



Fundación
J. L. Castaño

Para el desarrollo del Laboratorio clínico

ESTADÍSTICA BÁSICA APLICADA AL LABORATORIO CLÍNICO

Ed Cont Lab Clín; 30: 51 - 59

SEQC^{ML}

Sociedad Española de Medicina de Laboratorio

2016-2017

CARACTERÍSTICAS DE LOS TEST DIAGNÓSTICOS.

Inmaculada Pérez de Algaba Fuentes

UGC Intercentros de Laboratorio. Hospital Regional Universitario de Málaga. Málaga.

Boris Battikhi Vilar

UGC Intercentros de Laboratorio. Hospital Regional Universitario de Málaga. Málaga..

1. Introducción

Comenzaremos describiendo un caso descrito en una página web dedicada a la divulgación de la estadística:

Se trata de una pareja que acude a su médico tras una revisión de embarazo. Llegan muy agobiados, ya que en el informe de laboratorio le comunican que le ha dado positivo un test de SIDA. Ellos no pueden explicarse el resultado ya que en principio no pertenecen a grupos de riesgo. El facultativo, bastante formado, trata de tranquilizar a la pareja, les comenta que todos los laboratorios por pura estadística pueden dar resultados falsos positivos, y que este es probablemente su caso, por lo que les dicen que se realicen una nueva prueba.

Lo que dice este doctor es algo que conocemos en el laboratorio. Sabemos que cuando valoramos unos rangos de referencia, dejamos un 5 % de muestra que serían falsos positivos. Cuando mediante un test cuantitativo, queremos establecer un punto de corte, el problema es distinto, ya que, en función de donde fijemos el punto de corte, vamos a tener más falsos positivos o más falsos negativos. Esto lo podemos ver gráficamente (Figura 1). En esta figura representamos la distribución de resultados de test en individuos sanos y enfermos.

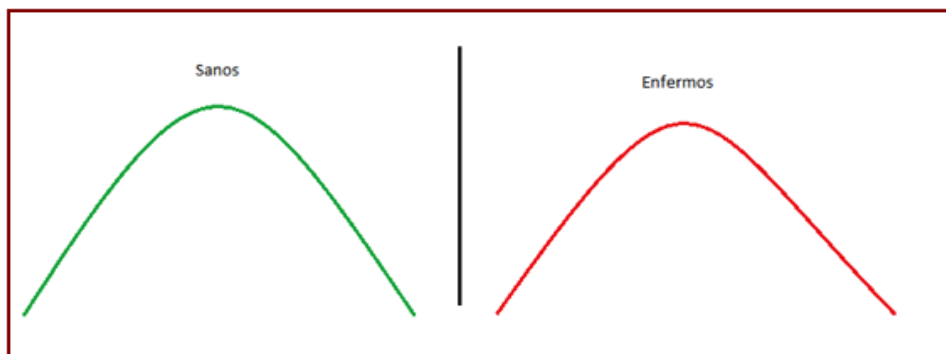


Figura 1: Distribución de resultados en una muestra sana y otra patológica.

Según este gráfico tendríamos un test perfecto en el que no habría falsos positivos ni falsos negativos, ya que la población sana está bastante separada de la población enferma. Pero esto no es lo habitual (Figura 2)

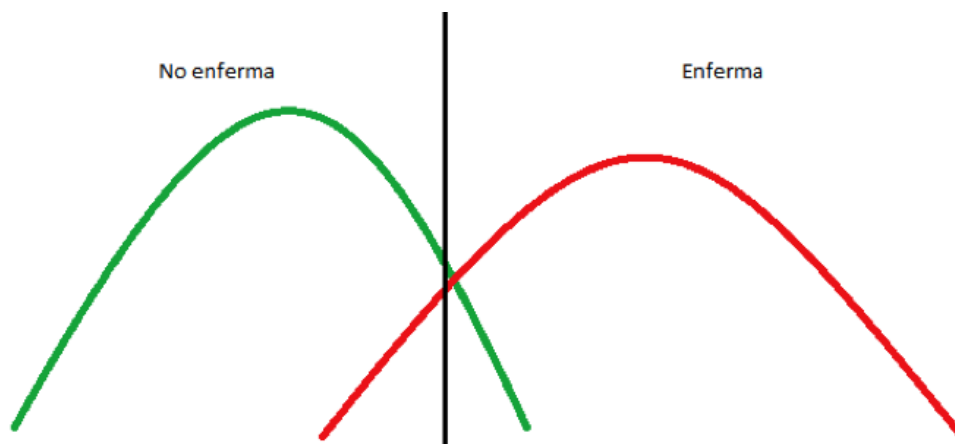


Figura 2: Distribución de población sana y enferma.

