

Atención a la Salud de la Comunidad

Tema 04. Pruebas diagnósticas



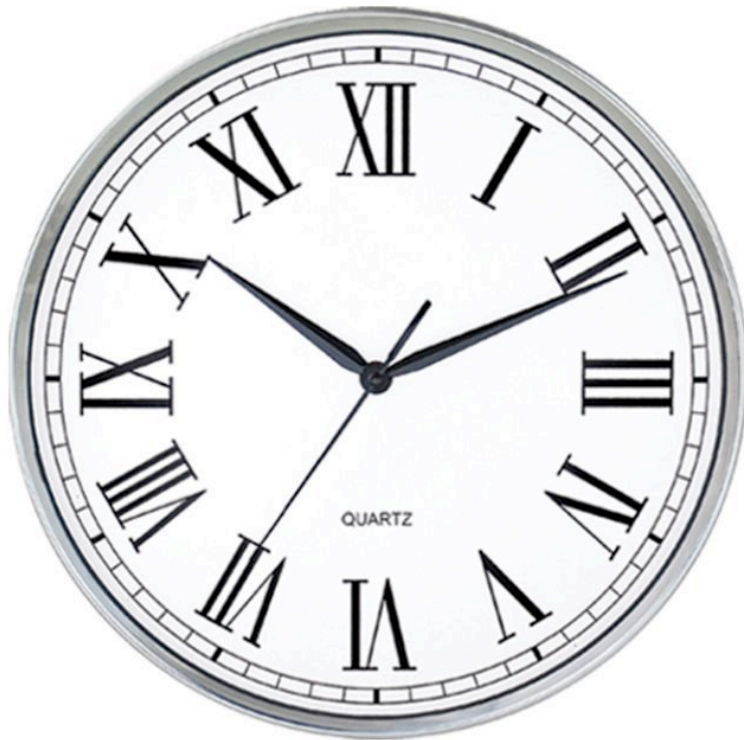
Celia Nespral Gaztelumendi
Miguel Santibáñez Margüello
Montserrat Bustamante Fonfría

DPTO. DE ENFERMERÍA

Este tema se publica bajo Licencia:

[Creative Commons BY-NC-SA 3.0](https://creativecommons.org/licenses/by-nc-sa/3.0/)

*"Hasta un reloj estropeado
acierta dos veces al día".*



Woody Allen
Anything Else
(*Todo lo demás*), 2003

“Lanzar una moneda al aire, cada vez que viene un paciente tiene una sensibilidad y especificidad del 50%”.



Objetivos:

1. Distinguir entre reproductitibilidad y validez en el diagnóstico.
2. Saber medir e interpretar la validez de una prueba diagnóstica.
3. Saber interpretar una curva ROC
4. Saber interpretar críticamente un artículo sobre pruebas diagnósticas

Para más información:

P. Fernández et al. Pruebas diagnósticas. Curvas ROC.

<http://www.fisterra.com/mbe/investiga/index.asp>

Argimón Pallàs, JM. Jiménez Villa, J. *Métodos de investigación clínica y epidemiológica*. 3ª ed. Madrid, Elsevier, 2006.

Domenech JM. Fundamentos de diseño y Estadística. UD 3. Teoría y calculo de probabilidades. Pruebas diagnósticas. 12ª ed. Barcelona: Signo; 2011.

Delgado M, Llorca J, Domenech JM. Estudios para pruebas diagnosticas y factores pronósticos. 3ª ed. Barcelona: Signo; 2008.

Reproductibilidad

La reproductibilidad recibe también los nombres de concordancia, acuerdo y fiabilidad (reliability).

Mide el grado de coincidencia de dos mediciones sobre la misma muestra de individuos. Puede ser **intraprueba (una prueba consigo misma)** o entre dos o más pruebas.

Para evaluar la coincidencia mas allá del azar, se utiliza el índice Kappa.

