

1) Suponed que tenemos una enfermedad X para la que disponemos de una prueba de referencia (positivo \Leftrightarrow enfermo), y queremos estudiar la bondad de un nuevo test diagnóstico (menos intrusivo, más barato,... lo que queráis).

a) Suponed que tomamos una muestra de 100 enfermos y una muestra de 200 sanos: de los primeros, 95 dan positivo en la nueva prueba y de los segundos, 6.

a. A partir de estos datos, ¿podemos estimar la sensibilidad de la nueva prueba? ¿Su especificidad? ¿Su VPP? ¿Su VPN? ¿La prevalencia de la enfermedad? ¿La tasa de positividad de la nueva prueba?

b. Estimad los valores que se puedan estimar.

La composición de la muestra en enfermos y sanos es artificial, la que nosotros hemos querido, y por lo tanto en principio no es representativa de la composición en enfermos y sanos de la población. Por lo tanto, a partir de la muestra no podemos estimar la probabilidad de estar enfermo, o sano, condicionada a nada: no se pueden estimar el VPP, el VPN ni la prevalencia.

Sí se pueden estimar la sensibilidad y la especificidad, porque suponemos que las muestras de enfermos y sanos son representativas de las poblaciones de enfermos y sanos, y por tanto las proporciones de positivos y negativos en ellas son representativas de las probabilidades de positivos y negativos en sus subpoblaciones.

- Sensibilidad: probabilidad de que un enfermo dé positivo: $95/100 = 0.95$

- Especificidad: probabilidad de que un sano dé negativo: $194/200 = 0.97$

¿Y la tasa de positividad? Como la composición en enfermos y sanos es artificial y la probabilidad de dar positivo depende de si se está enfermo o sano, la proporción de positivos en esta muestra no es representativa de la población, y por lo tanto tampoco se puede usar para estimar la proporción de positivos en la población.

b) Suponed que tomamos una muestra de 100 positivos de la nueva prueba y una muestra de 200 negativos: de los primeros, 95 están enfermos y de los segundos, 6.

a. A partir de estos datos, ¿podemos estimar la sensibilidad de la nueva prueba? ¿Su especificidad? ¿Su VPP? ¿Su VPN? ¿La prevalencia de la enfermedad? ¿La tasa de positividad de la nueva prueba?

b. Estimad los valores que se puedan estimar.

Mutatis mutandis.

La composición de la muestra en positivos y negativos es artificial, la que nosotros hemos querido, y por lo tanto en principio no es representativa de la composición en positivos y negativos de la población. Por lo tanto, a partir de la muestra no podemos estimar la probabilidad de dar positivo, o negativo, condicionada a nada: no se pueden estimar la sensibilidad, la especificidad ni la tasa de positividad.

Sí se pueden estimar los valores predictivos, porque suponemos que las muestras de positivos y negativos son representativas de las poblaciones de positivos y negativos, y por tanto las proporciones de enfermos y sanos en ellas son representativas de las probabilidades de enfermos y sanos en sus subpoblaciones.

- VPP: probabilidad de que un positivo esté enfermo: $95/100 = 0.95$

- VPN: probabilidad de que un negativo esté sano: $194/200 = 0.97$

¿Y la prevalencia? Como la composición en positivos y negativos es artificial y la probabilidad de estar enfermo varía según se dé positivo o negativo en la prueba, la proporción de enfermos en esta muestra no es representativa de la población, y por lo tanto tampoco se puede usar para estimar la proporción de enfermos en la población.

2) Una historia real. El jefe del servicio de urgencias de un hospital observó un número excesivo de falsos diagnósticos de apendicitis aguda (AA) en su servicio. Muchos enfermos operados de urgencia de AA resultaba que tenían el apéndice normal. Para evaluar la posibilidad de imponer que se realizase una ecografía a todos los enfermos con

síntomas de AA, llevó a cabo el estudio siguiente. Se realizó una ecografía a los primeros 200 pacientes que llegaron al servicio de urgencias con síntomas de AA, y un experto en diagnóstico por imagen usó las ecografías para diagnosticar si sufrían o no AA. Mientras, se siguió aplicando el protocolo de realizar una apendicectomía de urgencia en caso de diagnóstico clínico de AA. Tanto los médicos que atendieron cada caso como el experto en ultrasonidos (siempre el mismo) desconocían el diagnóstico de la otra parte.

