



theconversation.com

“Para entender una epidemia hay que tener la capacidad de describirla numéricamente.”

Ignacio de Blas

Los números detrás de la pandemia



En esta ocasión vamos a dejar atrás las palabras para centrarnos en las cifras porque para entender una epidemia hay que tener la capacidad de describirla numéricamente y para ello tendremos que definir una serie de medidas de enfermedad.

Pero antes de cuantificar el número de enfermos en una población necesitamos una herramienta que nos discrimine los individuos enfermos de los individuos sanos. Me estoy refiriendo a las pruebas diagnósticas. Un pequeño comentario antes de empezar a evaluar la fiabilidad del diagnóstico: tenemos que tener claro lo que estamos diagnosticando: enfermedad, infección o inmunidad frente al patógeno. El diagnóstico clínico permite diferenciar sanos y enfermos, pero en muchas ocasiones tiene que ser complementado con pruebas

de diagnóstico directo para identificar la presencia del patógeno que causa la enfermedad. En este grupo de pruebas directas se encuentran la PCR, las pruebas de antígenos y el cultivo microbiológico. Y por último hay un tercer grupo de pruebas indirectas que intentan determinar si el individuo ha estado en contacto con el patógeno en el pasado buscando elementos de la respuesta inmune frente a ese patógeno, fundamentalmente anticuerpos usando pruebas serológicas como ELISA y seroneutralización.

Pero lo que tienen en común todas estas pruebas diagnósticas es que son imperfectas y por eso decimos que un individuo es positivo a la prueba en lugar de decir que está infectado, o que es negativo para indicar que no está infectado. Puede parecer algo confuso porque la mayoría de las personas piensan que es lo mismo, y no es así.

Imaginemos que existiera una prueba diagnóstica perfecta a la que llamaremos prueba de oro (*gold standard*) con la que pudiéramos comparar nuestras pruebas diagnósticas. En ese caso podríamos aplicar ambas pruebas a un conjunto de individuos y obtendríamos cuatro posibilidades que vamos a mostrar en una tabla de 2x2.

Lo deseable es que solo hubiera **verdaderos positivos** (individuos enfermos que dan un resultado positivo con la prueba que estamos evaluando) y **verdaderos negativos** (individuos sanos con resultado negativo a la prueba evaluada). Eso solo ocurre con la prueba de oro que hemos imaginado en nuestra cabeza (aunque en muchos casos se utiliza la histopatología y la PCR como pruebas de oro asumiendo que son perfectas, sin serlo).

En la práctica hay varias alternativas a esa prueba de oro, la más sencilla sería seleccionar un grupo de individuos que tengamos plena seguridad de que están sanos (por ejemplo, procedentes de un país donde nunca se haya detectado la enfermedad) a los que tomaremos

muestras para ver cuántos son negativos (suponemos que la mayoría) y si hay alguno positivo. Esos individuos sanos con diagnóstico positivo es lo que denominamos **falsos positivos** y se pueden producir por muy distintos motivos (infección con un patógeno similar al que estamos estudiando, reacciones inmunitarias cruzadas con anticuerpos frente a otro patógeno, contaminaciones de la muestra, errores en el procesado, etc.).

Una vez estudiado este grupo de individuos sanos podemos plantearnos infectarlos experimentalmente. Lógicamente hay unas restricciones éticas que nos impiden hacerlo cuando trabajamos con seres humanos, pero en el caso de los veterinarios es viable, pero siempre con la supervisión de una comisión ética. Es decir, que nos aseguramos que los animales están realmente infectados y/o enfermos (según lo que estemos

“Especificidad (E) se define como la probabilidad de detectar correctamente a un individuo sano.”

		PRUEBA DE ORO	
		Enfermos (positivos)	Sanos (negativos)
Prueba a evaluar	Positivo	Verdaderos positivos (a)	Falsos positivos (b)
	Negativo	Falsos negativos (c)	Verdaderos negativos (d)

$$1) \quad S = P(\text{positivo}|\text{enfermo}) = \frac{P(\text{positivo, enfermo})}{P(\text{enfermo})} = \frac{a/N}{(a+c)/N} = \frac{a}{a+c} = \frac{\text{verdaderos positivos}}{\text{enfermos}}$$

$$2) \quad E = P(\text{negativo}|\text{sano}) = \frac{P(\text{negativo, sano})}{P(\text{sano})} = \frac{d/N}{(b+d)/N} = \frac{d}{b+d} = \frac{\text{verdaderos negativos}}{\text{sanos}}$$

diagnosticando). Ahora esperamos encontrar que los individuos sean positivos (verdaderos positivos), pero en la práctica algunos individuos tienen diagnóstico negativo (los **falsos negativos**) que pueden deberse a problemas con el tipo de muestra utilizada, mala conservación de la muestra, fallos durante el análisis laboratorio o toma de la muestra en un momento inadecuado (un ejemplo clásico ocurre con el diagnóstico de anticuerpos ya que tardan 7-10 días en aparecer tras la infección y si tomamos la muestra antes de que pase ese tiempo no se pueden detectar porque no están presentes todavía).