Este caso es más habitual. Vemos que hay un rango en el que los resultados se comparten entre la población sana y enferma. Un resultado en este rango podría ser tanto positivo como negativo.

Esto es lo que vamos a estudiar en este tema, intentar conocer esos falsos negativos o falsos positivos para poder interpretar correctamente un resultado.

En el laboratorio podemos usar un test para hacer un cribado en una población. Si estoy buscando una determinada patología me interesará no tener ningún falso negativo, ya que los falsos positivos después los puedo detectar con una prueba posterior. Mientras que a los negativos no los voy a estudiar más. En cambio si deseo diagnosticar una enfermedad, tengo que acudir a un test que no tenga falsos negativos ya que voy a dejar de diagnosticar a una parte de la población que está enferma.

2. Características pre-Test

Esto lo vamos a estudiar mediante la Sensibilidad y la Especificidad, y lo vamos a realizar estudiando las probabilidades.

La Sensibilidad es la probabilidad de tener un test positivo cuando se está enfermo, o dicho de otra forma el cociente entre los verdaderos enfermos y el total de enfermos (verdaderos enfermos más falsos enfermos), o también los que tienen un test positivo y están enfermos dividido por la suma de éstos y los que tienen un test negativo y están enfermos.

$$S = \frac{VP}{VP + FN}$$

Igualmente definimos la Especificidad como la probabilidad de tener un test negativo cuando se está sano. Es el ratio entre los verdaderos sanos y la suma de los verdaderos sanos y los sanos que tienen un test positivo. O, lo que es lo mismo, los sanos con test negativo dividido por la suma de los sanos con test negativo más los sanos con test positivo.

$$E = \frac{VN}{VN + FP}$$

Podemos poner un ejemplo: quiero estudiar un método de cribado para diagnóstico de infección por HIV. Para ello proceso 203 muestras mediante el test de cribado y por un test de referencia para diagnóstico de SIDA por ejemplo, una carga viral. Los resultados los podemos resumir en la Tabla 1.

		VIH		Total pacientes
		SI	NO	
Test de cribado	Negativo	A Falso negativo (FN) 12	B Verdaderos negativos (VN) 103	115
	Positivo	C Verdadero positivo (VP) 80	D Falso positivo (FP) 8	88

Tabla 1: Resultado del test de cribado y la prueba definitiva.

Con estos datos puedo calcular la sensibilidad y la especificidad.

$$S = \frac{VP}{VP + FN} = \frac{80}{80 + 12} \times 100 = 86,9 \%$$

$$E = \frac{VN}{VN + FP} = \frac{103}{103 + 8} \times 100 = 92,7 \%$$

Como podéis observar la sensibilidad depende de los falsos negativos y la especificidad depende de los falsos positivos. Estas propiedades son características de cada test y se conocen también como especificaciones pre-test.

Podemos definir la Probabilidad de falsos negativos (PFN) como el complementario de la sensibilidad o, lo que es lo mismo, la probabilidad de obtener en nuestro test un resultado negativo en un individuo enfermo, y sería

$$PFN = \frac{FN}{FN + VP} = \frac{12}{80 + 12} \times 100 = 13,1 \%$$

La sensibilidad más la PFN sería 100. Igualmente podemos definir la probabilidad de falsos positivos (FP) como el complementario de la especificidad. Sería la probabilidad de obtener un resultado positivo entre los individuos sanos. La fórmula es:

$$PFP = \frac{FP}{FP + VN} = \frac{8}{103 + 8} \times 100 = 7,3 \%$$

Igual que en la sensibilidad, la PFP más la especificidad es 100.

3. Características post-Test

Todos habéis escuchado alguna vez que un test determinado es útil para descartar una enfermedad. Por ejemplo, en el caso del NT-proBNP, se emplea un punto de corte en urgencias para descartar la enfermedad cardíaca por su elevado valor predictivo negativo.

Vamos a aclarar estos conceptos:

Definimos el Valor Predictivo Positivo (**VPP**) como la probabilidad de estar enfermo teniendo un test positivo. Su expresión matemática sería:

$$VPP = \frac{VP}{VP + FP}$$

Es mayor cuanto menor es el número de falsos positivos.