Las dos tablas siguientes recogen los resultados tras el análisis de los 200 casos (donde US+ significa que el experto en ultrasonidos diagnosticó AA y US- que no; MU+ significa que el médico de urgencias que atendió el caso diagnosticó AA y MU- que no; AA significa que se le realizó la apendicetomía y se descubrió que tenía AA, y no AA significa que o bien no se le realizó la apendicetomía, o si se le realizó, se descubrió que no tenía AA; vaya, no AA significa que no tuvo AA).

	AA	no AA		AA	no AA
US+	26	2	MU+	41	116
US-	15	157	MU-	0	43

a) ¿Por qué creéis que se escondió tanto al experto en ultrasonidos como a los médicos de urgencias los diagnósticos de la otra parte?

Para que no se copiaran. Vaaale, para que el diagnóstico de cada uno fuera independiente del diagnóstico del otro.

b) ¿Cuáles son la población objetivo y la población muestreada?

La población objetivo, la que realmente le interesa al jefe del servicio de urgencias, son los pacientes que llegan a su servicio de urgencias con síntomas de AA. En este caso, la población muestreada es la misma que la objetivo.

Comentario: Pero bueno, tampoco me he enfadado si como población objetivo tomáis la personas con síntomas de AA de todo el mundo mundial, pero el médico que plantea el estudio solo está interesado en los pacientes con síntomas de AA de su servicio de urgencias.

En todo caso, está mal tomar como población objetivo todo el mundo.

c) ¿Cómo fue la muestra tomada?

Transversal de conveniencia.

Comentario: Sujeto “la muestra tomada”, no “la muestra”.

d) Por la manera como se obtuvieron los datos, ¿se pueden estimar la especificidad y la sensibilidad del diagnóstico de AA tanto por parte del experto en ultrasonidos como por parte de los médicos de urgencias? ¿Y sus VPP y VPN? Justificad brevemente vuestra respuesta.

La muestra es transversal, por lo que esperamos que sea representativa de la población de pacientes con síntomas de AA en ese servicio de urgencias. Por lo tanto, podemos usarla para estimarlo todo.

e) Independientemente de la respuesta correcta a la pregunta anterior, calculad la especificidad, la sensibilidad y los VPP y VPN del diagnóstico de AA por parte del experto en ultrasonidos. Dad los detalles de los cálculos y explicad cada valor obtenido qué significa (para este caso concreto, no en general).

- Especificidad: $157/159 = 0.987$. Un 98.7% de los pacientes con síntomas de AA que acuden a ese servicio de urgencias y que no tienen AA, son diagnosticados sin AA por el experto en ultrasonidos.
- Sensibilidad: $26/41 = 0.634$. Un 63.4% de los pacientes con síntomas de AA que acuden a ese servicio de urgencias y que tienen realmente AA, son diagnosticados con AA por el experto en ultrasonidos.
- VPP: $26/28 = 0.929$. Un 92.9% de los pacientes con síntomas de AA que acuden a ese servicio de urgencias y que son diagnosticados con AA por el experto en ultrasonidos, tienen realmente AA.
- VPN: $157/172 = 0.913$. Un 91.3% de los pacientes con síntomas de AA que acuden a ese servicio de urgencias y que son diagnosticados sin AA por el experto en ultrasonidos, no tienen AA.

Comentario: No hace ninguna falta usar la fórmula de Bayes para calcular los VPP y VPN, se pueden calcular directamente de la tabla.

f) Independientemente de la respuesta correcta a (4), calculad ahora la especificidad, la sensibilidad y los VPP y VPN del diagnóstico de AA por parte de los médicos de urgencias en la muestra. Dad los detalles de los cálculos, no hace falta que expliquéis qué significa cada valor.

- Especificidad: $43/159 = 0.270$

- Sensibilidad: $41/41 = 1$
- VPP: $41/157 = 0.261$
- VPN: $43/43 = 1$

g) De los 200 sujetos estudiados, 41 tuvieron AA. Este 41/200, ¿de qué exactamente estima la prevalencia?

De la AA entre los pacientes que acuden con síntomas de AA a este servicio de urgencias.

Comentario: No “estiman” nada de la muestra. De la muestra no tenemos que estimar ningún valor, lo calculamos y ya está. Usamos la muestra para “estimar” cosas de la población objetivo.

h) En las urgencias pediátricas de ese hospital se observó que un 80 % de los niños que ingresaban con síntomas de AA en realidad sí que la tenían. Suponiendo que la especificidad y la sensibilidad del diagnóstico tanto por parte del experto en ultrasonidos como por parte de los médicos del servicio fueran las mismas en este contexto que en el de urgencias generales, ¿cuáles serían los valores predictivos positivo y negativo de ambos?

