo puedo distinguirlas? Yo sabía por tacto y experiencia que la variable dependiente seguía esa distribución, pero debía demostrarlo con evidencia. Cuando llegué a la binomial negativa, debía buscar la parametrización correcta, y esto no fue tarea fácil. Particularmente, la regresión binomial negativa tiene más de 20 parametrizaciones diferentes, y cada una sirve para algo específico. Si no hubiera investigado esto, podría haber corrido el comando `nbreg` en Stata y se acababa el problema. No obstante, ese comando modela por defecto una parametrización NB2, es decir, la varianza es una función cuadrática de la media. Al hacer eso, me di cuenta de que la sobredispersión no se solucionaba, por ende, tuve que estimar el exponente utilizando una NB-P. Todo esto mientras me preguntaba si las investigaciones que utilizaban este tipo de métodos hicieron lo mismo que yo, ¿estarán seguros de que usaron la parametrización correcta?

Aprender Estadística se asemeja a introducirse en el arte de la magia. Tiene la virtud de impresionar y dar un gran espectáculo, pero la gran desventaja es que

mientras más trucos se aprenden, más se pierde la capacidad de asombro. Después de pasar por el programa, veía errores en todos lados, ¿por qué hace una regresión lineal con esos datos?, ¿por qué está hablando de que la muestra es representativa si quiere hacer comparaciones?, ¿para qué ponen la media y la desviación estándar de una variable binaria en la tabla de descriptivos? En contraste, también me he preguntado ¿por qué le añaden tanta complejidad a algo que puede ser tan simple? Ahora que soy candidato a Doctor en Políticas Públicas y Magíster en Bioestadística, sigo sin entender fórmulas con integrales dobles y sigo sin saber cuándo es el mejor momento para añadir una constante a ambos lados de la ecuación. Sin embargo, entiendo que todas las investigaciones tienen fortalezas y debilidades, y soy capaz de reconocer la diferencia entre error y limitación.

La solución a todos nuestros problemas está en hacer un buen diseño, idealmente desde la extracción de la muestra. Si esto no es posible, debemos estar alerta de todos los errores no forzados que tuvimos que cometer en el trayecto. Porque finalmente los métodos estadísticos deberán ser confecciones de sastrería para llevar a cabo nuestro diseño de investigación. Estoy consciente de que puede ser difícil considerando el arduo proceso que conlleva conseguir una publicación. Desde hace un tiempo que trabajo haciendo asesorías estadísticas. Mi labor consiste en ayudar a que otras personas se gradúen o publiquen. Perdí la cuenta de las veces que me han pedido analizar datos para ver si encuentro algo interesante, ojalá un p valor que muestre significancia. Ser Magíster en Bioestadística me permite tomarme la molestia de investigar métodos que sean adecuados para dar cuenta de los posibles errores que estoy cometiendo, haciendo de esto un proceso reflexivo y no automático.

Información de la tesis:

Título: “Relación entre el consumo de alcohol y otras drogas con el número de parejas sexuales en el último año en jóvenes residentes en Chile de 15 a 29 años”

Año: 2020

Autor: José Antonio Ruiz-Tagle Maturana

Profesor guía: Dr. Dante Cáceres Lillo, Programa de Salud Ambiental, Escuela de Salud Pública

1. Tres olvidos

Hacer epidemiología en un hospital es un trabajo anfibio. Puedo respirar con pulmones de médico y bucear números con piel húmeda de bioestadístico. Tratar de ayudar a que las acciones del hospital sigan a la realidad, aunque sea su pálida traza en cifras. Este marzo ese gran trabajo se volvió una experiencia intensa, permanente y, como tal, llena de aprendizajes.

La reforma de salud del 2005 hizo que intentar hacer epidemiología dentro de un hospital sea una intención anacrónica y desubicada. A partir de la radical separación entre salud pública y establecimientos, se retiró la “inteligencia sanitaria” a las dependencias de la Subsecretaría de Salud Pública y las Secretarías Regionales Ministeriales (SEREMIs), dejando a los hospitales sólo como establecimientos autogestionados en red, una palabra elegante para llevarlos a la privatización.

COVID con toda su desgracia ha venido a recordarnos que un sector público de salud vigoroso es indispensable y éticamente justo. Que los afanes privatizadores son sueños tecnocráticos e ignorantes.