Interpretación Fleiss

kappa	grado
≥ 0.75	excelente
0.40 - 0.75	moderado
≤ 0.40	deficiente

Interpretación Landis y Koch

kappa	grado de acuerdo
< 0	sin acuerdo
0 - 0,2	insignificante
0,2 - 0,4	bajo
0,4 - 0,6	moderado
0,6 - 0,8	bueno
0,8 - 1	muy bueno

Reproductibilidad

Con frecuencia se suele conceder mucho menos valor a la reproductibilidad que a la validez en el diagnóstico.

Esto es del todo incorrecto-> PARA PODER VALORAR LA VALIDEZ HAY QUE TENER ANTES REPRODUCTIBILIDAD (La reproductibilidad es un requisito previo para la Validez).

Si mi prueba cada vez que la hago me da un resultado diferente: ¿para que continuar?

Validez

Se entiende por validez en el diagnóstico, el grado en que una medición coincide con la verdad (con el gold standard).

-*Parámetros de validez interna*. Sensibilidad y especificidad.

-*Valores predictivos*. Totalmente relacionados con la prevalencia. Formula de Bayes.

-*Razones de verosimilitud*. Óptica de la Medicina Basada en la evidencia.

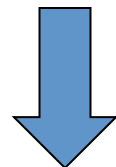
Parámetros de Validez Interna

Sensibilidad/ Especificidad

Matemáticamente son proporciones-> = tienen un denominador en el que se incluye el numerador

Esto es muy importante, porque se pueden calcular los IC95% (metodo Wilson).

También es muy importante, porque se puede calcular a priori el tamaño de la muestra para obtener unos IC95% mas o menos precisos.



	Enfermos	Sanos	Total
Test Dº +	VP (a)	FP (b)	a+b
Test Dº -	FN (c)	VN (d)	c+d
	a+c	b+d	N

VP: verdaderos positivos

VN: verdaderos negativos

FP: falsos positivos

FN: falsos negativos

Parámetros de VALIDEZ INTERNA:

Sensibilidad: probabilidad de que un individuo con enfermedad sea clasificado como enfermo; o proporción de verdaderos positivos entre los enfermos.

Una prueba muy sensible halla muchos positivos y pocos falsos negativos. Su utilidad principal es la de descartar la enfermedad.

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

La sensibilidad es especialmente importante cuando es grave no diagnosticar una enfermedad y/o cuando el pronóstico mejora mucho con el tratamiento precoz ⇒ Las pruebas de cribado de una enfermedad (población sana) requieren alta sensibilidad, si bien luego ha de poder confirmarse el diagnóstico con una prueba mas específica.

Parámetros de VALIDEZ INTERNA:

Especificidad: probabilidad de que un individuo sin enfermedad sea clasificado como sano; o proporción de verdaderos negativos en no enfermos (**sanos**). Una prueba muy específica halla muchos verdaderos negativos y pocos falsos positivos.

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

Como una prueba específica tiene pocos falsos positivos, su utilidad principal será la de confirmar un diagnóstico.

Es importante cuando la enfermedad cambia la vida del enfermo o tiene cierto estigma (Sida) o cuando las consecuencias del tratamiento son relevantes...

Parámetros de VALIDEZ INTERNA:

La sensibilidad y especificidad nos habla de la validez interna de la prueba, es decir, de la exactitud diagnóstica (diagnostic accuracy), eficacia, o rendimiento diagnóstico (si acierta en lo que quiere medir y por tanto es un buen instrumento de medición).

No es posible por tanto dar cifras de los niveles aceptables de estos índices ya que dependen de los objetivos de la prueba.

No obstante, en general se considera que una prueba tiene una exactitud aceptable si su sensibilidad y especificidad alcanzan el umbral de 0.8.

Las preguntas que te interesan son:

Si el test me da positivo, que probabilidad tengo de estar enfermo?

Si el test me da negativo, que probabilidad tengo de estar sano?

La respuesta a estas preguntas está en el VPP y en el VPN

Estudio del poder predictivo de una prueba

Las preguntas que a un clínico le interesan son:

Si el test me da positivo, que probabilidad tiene el paciente de estar enfermo?