Para hacer un ejemplo de cada, para el experto en ultrasonidos voy a usar una tabla y para el médico de urgencias voy a usar la fórmula de Bayes.

Empecemos con el US. Supongamos que llegan 1000 niños con síntomas de AA. Como la prevalencia de la AA entre ellos es del 80 %, esperamos que 800 tengan AA y 200, no. Como la sensibilidad de 0.634, de los 800 que tiene apendicitis, $800 \cdot 0.634 = 507$ (en números redondos) serán diagnosticados (correctamente) que tienen apendicitis y el resto ($800 - 507 = 293$) serán diagnosticados (erróneamente) como que no.

Con una especificidad de 0.987, de los 200 que no tiene AA, serán diagnosticados (correctamente) como libres de AA $200 \cdot 0.987 = 197$ y los 3 restantes (erróneamente) con AA. El resultado es la tabla siguiente:

	AA	no AA	Total
US+	507	3	510
US-	293	197	490
Total	800	200	1000

De aquí obtenemos, para las ecografías, $VPP = 507/510 = 0.994$ y $VPN = 197/490 = 0.402$.

Por lo que refiere al MU:

$$VPP = \frac{P(MU + |AA)P(AA)}{P(MU + |AA)P(AA) + P(MU + |No AA)P(No AA)} = \frac{1 \cdot 0.8}{1 \cdot 0.8 + (1 - 0.27) \cdot 0.2} = 0.846$$

$$VPN = \frac{P(MU - |No AA)P(No AA)}{P(MU - |No AA)P(No AA) + P(MU - |AA)P(AA)} = \frac{0.27 \cdot 0.2}{0.27 \cdot 0.2 + 0 \cdot 0.8} = 1$$

i) A la vista de los resultados obtenidos, ¿creéis que valía la pena contratar el experto en ultrasonidos para complementar el diagnóstico clínico de AA en el servicio de urgencias general? (Pensad que el dinero que se destine a este contrato y a realizar las ecografías no se podrá destinar a otros conceptos, así que sólo valdrá la pena si el beneficio para los pacientes es claro.) En caso afirmativo, ¿cómo recomendaríais que se complementasen ambos diagnósticos?

El objetivo es evitar operaciones, por su mayor coste económico para el hospital y el coste humano para el paciente y sus familiares.

Si el MU dice que el paciente no tiene AA, es muy probable que no la tenga ($VPN = 1$). Pero si el MU dice que el paciente sí que tiene AA, en realidad solo tiene una probabilidad del 26.1 % de tenerla: 3 de cada 4 diagnósticos de AA del MU son falsos positivos. En cambio, si el US dice que el paciente tiene AA, tiene una probabilidad del 92.9 % de no tenerla y si dice que no la tiene, tiene una probabilidad del 91.3 % de no tenerla, bastante equilibrado. Entonces, es conveniente usar US para confirmar el diagnóstico de AA del MU, lo que reducirá los falsos positivos.

Comentario: No como segunda opinión siempre: solo para los positivos según el MU. Para los negativos según el MU, el US empeora el diagnóstico.

Por lo tanto, es recomendable contratar un experto en US para evitar falsos positivos, y el protocolo para una persona (adulto) con síntomas de AA debería ser:

- Pasa por el MU. Si le diagnostica que no tiene AA, se considera que no tiene AA (tiene otra cosa que habrá que diagnosticar).

- Si el MU cree que tiene AA, se le hace una ecografía. Si se confirma el diagnóstico, va a quirófano. Si no, se considera que no tiene AA (tiene otra cosa que habrá que diagnosticar).

j) (Opcional) ¿Creéis que valdría la pena contratar el experto en ultrasonidos para complementar el diagnóstico clínico de AA en el servicio de urgencias pediátricas? En caso afirmativo, ¿cómo recomendaríais que se complementasen ambos diagnósticos?

Al ser la prevalencia muy alta, el VPP del MU ya es muy alto, un 85%. En cambio, ahora el VPN del diagnóstico del especialista en ultrasonidos es de 0.40, por lo tanto el peligro de un falso negativo es alta. Es mucho más difícil de justificar en este caso el uso de la ecografía para confirmar el diagnóstico.