Pero también COVID nos ha enseñado que los hospitales requieren epidemiología y que son productores de valiosas cifras epidemiológicas. No es casual que la unidad en la que trabajo (somos dos personas) haya nacido en plena pandemia, en octubre recién pasado.

Hay tres fuertes razones conceptuales tras esta marginalidad de la epidemiología en la salud chilena y en los establecimientos de salud. La primera es que, de acuerdo con el informe “Invertir en salud” del Banco Mundial de 1993, libro canónico para el Ministerio de Salud (MINSAL) desde esa fecha, debíamos olvidarnos en los hospitales de hacer epidemiología. A lo más deberíamos calcular AVISA para organizar el sistema de salud. Todo el resto lo haría la oferta y la demanda. Por supuesto estos apurados salubristas vueltos economistas nunca leyeron a la Sra. Joan Robinson, y no se enteraron de que el único lugar en que la demanda intersecta a la oferta al precio de equilibrio es en la pizarra de una cátedra de microeconomía.

La segunda, es que desde los 60 se nos viene diciendo que olvidemos las enfermedades transmisibles, que son materia superada. Pese al cólera, a la Tubercu-

losis (TBC), al Virus de la Inmunodeficiencia Humana (VIH), a la resistencia antibiótica microbiana. La preocupación por las bacterias y los virus parecía una cosa digna de los historiadores de la medicina. Por supuesto nadie leía a Paul de Kruif y los llamados de Paul Farmer, la emergencia y reemergencia de enfermedades parecían inquietud sólo de una medicina tropical. Una rápida ojeada a los clásicos de la epidemiología, el Rothmann y el Zklo, deja un sabor amargo. No tienen una línea de enfermedades transmisibles. Simplemente para ellos las transmisibles no son epidemiología.

Y el tercer agente de olvido, fue la aparición de una figura numérica más atractiva que las viejas tasas de ataque, las R_0 y las encuestas de campo. La llegada de los cuadros de mando integral copó el mundo de las cifras. No es extraño hoy que el DEIS aparezca como un aparato obsoleto (y sí, tiene rasgos epistemológicos que perturban), que los Compromisos de Gestión (COMGES) aún hoy, en medio de esta marea avasalladora, sigan siendo reverenciados como los indicadores y que haya debido travestirse a EPIVIGILA (sistema on-line de notificación de enfermedades transmisibles) en un registro informático de gestión de las crisis, volviéndose el sistema de uso más intensivo en hospitales y que el Registro Nacional de Inmunizaciones sea tan crucial como para que cualquier ciudadano requiera estar allí para moverse de una región a otra.

2. Construidos, situados, rodeados de incertidumbre y para la acción: así son los números

Para el sentido común las estadísticas están dadas o sólo se limitan a reflejar hechos dados, son universales, duras, exactas e independientes. La pandemia ha venido a confirmarnos lo equivocado de ese enfoque y a confirmar lo que varios investigadores de la historia de la Estadística nos han enseñado:

- Los números son contruidos, procesados por redes laboriosas y extensas. Sin esas redes múltiples, es imposible tener un dato. Los números son parte de la construcción de un mundo común, son parte de la democracia y el alfabetismo estadístico es un aspecto crucial de la política.
- Los números jamás pierden las marcas de su origen, por más que se vuelvan móviles inmutables, siem-

pre tienen sabor y color local, pues se han producido para responder a problemas locales. Los números no sólo poseen localidad sino también historicidad. Mas que un panóptico, lo que vemos es casi siempre un oligóptico.

- Toda medida debe expresar su incertidumbre. Es la regla básica de la metrología. Números que aparecen desnudos como una flor en el desierto, sin referencias y marcas de sus limitaciones, son flores frágiles.
- Los números dependen de las técnicas. Si al principio todo era PCR, poco a poco nos interesaron los censos serológicos, los test rápidos, las CNAF, protagonistas todos de equipamientos y abordajes que hasta el 2020 eran absolutamente impensadas estadísticamente.

3. Los números que hemos usado

La introducción de variables en el tiempo ha sido la principal dimensión. La semana epidemiológica se ha quedado corta, hay que decirlo. La modelación de los datos mediante suavizadores sencillos como medias móviles ha permitido entender fenómenos de máximos y mínimos y los puntos de inflexión, nodos cruciales para definir nuestras acciones. El análisis de situación se ha vuelto un punto central del trabajo hospitalario y en eso, las cifras en serie de tiempo juegan un rol predominante. La actualidad de los datos es clave, y esto significa un desafío cuando los productores de esos números por lo general no tienen entrenamiento estadístico. La administración de esos datos en media hora o menos pasa a ser pieza fundamental, para ponerlos en regla con las duras exigencias de los procesadores. Sin un buen software trabajado en modo *batch* hacer eso es imposible.