El Valor Predictivo Negativo (**VPN**) es la probabilidad de no estar enfermo teniendo un test negativo:

$$VPN = \frac{VN}{VN + FN}$$

Su valor es más alto cuanto menor es el número de falsos negativos. En el caso de disponer de un test con un número de falsos negativos muy bajo se podría emplear como test para descartar una determinada patología.

Los valores predictivos dependen de la prevalencia y, por tanto, pueden ser diferentes en distintas poblaciones:

$$VPP = \frac{\text{Prevalencia} \times \text{Sensibilidad}}{\text{Prevalencia} \times \text{Sensibilidad} + (1 - \text{Prevalencia}) \times \text{Falsos Positivos}}$$

Un test con alta especificidad (baja proporción de falsos positivos) que dé un resultado positivo, indica que existe alta probabilidad de enfermedad.

$$VPN = \frac{(1 - \text{Prevalencia}) \times \text{Especificidad}}{(1 - \text{Prevalencia}) \times \text{Prevalencia} + \text{Prevalencia} + \text{Falsos Negativos}}$$

Un test con alta sensibilidad (baja proporción de falsos negativos) que dé un resultado negativo, indica que existe alta probabilidad de no enfermedad.

Cuando un facultativo pide un test de forma arbitraria en la población general para una determinada patología, al ser baja la prevalencia los valores predictivos serán bajos. En cambio, si selecciona una población en la que se cumplan unos requisitos de signos y síntomas de la enfermedad en estudio, lo que está haciendo es aumentar la prevalencia en la población en la que va a solicitar el test, de esta forma lo que hace es aumentar los valores predictivos.

La razón de verosimilitud positiva (LR+) mide cuántas veces es más probable un resultado positivo en un enfermo en relación a un sano:

$$LR+ = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}} = \frac{\text{Verdaderos Positivos}}{\text{Falsos Positivos}}$$

Conocida la LR+, podemos conocer la diferencia entre la probabilidad pre-test (prevalencia) y la post-test (VP+). A mayor LR+, mayor VP+

$$VP+ = \frac{\text{Prevalencia}}{\text{prevalencia} + \frac{1 - \text{Prevalencia}}{LR+}}$$

También existe una razón de verosimilitud negativa (LR-), la cual mide cuantas veces es más probable un resultado negativo en un sano en relación a un enfermo:

$$LR- = \frac{\text{Especificidad}}{1 - \text{Sensibilidad}} = \frac{\text{Verdaderos Negativos}}{\text{Falsos Negativos}}$$

A mayor LR- , mayor VP-:

$$VP- = \frac{1 - \text{Prevalencia}}{(1 - \text{Prevalencia}) + \frac{(1 - \text{Prevalencia})}{LR-}}$$

Vamos a poner un ejemplo para afianzar los conocimientos de una forma práctica. Vamos a usar unos datos publicados en el BMJ(1998;316:577-82). En este trabajo estudian el empleo de la proteína S100, en el diagnóstico de la enfermedad de Creutzfeldt-Jakob (CJ). El diagnóstico definitivo es fundamentalmente clínico mediante electroencefalograma y los síntomas característicos de la enfermedad. La incidencia en Europa es del 0,5 a 1 x10⁶. El estudio se realiza en dos poblaciones en las que las prevalencias son diferentes. Los resultados se muestran en la Tabla 2:

S100 (pg/mL)	CJ		S100 (pg/mL)	CJ	
	SI	No		SI	No
>213	84	14	>213	84	20412
<213	24	60	<213	24	87480
	108	74		108	107892

Tabla 2: Ejemplo de dos estudios el de la derecha con una prevalencia de 0.59 y el de la izquierda con una prevalencia de 108100.

La prevalencia en el primer caso es de 0,59 y en el segundo caso es de 0,001.

Los resultados se muestran en la Tabla 3:

Prevalencia	0,59	0,001
Sensibilidad	0,78	0,78
Especificidad	0,81	0,81
VPP	0,86	0,004
VPPN	0,71	0,9997

Tabla 3: Resultados del estudio.

Como podéis ver en este ejemplo, los valores predictivos están muy influenciados por la prevalencia, mientras que los valores pre-test son característicos del ensayo.