A partir de esta información podemos calcular dos indicadores de fiabilidad de las pruebas diagnósticas: sensibilidad y especificidad. Empezamos a usar los números.

La **sensibilidad (S)** es la probabilidad de detectar correctamente a un individuo enfermo y podemos expresarlo también como la probabilidad de que un verdadero po-

sitivo padezca la enfermedad o como la probabilidad de obtener un diagnóstico positivo cuando el individuo está enfermo. Estamos hablando de una probabilidad condicional que se expresaría como P(positivo|enfermo), y que tras simplificar corresponderá al cociente entre el número de verdaderos positivos (enfermos detectados por la prueba) y el número total de enfermos reales (ecuación 1).

Es evidente que una prueba muy sensible dará pocos falsos negativos y por tanto tendrá utilidad para descartar la presencia de enfermedad, ya que ante un resultado negativo podremos estar bastante seguros de que se tratará de un individuo sano.

El segundo indicador es la **especificidad (E)** que se define como la probabilidad de detectar correctamente a un individuo sano, y que podemos expresar también como la probabilidad de que un verdadero negativo esté sano o como la probabilidad de obtener un diagnóstico negativo cuando la enfermedad esté ausente, lo que en términos de probabilidad condicional equivale a P(negativo|sano) y se calcula como el cociente entre el número de verdaderos negativos (sanos detectados por la prueba) y el número total de sanos reales (ecuación 2).

Lógicamente una prueba muy específica dará pocos falsos positivos y por tanto tendrá utilidad para confirmar la presencia de enfermedad, ya que gran parte de los positivos serán individuos realmente enfermos.

Para terminar de entenderlo pondré un sencillo ejemplo para detectar infectados de SARS-CoV-2 con una muestra nasofaríngea utilizando una PCR (ver tabla).

		PRUEBA DE ORO		
		Infectados	No infectados	TOTAL
Prueba a evaluar (PCR)	Positivo	990 Verdaderos positivos	200 Falsos positivos	1.190 Positivos
	Negativo	10 Falsos negativos	3.800 Verdaderos negativos	3.810 Negativos
	TOTAL	1.000 infectados	4.000 no infectados	5.000

Aplicando las fórmulas anteriores veremos que la sensibilidad es 99% (=990/1.000) y la especificidad es 95% (=3.800/4.000). Pueden parecer valores muy altos (y lo son) e indicarían que la prueba es muy fiable, pero a pesar de ello en la muestra analizada hay 10 individuos que estaban infectados y que hemos diagnosticado como negativos (falsos negativos). Puede que no parezcan demasiados, pero son personas infectadas a los que les decimos que no tienen el virus y que vuel-

ven a la calle donde pueden infectar a otras personas. Para evitar ese problema con la sensibilidad, a los contactos estrechos con un positivo se les obliga a mantenerse en cuarentena a pesar de que ellos tengan un diagnóstico negativo. Cuando estamos rastreando en busca de infectados es muy importante que la prueba sea muy sensible para minimizar el número de individuos infectados que sean negativos y que no seríamos capaces de detectar.



“La sensibilidad (S) es la probabilidad de detectar correctamente a un individuo enfermo.”

El otro problema es que hay 200 personas a las que les hemos dicho que están infectadas sin ser cierto (falsos positivos). En este caso la consecuencia es que vamos a ponerles en aislamiento durante 14 días sin ser realmente necesario y además vamos a investigar a sus contactos estrechos (que también se mantendrán confinados en sus casas). Es el precio que hay que pagar por usar una prueba inespecífica, pero más vale prevenir. En este caso no es un precio demasiado alto, pero en los planes de erradicación en sanidad animal un resultado positivo conlleva el sacrificio del animal positivo o incluso de toda la explotación ganadera, así que las pruebas tienen que ser muy específicas para evitar sacrificar animales que realmente están sanos.

Por lo tanto, cuando medimos la cantidad de enfermos en una población lo que realmente estamos haciendo es estimar la cantidad de positivos, y a esa proporción de positivos en la población la denominamos **prevalencia aparente**, que es la probabilidad de diagnosticar como positivo a un individuo de la población y se calcula dividiendo número total de individuos positivos por el número total de individuos. En nuestro ejemplo sería 23,8% (=1.190/5.000).

$$P_{\text{aparente}} = P(\text{positivo}) = \frac{a+b}{N} = \frac{\text{positivos}}{\text{población total}}$$

Sin embargo, lo que realmente querríamos conocer es la **prevalencia real o verdadera** que mide la frecuencia de aparición de una enfermedad en una población (la

probabilidad de estar enfermo) y se calcularía dividiendo el número total de enfermos por total de individuos. En nuestro caso es 20% (=1.000/5.000).