Los indicadores con lo que hemos trabajado en el mes de marzo han sido:

- Positividad local de PCR (*Polymerase Chain Reaction*): serie de tiempo desde marzo de 2020.
- Consultas totales, respiratorias, sospechas de COVID y hospitalizaciones en los tres hospitales y el consolidado de la atención primaria de Valparaíso y San Antonio desde marzo de 2020.
- Ocupación de camas críticas, básicas y en la Urgencia de Adultos del Hospital desde diciembre de 2020.
- Ocupación de equipos de soporte ventilatorio en el Hospital desde enero de 2020.
- Cifras comunales de casos sintomáticos desde marzo de 2020 (tomadas del *Github* del Ministerio de Ciencias).

- Positividad a COVID de las mujeres que acuden por trabajo de parto a nuestro Hospital.
- Mortalidad COVID + y - en nuestro Hospital.

Estos números se han transformado en un reporte dos veces al día, am y pm, distribuido en forma de archivo pdf.

He abusado de Stata y de \LaTeX . Comandos más aplicados: `tsset`, `tsmooth`, `ma`, `tabout` (para sacar a \LaTeX) y he programado las capturas desde Excel, usando como parámetros el `cellrange` y las `sheet`. Para hacer informes regulares, he usado `beamer` de \LaTeX y a la fecha le he agregado la hora `\currenttime`. Los `beamer` funcionan muy bien en celulares. También he abusado de hacer un `summarize` de las fechas y armar una `local` con su valor máximo, para poner al pie del gráfico el valor actual y la fecha de ese momento.

El comando `import delimited "https://raw.githubusercontent.com/MinCiencia/Datos-COVID19/master/output/producto15/FechaInicio Sintomas_std.csv", clear charset("latin1")` permite descargar automáticamente desde el sitio del Ministerio de Ciencias la base de casos comunales semanales por fecha de inicio de síntomas, por ejemplo, y empezar a trabajar con la base de datos.

Adjunto mis dos joyitas: la serie de tiempo con la positividad poblacional (Figura 1) y el uso de Ventilación Mecánica (VM) y Cánula Nasal de Alto Flujo dentro del Hospital (CNAF) (Figura 2). Un enigma epidemiológico es que el máximo uso de CNAF fue el 10 de abril y al día siguiente registramos la mayor positividad poblacional de PCR.

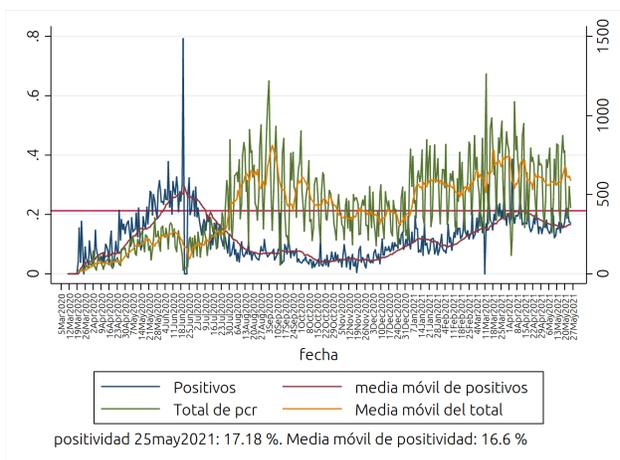


Figura 1: Positividad de PCR a nivel de SSVSA.

4. Lecciones

Todas las enfermedades, con mayor énfasis las infecciosas, no sólo han cambiado la medicina sino el

mundo. Me atrevo a presagiar que las estadísticas se volverán más democráticas y accesibles, que el uso de software se tornará más extenso y que los programadores proliferarán. Pero también sueño con que las estadísticas ambientales continúen lo que hemos hecho con los datos de salud: que sepamos usarlas en tiempo real para abordar una enfermedad severa y con ellas movernos en un movimiento colectivo coherente y ordenado a un planeta sustentable.