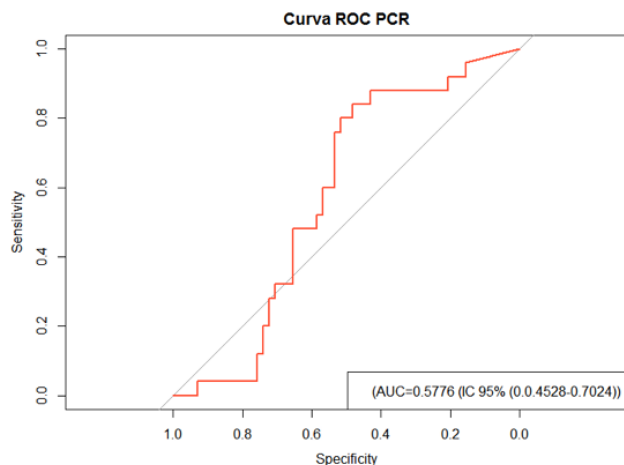
Para el estudio de estos parámetros necesitamos los datos del test en estudio, y los datos de un método definitivo de diagnóstico, para poder considerar los datos como verdaderos o falsos. El método definitivo puede ser un "gold estándar" o bien, un método que esté en uso y otro que queramos estudiar.

Los datos en estos estudios pueden ser cualitativos o cuantitativos. Sin embargo, cuando los datos son cualitativos, nos interesa establecer un punto de corte para el cual la sensibilidad y especificidad sean máximas. Esto se hace mediante el estudio de curvas ROC.

4. Curvas ROC

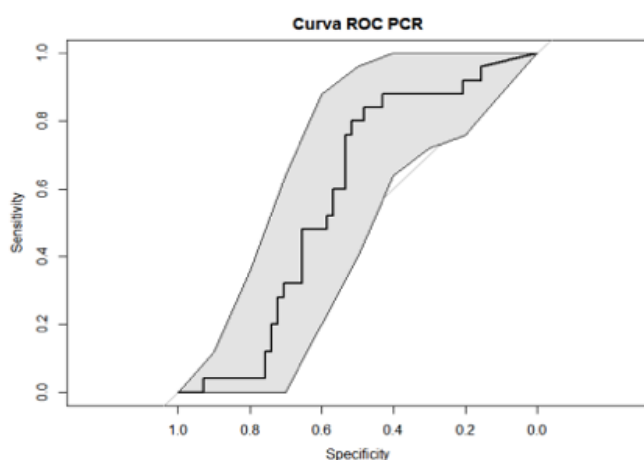
El nombre proviene del inglés "Receiver Operating Characteristic". Estas curvas relacionan la especificidad y la sensibilidad. Esta última se representa en el eje de la "y" y en el eje de la "x" se representa 1 menos la especificidad. Lo que nos interesa es el área que se sitúa bajo la curva, que puede tomar valores entre 0 y 1. Si el área bajo la curva (AUC) es 1, el test discrimina perfectamente, y la sensibilidad y especificidad tomarán valores de 1 (o 100 %). Si el valor es de 0,5, no hay discriminación o, lo que es lo mismo, la probabilidad de tener un test positivo estando enfermo es la misma que tener el test positivo estando sano.

En la Figura 3 podéis observar una curva ROC. Ya hemos comentado que en el eje de abscisas se representa 1 menos la especificidad. Sin embargo en la curva sólo figura la especificidad, pero fijaros que el eje va desde 1 a 0.



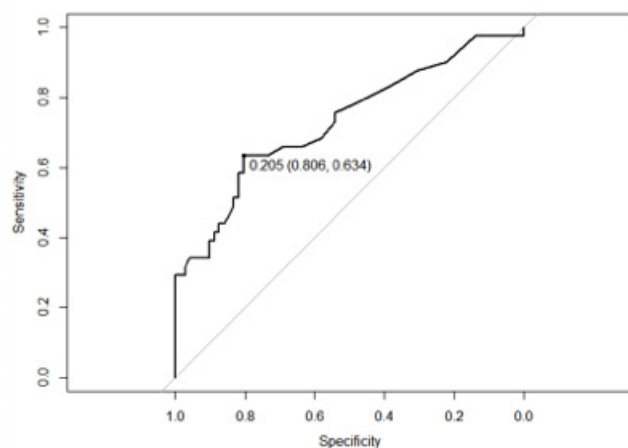
Cuando veáis un área bajo la curva en un trabajo fijaros que se informa ésta junto a su intervalo de confianza, ya que si este contiene el 0,5, la curva no tiene poder de discriminación como ya hemos comentado. En este caso, observad como el IC 95 % contiene el 0,5.

Figura 3: Curva ROC.