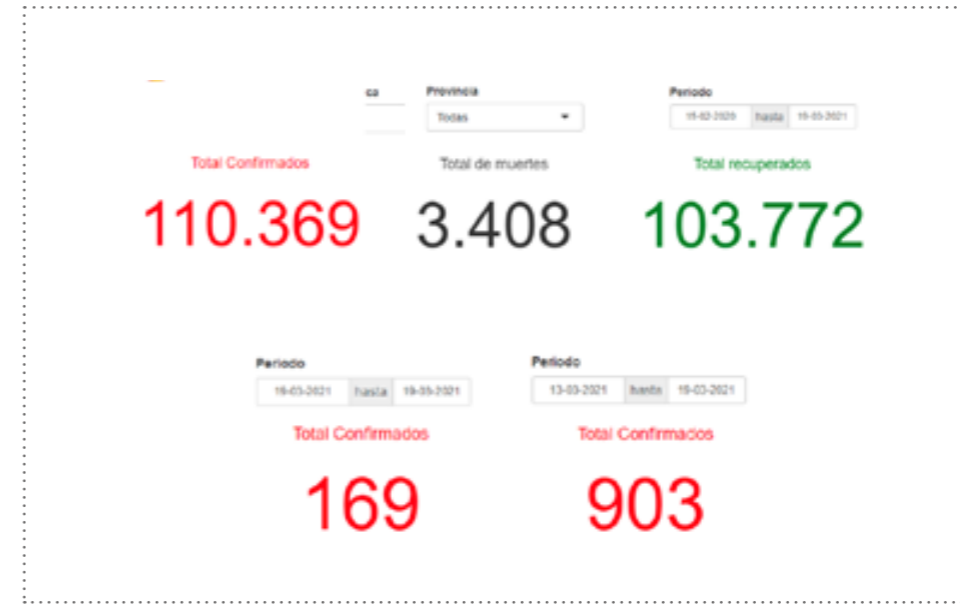
$$P_{\text{real}} = P(\text{enfermo}) = \frac{a+c}{N} = \frac{\text{enfermos}}{\text{población total}}$$

Vemos que al calcular la prevalencia aparente estamos considerando que hay más enfermedad de la que realmente hay por eso es importante estimar la prevalencia real en función de la prevalencia aparente, la sensibilidad y la especificidad utilizando la siguiente fórmula que nos permite comparar la proporción de enfermos entre poblaciones que han sido diagnosticadas con pruebas diagnósticas con distinta fiabilidad.

$$P_{\text{real}} = \frac{P_{\text{aparente}} + \text{Especificidad} - 1}{\text{Sensibilidad} + \text{Especificidad} - 1}$$

Pensad por un momento en un mismo grupo de personas que es diagnosticada con dos pruebas diferentes; por ejemplo, la PCR de sensibilidad de 99% y especificidad de 95%, y una prueba de antígenos con sensibilidad del 97% y especificidad del 96%. Con la PCR la prevalencia aparente era 23,8%, sin embargo, con la prueba de antígenos sería del 20,2%. ¿Cuál se equivoca? Ambas, y la solución es calcular la prevalencia real teniendo en cuenta la fórmula anterior. Os invito a que hagáis los cálculos con los datos de la tabla.

		PRUEBA DE ORO		
		Infectados	No infectados	TOTAL
Prueba a evaluar (PCR)	Positivo	990	200	1.190
	Negativo	10	3.800	3.810
	TOTAL	1.000	4.000	5.000
Antígenos	Positivo	850	160	1.010
	Negativo	150	3.840	3.990



Mapa de casos de COVID-19 en Aragón.

Hemos hablado de la prevalencia como probabilidad de que un individuo esté enfermo, infectado o positivo (según lo que estemos diagnosticando) en un momento dado (prevalencia puntual) o a lo largo de un periodo de tiempo (prevalencia acumulada).

Tomemos como referencia los datos de la covid-19 en Aragón del día 19 de marzo de 2021 (disponibles en esta dirección: <https://datacovid.salud.aragon.es/covid>). Hasta ese día en Aragón se han confirmado 110.369 casos (positivos), de los cuales se han recuperado 103.772 y han fallecido 3.408 personas.

Pero si buscamos sólo los datos de ese día comprobamos que ese día se han producido 169 nuevos casos y en la última semana han sido un total de 903.