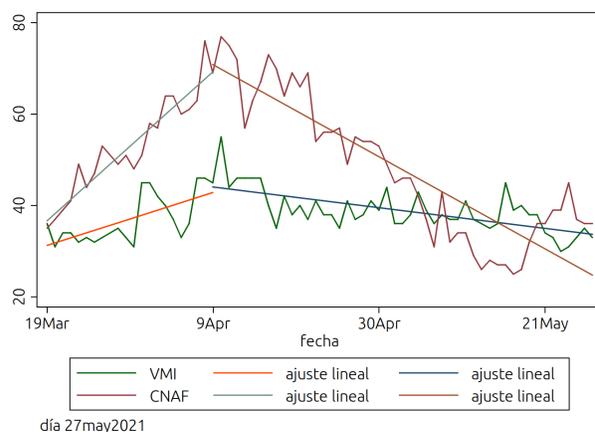


Figura 2: Evolución del uso de CNAF y VM.

Sobre los exámenes diagnósticos y la positividad en pandemia

MAURICIO FUENTES ALBURQUENQUE
PROGRAMA DE BIOESTADÍSTICA, ESCUELA DE SALUD PÚBLICA
FACULTAD DE MEDICINA, UNIVERSIDAD DE CHILE

Es aceptado que el test RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa) es el mejor método disponible para la detección del virus SARS-CoV-2 [1, 2, 3, 4]. Aunque sea considerado como *gold standard* o referencia, el test RT-PCR, como toda prueba diagnóstica, no es perfecto y también arroja resultados incorrectos. Según se ha podido determinar, el mejor desempeño se obtiene durante la primera semana de aparición de los síntomas, con una sensibilidad del orden de 85-90% y una especificidad de 99% o más. Los diagnósticos incorrectos se deben en gran medida a la etapa de la infección en la que se toma el examen, problemas en la muestra biológica o errores de procedimiento del personal técnico [1, 5, 6].

En este artículo intentaré explicar los aspectos centrales del desempeño de un test diagnóstico: sensibilidad, especificidad y los valores predictivos. También incluyo algunos cálculos y una que otra reflexión sobre la positividad como indicador para evaluar el estado de la pandemia de COVID-19. Para efectos de lo que pretendo exponer me referiré siempre al test RT-PCR, ya que es el principal examen usado para confirmar casos y con él se construyen los indicadores de positividad [7, 8, 9, 10, 11].

1. Lo sensible y lo específico del test

La *sensibilidad* es la probabilidad de que una persona contagiada dé un resultado positivo. Dicho de otra forma, si la sensibilidad del test RT-PCR es de 90% (o 0,9 si lo expresamos estrictamente como probabilidad) quiere decir que, de 100 personas contagiadas, en 90 de ellas el resultado será positivo. Esto implica que, a su vez, 10 de ellas obtendrán un resultado negativo. Estos 10 diagnósticos incorrectos se conocen como *falsos negativos*.

Por otro lado, la *especificidad* es la probabilidad de que una persona libre del virus dé un resultado negativo. Para el test RT-PCR, una especificidad de 99% significa que, al tomarlo a 100 personas sanas, sólo para una de ellas dará positivo, diagnóstico fallido conocido como *falso positivo*.

Según lo anterior, podemos decir que el test es más *sensible* mientras mejor es su capacidad de diagnosticar correctamente a las personas infectadas, y es más *específico* cuanto mejor es para diagnosticar correctamente a las personas no infectadas. El problema es que ambas cosas no varían siempre en el mismo sentido, existiendo exámenes muy sensibles y muy específicos al mismo tiempo, muy sensibles y poco específicos, o muy específicos y poco sensibles.

2. ¿Cuánto creerle al test? Valores predictivos y algo sobre la positividad

Para entender qué tan bien funciona un test diagnóstico en la práctica no basta con conocer sus indicadores de calidad “de fábrica”, como son la sensibilidad y especificidad. También debemos saber qué tan confiables son sus resultados cuando se aplican en una población o grupo determinado, y para esto usamos los valores predictivos.

El *valor predictivo positivo* (VPP) es la probabilidad de que una persona cuyo resultado fue positivo esté realmente contagiada. Nuevamente podemos facilitar la explicación pensando que, de 100 personas con resultado positivo, cuántas de ellas están realmente con el virus. Naturalmente, mientras más alto este indicador (más cercano a 1 o a 100 %) más confiable es el resultado positivo para indicar o “predecir” (de ahí su nombre) el verdadero estado de la persona.

El *valor predictivo negativo* (VPN), por su parte, es la probabilidad de que una persona con resultado negativo del examen esté realmente libre del virus. Es decir, de 100 personas que dieron negativo, cuántas de ellas de verdad no están contagiadas.