Si el test me da negativo, que probabilidad tiene el paciente de estar sano?

La respuesta a estas preguntas está en el VPP y en el VPN

	Enfermos	Sanos	Total
Test Dº +	VP (a)	FP (b)	a+b
Test Dº -	FN (c)	VN (d)	c+d
	a+c	b+d	N

VP: verdaderos positivos

VN: verdaderos negativos

FP: falsos positivos

FN: falsos negativos

Estudio del poder predictivo de una prueba

Valor predictivo positivo: probabilidad de estar enfermo cuando el test es positivo. Se calcula dividiendo los verdaderos positivos entre el total de resultados positivos.

$$\text{VPP} = \frac{a}{a + b} = \frac{\text{VP}}{\text{VP} + \text{FP}}$$

Valor predictivo negativo: probabilidad de estar sano cuando el test es negativo. Se calcula dividiendo los verdaderos negativos entre el total de resultados negativos.

$$\text{VPP} = \frac{d}{c + d} = \frac{\text{VN}}{\text{VN} + \text{FN}}$$

Probabilidad a priori o pre-test= probabilidad que tienes de estar enfermo antes de hacerte el test. Es la prevalencia de la enfermedad en la población a la que perteneces.

Probabilidad a posteriori o posttest= probabilidad que tienes de estar enfermo después de hacerte el test. Es el VPP.

Como es lógico, lo interesante es que la probabilidad posttest sea mucho mayor que la probabilidad pretest.

La prevalencia influye en el VPP (a mayor prevalencia mayor VPP y viceversa).

El entendimiento de este concepto es fundamental para la aplicación clínica de la prueba.

Relación del VPP con la PREVALENCIA

Una prueba con una sensibilidad=80% y especificidad=90%, se aplica a una población con una **prevalencia de enfermedad=50%**

	Enfermos	Sanos		Prevalencia=100/200=50%
Test+	80	10	90	$S = 80/100 = 80\%$
Test-	20	90	110	$E = 90/100 = 90\%$
	100	100	200	$VPP = 80/90 = 88.8\%$

* Si la misma prueba se aplica a una población con una prevalencia menor (5%) en una población de 2000 personas, disminuirá el VPP, aún manteniendo la misma sensibilidad y especificidad

	Enfermos	Sanos		$P = 5\% \Rightarrow 0.05 = (\text{total enfermos})/2000$
Test+	80	190	270	Total enfermos = 100
Test-	20	1710	1730	$S = 80\% \Rightarrow 0.80 = VP/100 \Rightarrow VP = 80$
	100	1900	2000	$E = 90\% \Rightarrow 0.90 = VN/1900 \Rightarrow VN=1710$
				VPP = 80/270 = 30%

Teorema de Bayes

Lo que hemos hecho , es aplicar el teorema de Bayes. Es decir, conociendo la sensibilidad y especificidad de la prueba, podemos generar los valores predictivos positivos y negativos para cualquier prevalencia de enfermedad (bastaría con adaptar la tabla de contingencia 2x2).

Una prueba son $S=0.99$ y $E=98\%$ (casi perfecta), aplicada a una población con prevalencia 1 por mil, tiene solo un VPP del 4.7% (Prevalencia del Sida en población general española: de cada 100 positivos en la prueba, solo 5 estarán infectados)

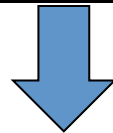
Antes de aplicar una prueba a una determinada población, es de suma importancia tener en cuenta los valores predictivos y negativos esperados, con el fin de disponer de una valoración previa de la utilidad real de la prueba.