También podemos dibujar el intervalo de confianza en la curva (Figura 4).

Figura 4: Curva ROV con IC 95 %.



Sería interesante ver el punto de corte. Ya hemos comentado que una posibilidad sería establecer el punto con máxima sensibilidad y especificidad (Figura 5).

Figura 5: Punto de corte con el IC con mejor sensibilidad y especificidad.

Pero también hemos comentado que podríamos elegir este corte dependiendo del uso que demos al test. Si este es de cribado, nos interesa la máxima sensibilidad, mientras que si es diagnóstico, nos interesa que la especificidad sea máxima. R nos ofrece la posibilidad de elegir el punto que más nos interesa. En el cuadro de texto 1, podemos ver las especificidad con su IC 95 % en función de la sensibilidad.

```

95% CI (100 stratified bootstrap replicates):
  se sp.low sp.median sp.high
0.0 1.00000 1.0000 1.0000
0.1 0.64610 0.7586 0.8793
0.2 0.62070 0.7414 0.8457
0.3 0.55990 0.6897 0.8022
0.4 0.51720 0.6552 0.7849
0.5 0.47370 0.6121 0.7241
0.6 0.43100 0.5517 0.6815
    
```

Cuadro de texto 1: Puntos de corte de especificidad con IC 95% en función de la sensibilidad.

También podríamos calcular la especificidad con su IC en función de la sensibilidad. Podríamos además conocer la potencia de la curva. R nos da también este dato (Cuadro de texto 2).

```

One ROC curve power calculation

ncases = 25
ncontrols = 58
auc = 0.5775862
sig.level = 0.05
power = 0.20934
    
```

Cuadro de texto 2: Datos de la potencia de la curva.

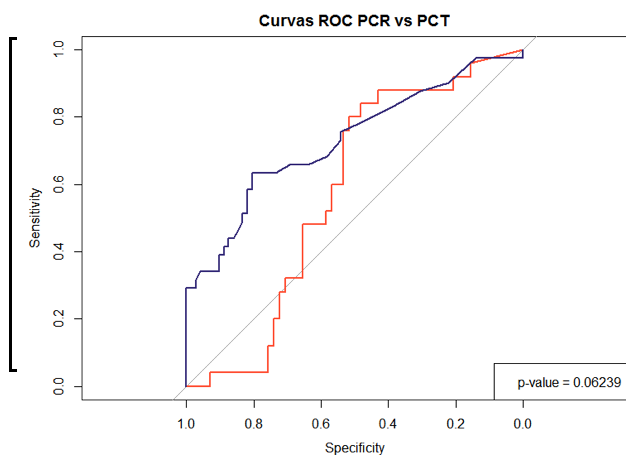
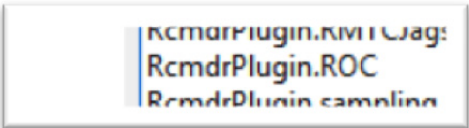


Figura 6: Comparación de curvas ROC.

Además podríamos comparar dos curvas, tanto gráficamente como mediante un test de comparación. En la Figura 6 se muestran las dos curvas y una “p” que, como ya hemos comentado, es la probabilidad de cometer un error al rechazar la hipótesis nula. En nuestro caso la hipótesis nula es que las dos curvas son iguales.

Podéis ver en el gráfico las dos curvas y la “p” resultante del test (0,062), muy alta, por lo que no podemos rechazar la hipótesis nula y decir que las dos curvas son iguales.

Por último comentar que para trabajar con Curvas ROC, tenemos como siempre dos posibilidades: mediante Rcmdr y mediante Rstudio. En el primer caso necesitaremos el paquete RcmdrPlugin.ROC y, si lo hacemos con, Rstudio deberíamos cargar el paquete pROC



```
RcmdrPlugin.RMTCag:  
RcmdrPlugin.ROC  
RcmdrPlugin.camling
```

En el caso clínico construiremos las curvas con ambos paquetes.

EDUCACIÓN CONTINUADA EN EL LABORATORIO CLÍNICO COMITÉ DE EDUCACIÓN

D. Balsells, B. Battikhi (*Residente*), R. Deulofeu, M. Gassó, N. Giménez, J.A. Lillo, A. Merino, A. Moreno, A. Peña (*Residente*), M. Rodríguez (*Presidente*), N. Rico, MC. Villà.

ISSN 1887-6463 – Abril 2017 (recibido para publicación Marzo 2017).