Para efectos de simplificar el siguiente ejercicio, vamos a suponer que el examen RT-PCR se hace siempre en las mismas condiciones (por ejemplo, durante la primera semana de los síntomas) y que tiene una sensibilidad de 90 % y una especificidad de 99 %. Quiero aclarar desde ya que los valores que usaré de prevalencia, aunque basados en cierta lógica, son arbitrarios al fin y al cabo.

Supongamos que hacemos una campaña de testeo en la comunidad y examinamos a 1000 personas, de las cuales 100 están contagiadas (con esto estamos suponiendo una prevalencia de casos de 10 %). De acuerdo a la sensibilidad y especificidad del test, los resultados positivos y negativos se distribuirán como se muestra en el diagrama de la Figura 1.a): 90 positivos y 10 negativos en las personas contagiadas, 9 positivos y 891 negativos en las personas sin el virus.

Si sumamos todos los positivos, éstos son 99 en total, de los cuales 90 corresponden a personas contagiadas. Esto quiere decir que el 91 % (90 de 99) de las personas que dan positivo están realmente contagiadas, esto es, 91 contagiadas por cada 100 resultados positivos. Ahora, si hacemos un cálculo similar con los negativos, éstos son en total 901, de los cuales 891 están efectivamente libres del virus. Esto representa un 99 %, es decir, 99 realmente libres del virus de cada 100 resultados negativos. Es claro que en este escenario sería

mucho más confiable un resultado negativo del test que uno positivo, ya que el primero tiene un 99 % de probabilidad de ser cierto (VPN) y el segundo sólo un 91 % (VPP). Notar que la positividad en este grupo es del 10 % (99 resultados positivos en las 1000 personas examinadas), coincidente con la prevalencia de casos.

Pensemos ahora en un grupo de 1000 pacientes que consultan por síntomas asociados a COVID-19. Supongamos que 800 de estos pacientes están contagiados (prevalencia de 80 %), y que al examinarlos con RT-PCR los resultados positivos y negativos se distribuyen como se muestra en la Figura 1.b). Haciendo cálculos similares obtenemos un VPP de 99,7 % (722 contagiados de 722 positivos) y un VPN de 71 % (198 libres del virus de 278 negativos). En este caso la credibilidad del resultado negativo caería drásticamente, indicando que casi un 30 % de los resultados negativos son pacientes que sí están contagiados, es decir, falsos negativos. Además, la positividad en este grupo es del 72 % (722 resultados positivos de 1000 exámenes), más baja que la prevalencia real.

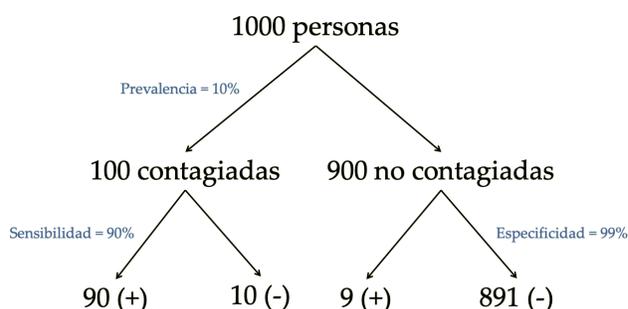


Figura 1.a): Resultados de testear a mil personas con 10 % de prevalencia de la infección, considerando una sensibilidad de 90 % y una especificidad de 99 %.

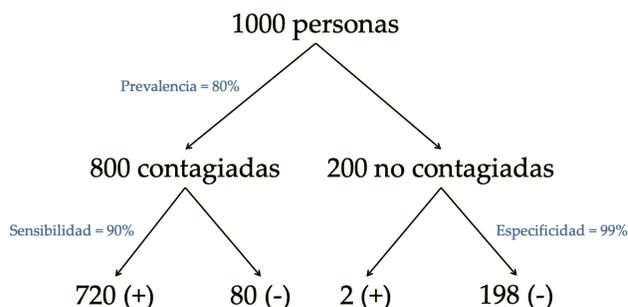


Figura 1.b): Resultados de testear a mil personas con 80 % de prevalencia de la infección, considerando una sensibilidad de 90 % y una especificidad de 99 %.