Las pruebas diagnósticas desde la MBE www.cebm.net

LA MBE promueve la obtención de índices de fácil cómputo que a la vez resulten sumamente informativos en términos clínicos y de gestión sanitaria

Las razones de verosimilitud o cocientes de probabilidad (likelihood ratios, LR) relacionan los conceptos de sensibilidad y especificidad en un solo índice.

$$RV_{+} = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}}$$



$$RV_{-} = \frac{1 - \text{Sensibilidad}}{\text{Especificidad}}$$



	Enfermos	Sanos	Total
Test Dº +	VP (a)	FP (b)	a+b
Test Dº -	FN (c)	VN (d)	c+d
	a+c	b+d	N

VP: verdaderos positivos

VN: verdaderos negativos

FP: falsos positivos

FN: falsos negativos

Las pruebas diagnósticas desde la MBE www.cebm.net

Ejemplo. VIH y ELISA. LR+ 49.5.

Si nos fijamos bien, quiere decir que en enfermos, la probabilidad de un resultado + es 49.5 veces superior que en sanos

Cuanto mayor es respecto a 1, mayor certidumbre que un paciente con resultado + presenta la enfermedad.

$$RV_{+} = \frac{\text{Sensibilidad}}{1 - \text{Especificidad}}$$

+ en enf / + en sanos

	Enfermos	Sanos	Total
Test Dº +	VP (a)	FP (b)	a+b
Test Dº -	FN (c)	VN (d)	c+d
	a+c	b+d	N

VP: verdaderos positivos

VN: verdaderos negativos

FP: falsos positivos

FN: falsos negativos

Casi cualquier prueba, distingue entre sanos y gravemente enfermos



Diseño de estudios para valorar pruebas diagnósticas

La composición del grupo de enfermos influye en la sensibilidad.

La composición del grupo de sanos influye en la especificidad.

O sea, la sensibilidad y especificidad también dependen de la prevalencia*

La correcta elección de los sujetos de estudio, que es crucial, comporta aplicar la prueba a pacientes que representen a la población en la que posteriormente se aplicara dicha prueba (Hulley y Cummings, 1993).

Es decir, tanto la muestra de enfermos como de no enfermos, deben ser representativas de los sujetos a los que se aplicara la prueba en condiciones habituales o de rutina.

* Brenner H, Gefeller O. Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence. Stat Med. 1997; 16:981-91.

Curvas ROC

Para que proceda una curva ROC, tu prueba a estudio debe de dar un valor cuantitativo.

Ej: H Pylori y PCR a tiempo real.

Es decir:

Gold Standard= Enf/ No Enf (Variable cuantitativa binaria)

Prueba a estudio= variable cuantitativa continúa

Curvas ROC

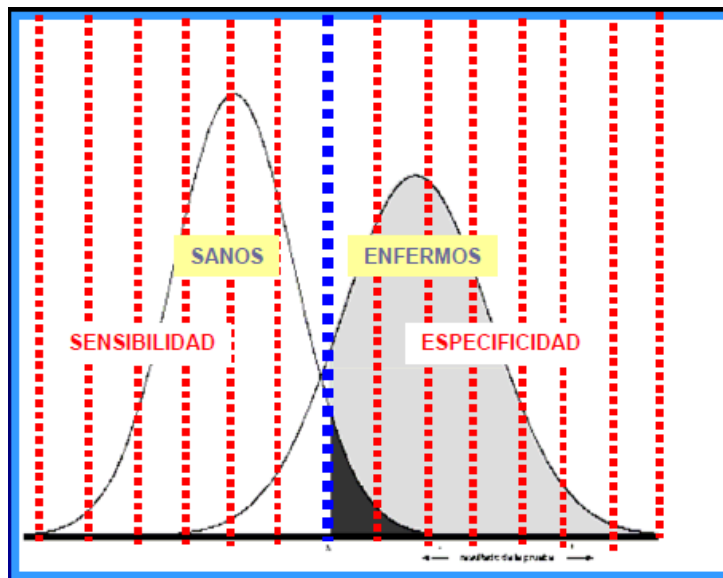
Ejemplo. Niveles de Glucosa y diabetes.

La población de sanos tendrá unos niveles de glucemia.

La población de enfermos otros niveles.