Para saber la **prevalencia puntual** del día 19 de marzo necesitamos saber los casos activos en ese momento, que será el total de casos acumulados menos los que han fallecido y los que se han recuperado, es decir que ese día hay 3.189 casos activos (=110.369 - 3.408 - 103.772), y para poderlo comparar con otras regiones tenemos que convertir esta frecuencia absoluta en una frecuencia relativa dividiendo por el total de la población de Aragón (1.311.212 habitantes). El resultado es una proporción de casos igual a 0,243%, y como es un valor muy pequeño lo vamos a expresar como casos por 100.000 habitantes de manera que el 19 de marzo había 243 casos activos por 100.000 habitantes.

Esa es una medida transversal, pero también nos interesa ver cuál ha sido la evolución en un periodo de tiempo; por ejemplo, desde que empezó la epidemia en marzo de 2020. En ese caso utilizaremos la **prevalencia acumulada** que es una medida longitudinal de la enfermedad, y dividiremos el total de casos que se han producido por la población de Aragón, y vemos que el 8,42% (=110.369/1.311.212) de la población aragonesa ha sido diagnosticada como caso desde que se inició la pandemia (aunque es muy probable que sean muchos más debido a la existencia de los falsos negativos y a

“A los epidemiólogos nos gusta trabajar con las incidencias acumuladas semanales porque nos indican la velocidad de propagación de la enfermedad y cómo varía a lo largo del tiempo.”



“A partir de estos números podemos valorar la eficacia de las medidas de prevención, el impacto de la enfermedad y la evolución de la epidemia.”

que algunos han cursado la infección de forma asintomática y no se les realizó ninguna prueba diagnóstica).

Hay otra medida de la enfermedad muy importante para saber la evolución de la enfermedad a lo largo de tiempo y poder comparar la velocidad de aparición de nuevos casos. Se trata de la **incidencia** que se calcula dividiendo el número de nuevos casos que se producen en un periodo de tiempo por la población en riesgo al inicio de ese periodo. Se puede calcular la **incidencia diaria**, pero es más habitual calcular la **incidencia acumulada** cada 7 días (IA7) o cada 14 días (IA14) para “compensar” las variaciones diarias.

Volviendo a nuestro ejemplo vemos que la incidencia diaria sería de 0,0129% (=169/1.311.212) o 12,9 nuevos casos diarios por cada 100.000 habitantes, mientras que la incidencia acumulada a 7 días (IA7) sería de 0,0689% (=903/1.311.212) o 68,9 nuevos casos semanales por cada 100.000 habitantes.

Para finalizar este breve repaso de epidemiología descriptiva, hay dos indicadores de gran interés relacionados con la probabilidad de morir por la enfermedad. El primero es la **mortalidad** definida como la probabilidad de que un individuo de la población muera a consecuencia de la enfermedad en un momento o periodo de tiempo. Se calcula como el cociente entre el número de muertes causadas por la enfermedad en el periodo de tiempo estudiado y la población en riesgo de morir por esa enfermedad en ese mismo periodo (los casos).

$$\text{Mortalidad} = \frac{\text{nº de muertos}}{\text{Población en riesgo}}$$

Por tanto, la mortalidad acumulada en Aragón por covid-19 será del 0,26% (=3.408/1.311.212).

La otra medida es la **letalidad** y es la probabilidad de que un caso muera en un periodo de tiempo determinado.

La calcularemos dividiendo el número de muertos por la enfermedad y el número de individuos enfermos en un periodo de tiempo establecido. Aunque también se podría calcular dividiendo la mortalidad por la prevalencia (a la que en ocasiones también se le denomina morbilidad).

$$\text{Letalidad} = \frac{\text{nº de muertos}}{\text{nº de casos}} = \frac{\text{Mortalidad}}{\text{Prevalencia}}$$

En nuestro ejemplo la letalidad es del 3,09% (=3.408/110.369), lo que indicaría que si alguien se infecta con SARS-CoV-2 tendría más de un 3% de posibilidades de morir por culpa de la covid-19. En la práctica esa probabilidad depende de la edad, y en personas mayores de 70 años la probabilidad de morir cuando se infectan puede ser incluso superior al 20%.

A los epidemiólogos nos gusta trabajar con las incidencias acumuladas semanales porque nos indican la velocidad de propagación de la enfermedad y cómo varía a lo largo del tiempo, lo que nos permite saber si estamos en la fase de progresión o de regresión de la curva epidémica. Y también nos resulta muy interesante conocer la letalidad ya que podemos pronosticar la evolución de la enfermedad en los individuos infectados.

Hay muchos más números con los que trabajar en una epidemia, pero los principales son los que hemos visto, y a partir de ellos podemos valorar la eficacia de las medidas de prevención, el impacto de la enfermedad y la evolución de la epidemia.

Ignacio de Blas
Dpto. de Patología Animal
Facultad de Veterinaria
Universidad de Zaragoza