Si ahora juntamos ambos grupos en uno solo de 2000 personas, tenemos que la prevalencia de casos en este total es de 45 % (900 personas contagiadas de las 2000). Si sumamos los resultados positivos y negativos y hacemos los cálculos respectivos, tenemos un VPP de

99 % (de los 821 positivos, 810 son personas contagiadas) y un VPN de 92 % (1089 libres del virus de 1179 resultados negativos). La positividad es de 41 % (821 resultados positivos de 2000 exámenes).

De lo anterior podemos hacer algunas observaciones importantes:

i) Los valores predictivos cambian según la prevalencia en la población

El desempeño y la capacidad predictiva de los resultados positivos y negativos varían según el porcentaje de personas infectadas en el grupo testeado. En particular, para una prevalencia relativamente baja el VPN es superior al VPP, es decir, los resultados negativos son más confiables que los positivos. En esta situación de baja prevalencia la falla en los diagnósticos tiene como consecuencia que aumentan los falsos positivos, es decir, habrá más personas no infectadas diagnosticadas como contagiadas, las que deberán realizar cuarentenas y tomar las medidas establecidas para los casos de COVID-19.

En contraste, cuando la prevalencia de casos es alta, como podría darse en medio de la pandemia en las personas que consultan por síntomas, los resultados positivos son mucho más probables de ser ciertos que los negativos. La consecuencia de esto para un virus altamente contagioso puede ser mucho más peligrosa que en la primera situación, dado que aumenta el número de falsos negativos, es decir, personas diagnosticadas como no contagiadas que quedarían sin ninguna medida y podrían seguir haciendo sus actividades habituales y contagiando a otras.

Del Teorema de Bayes tenemos que los valores predictivos positivo y negativo están dados por

$$VPP = \frac{S \cdot Prev}{S \cdot Prev + (1 - E) \cdot (1 - Prev)} \quad (1)$$

$$VPN = \frac{E \cdot (1 - Prev)}{E \cdot (1 - Prev) + (1 - S) \cdot Prev} \quad (2)$$

donde S : sensibilidad del test; E : especificidad del test; $Prev$: prevalencia de casos en el grupo de personas testeadas.

En la Figura 2 se muestra la relación entre la prevalencia y los valores predictivos, considerando siempre una sensibilidad de 90 % y una especificidad de 99 %. Por ejemplo, para una prevalencia de 50 % el VPN desciende hasta 90 %, lo que significa que 10 de cada 100 resultados negativos corresponden a personas infectadas.

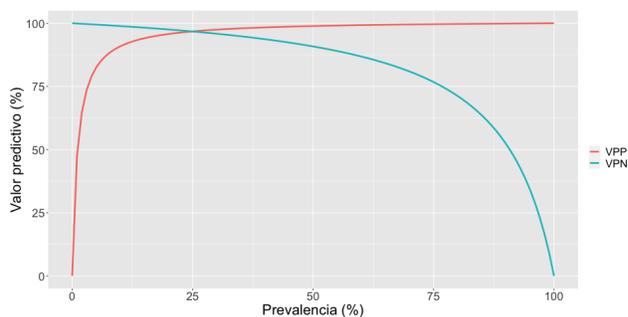


Figura 2: Relación entre la prevalencia del virus y los valores predictivos positivo (VPP) y negativo (VPN) del test, considerando una sensibilidad de 90 % y una especificidad de 99 %.

ii) Positividad y prevalencia

Cuando la prevalencia es baja la positividad del test se aproxima bien a dicha prevalencia, mientras que cuando es alta la positividad subestima la prevalencia real. Esto es coherente con el resultado anterior y debiera llevarnos a ser cautos con la interpretación que se hace de la positividad.

También podemos tener una idea de cómo es esta relación, ya que la positividad equivale a la probabilidad de un resultado positivo, dada por

$$P(+) = S \cdot Prev + (1 - E) \cdot (1 - Prev) \quad (3)$$

$$= [S - (1 - E)] \cdot Prev + (1 - E)$$

En la Figura 3 vemos dicha relación que, como indica la última expresión de la ecuación 3, es una recta cuya pendiente es la diferencia entre la sensibilidad (S) y la proporción de falsos positivos ($1 - E$). Se muestra además con una línea discontinua el caso en que la prevalencia es igual a la positividad, que sería el caso ideal si el test no fallara nunca. Vemos que la recta continua indica que la positividad es menor a la prevalencia, con una diferencia que va en aumento a medida que la prevalencia es mayor. Esto sugiere que el indicador de positividad estaría ocultando o distorsionando lo que está ocurriendo en la población estudiada, y esa distorsión aumenta proporcionalmente al aumento del propio indicador.

iii) ¿Positividad total o separada?

Cuando testeamos dos grupos con prevalencias muy distintas, éstas se combinan y pueden llevar a resultados que no representan ni a una ni a la otra. En el ejemplo mostrado ocurre esto cuando consideramos el total de 2000 pacientes, donde la prevalencia total es 45 % y se obtiene una posibilidad total de 41 %. Dichos grupos son iguales en tamaño y por lo tanto el

total se divide en un 50 % con prevalencia baja y 50 % con prevalencia alta. Sin embargo, si estos porcentajes varían, la positividad puede estar influenciada por el grupo con mayor presencia y distar aún más de la prevalencia real.

En Julio de 2020 se implementó la Estrategia Nacional de Testeo, Trazabilidad y Aislamiento (TTA), y una de sus acciones es la Búsqueda Activa de Casos (BAC), la que tiene como objetivo la detección de casos de COVID-19 que el sistema de salud no ha pesquisado por consulta espontánea [12, 13, 14, 15]. Un indicador reportado desde entonces es el porcentaje de tests RT-PCR por BAC respecto al total de tests RT-PCR realizados, tanto por comuna como por región [8, 9]. Este porcentaje ha ido aumentando en el tiempo, llegando a alcanzar el 100 % en algunas comunas.

Si por búsqueda activa se toma un porcentaje importante del total de exámenes, y la positividad en este grupo es muy pequeña, entonces la positividad total disminuye. Un aspecto importante para reflexionar entonces es si la positividad total es un indicador adecuado, dado que podemos estar mezclando dos poblaciones distintas en términos de su prevalencia de casos.

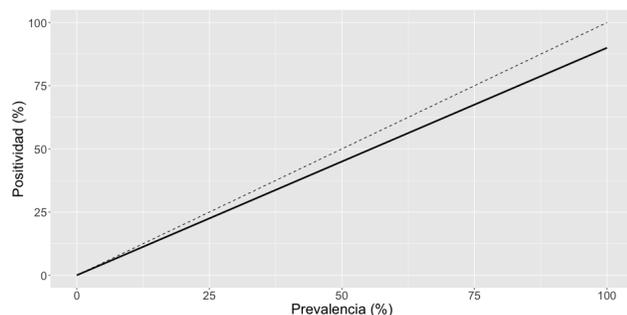


Figura 3: Relación entre la prevalencia de casos y la positividad del test, considerando una sensibilidad de 90 % y una especificidad de 99 %. La línea discontinua indica la recta de identidad, que corresponde al caso ideal en que la prevalencia coincide con la positividad (un test que no falla nunca).

3. Comentarios finales

He intentado mostrar y explicar los indicadores de desempeño de los exámenes diagnósticos, esto es, sensibilidad, especificidad y los valores predictivos, usando ejemplos sencillos y refiriéndome al test RT-PCR usado para detectar el virus SARS-CoV-2.

Una observación importante es que el resultado del test diagnóstico usado será más o menos confiable no sólo según sus características propias de calidad (sen-

sibilidad y especificidad), sino también de acuerdo a la población o grupo de personas a la que se aplique. En términos generales, cuando la prevalencia de casos es baja debemos cuestionarnos la interpretación de los resultados positivos. Sin embargo, cuando la prevalencia es alta (por ejemplo, en pacientes que consultan por síntomas de COVID-19) aumentan los falsos negativos, con el consecuente peligro del subreporte de un virus tan contagioso como éste, al dejar sin medidas a personas infectadas que pueden propagar más la enfermedad.

Por otro lado, el indicador de positividad usado para la toma de decisiones de control epidemiológico (junto a otros indicadores) tiene ventajas incluso por sobre el número de casos, ya que éstos dependen del número de exámenes practicados. Si los exámenes se aplicaran a una muestra aleatoria de la población y se corrigiera por la sensibilidad y especificidad del test, la positividad sería una buena estimación de la prevalencia de casos. Sin embargo, esto no ocurre ya que es impracticable aplicar exámenes a una muestra aleatoria. Entonces, cabe preguntarse si la positividad debería ser tratada con mayor cautela para saber a qué grupos de la población representa.

La búsqueda activa de casos (BAC) es una muy buena y necesaria estrategia, pero combinar la positividad obtenida a través de la BAC con la positividad de las personas que consultan puede estar llevando a un indicador que no está entregando la información que se necesita, ya que se pueden estar combinando grupos con prevalencias muy distintas. Si fuese así, quizás sea mejor reportar dos positivities separadas: una para las personas que consultan en los centros de salud y otra para la comunidad obtenida a través de la búsqueda activa de casos, siempre resguardando el seguimiento y aislamiento de los casos positivos y sus contactos.

Agradecimientos

Quiero agradecer a mis queridas colegas Sandra Flores del Programa de Bioestadística, por la revisión que hizo a la apurada y desordenada primera versión de este manuscrito, y Tania Alfaro del Programa de Epidemiología, por la revisión al documento ya más avanzado. La mirada crítica, observaciones y comentarios de ambas fueron esenciales para corregir y “podar” el documento, además de evitar aberraciones en el lenguaje y ayudar a poner los énfasis donde era necesario.

Referencias

- [1] L. S. Langa, L. V. Sallent, and S. R. Díez, “Interpretación de las pruebas diagnósticas de la COVID-19,” *FMC : Formación médica continuada en atención primaria*, vol. 28, pp. 167–173, mar 2021.
- [2] J. Dinnes, J. J. Deeks, S. Berhane, M. Taylor, A. Adriano, C. Davenport, S. Dittrich, D. Empeador, Y. Takwoingi, J. Cunningham, S. Beese, J. Domen, J. Dretzke, L. Ferrante di Ruffano, I. M. Harris, M. J. Price, S. Taylor-Phillips, L. Hooft, M. M. Leeflang, M. D. McInnes, R. Spijker, A. Van den Bruel, and Cochrane COVID-19 Diagnostic Test Accuracy Group, “Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection,” *Cochrane Database of Systematic Reviews*, no. 3, 2021.
- [3] N. N. Y. Tsang, H. C. So, K. Y. Ng, B. J. Cowling, G. M. Leung, and D. K. M. Ip, “Diagnostic performance of different sampling approaches for SARS-CoV-2 RT-PCR testing: a systematic review and meta-analysis,” *The Lancet Infectious Diseases*, jul 2021.
- [4] B. D. Kevadiya, J. Machhi, J. Herskovitz, M. D. Oleynikov, W. R. Blomberg, N. Bajwa, D. Soni, S. Das, M. Hasan, M. Patel, A. M. Senan, S. Gorantla, J. McMillan, B. Edagwa, R. Eisenberg, C. B. Gurumurthy, S. P. M. Reid, C. Punyadeera, G. Chang, Linda, and H. E., “Diagnostics for SARS-CoV-2 infections,” *Nature Materials*, vol. 20, no. 5, pp. 593–605, 2021.
- [5] N. Sethuraman, S. S. Jeremiah, and A. Ryo, “Interpreting Diagnostic Tests for SARS-CoV-2,” *JAMA*, vol. 323, pp. 2249–2251, jun 2020.
- [6] Ministerio de Sanidad de España and Instituto de Salud Carlos III, “Estrategia de detección precoz, vigilancia y control de COVID-19,” 2021.
- [7] Ministerio de Salud, “Actualización de la definición de caso sospechoso, probable y confirmado para vigilancia epidemiológica ante pandemia de COVID-19 en Chile,” 2021.
- [8] C. e. I. Ministerio de Ciencias, Tecnología, “MinCiencia/Datos-COVID19,” 2021.
- [9] Ministerio de Salud, “Informes epidemiológicos COVID-19,” 2021.
- [10] Ministerio de Salud, “Informes de indicadores Estrategia Nacional de Testeo, Trazabilidad y Aislamiento,” 2021.
- [11] “iCOVID Chile,” 2021.
- [12] Ministerio de Salud, “Noticias regionales: Equipos de Salud de Atención Primaria avanzan en Búsqueda Activa de Casos,” 2021.
- [13] Ministerio de Salud, “Destacados: Subsecretaria Daza entrega balance de Búsqueda Activa para Residencias Sanitarias,” 2020.
- [14] Ministerio de Salud, “Departamento de Epidemiología. Manual operativo para la búsqueda activa de casos en la comunidad y diagnóstico precoz COVID-19,” 2020.
- [15] Ministerio de Salud, “Departamento de Epidemiología. Protocolo de coordinación para acciones de vigilancia epidemiológica durante la pandemia COVID-19 en Chile: Estrategia Nacional de Testeo, Trazabilidad y Aislamiento,” 2020.