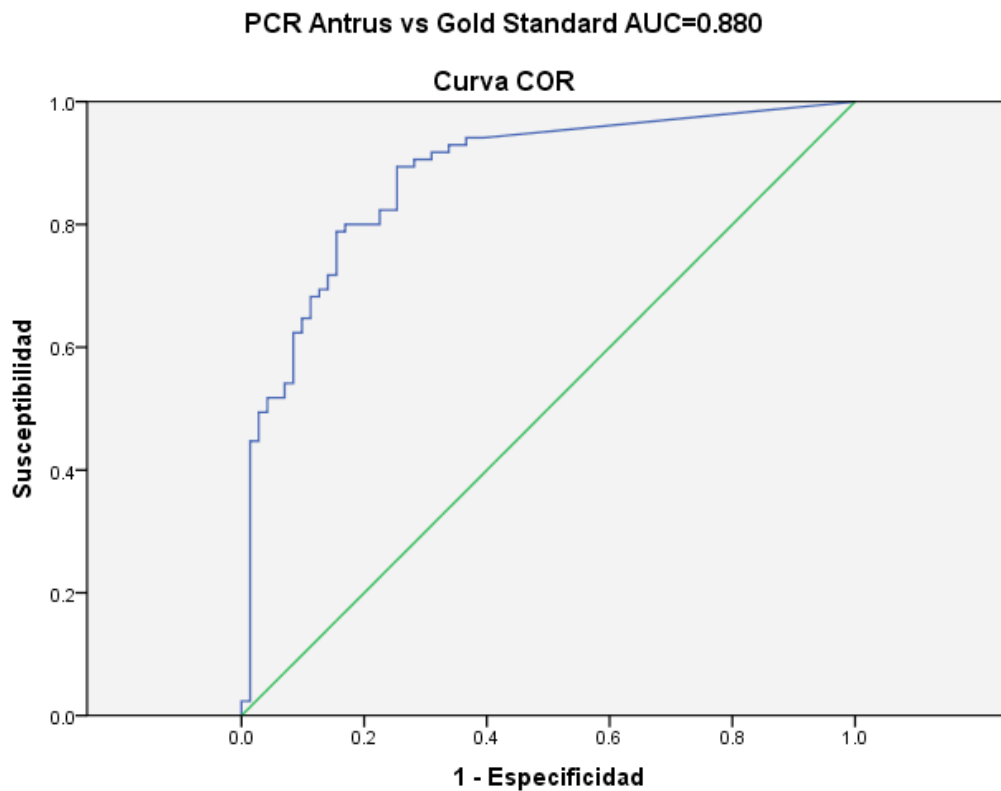
Ambos niveles se cruzarán

Cuanto más altos sean tus niveles, más probabilidad será que pertenezcas a la población de enfermos (más especificidad) pero a costa de dejar atrás enfermos detrás.



Curvas ROC

La curva ROC lo que te hace es darte una idea global de la prueba, el peor valor es 0.5, el mejor es 1.



Los segmentos diagonales son producidos por los empates.

Despues viene la pregunta del millón, qué punto de corte elijo...

Se suele coger el de máx área, pero no hay una única respuesta para esto, depende de lo que te supongan los falsos negativos y positivos

Si sensibilidad elevada, habrá pocos FN-> los negativos serán verdaderos negativos. Priorizaras esto cuando priorices descartar la enfermedad con gran certitud.

Si especificidad elevada, habrá pocos FP -> los positivos serán VP. Priorizaras esto cuando priorices diagnosticar la enfermedad con gran certitud.

	Enfermos	Sanos	Total
Test Dº +	VP (a)	FP (b)	a+b
Test Dº -	FN (c)	VN (d)	c+d
	a+c	b+d	N

VP: verdaderos positivos

VN: verdaderos negativos

FP: falsos positivos

FN: falsos negativos

Standards for Reporting of Diagnostic Accuracy -STARD ***(Estándares para los informes validez diagnóstica)***

Se trata de una **lista de 25 puntos** que cualquier artículo de este tema debe cumplir para llegar al standard de calidad en pruebas diagnósticas.

BMJ Group requires compliance to the following reporting guidelines, please upload your completed check-list with your submission and label it " Research checklist":

Guidance and forms are available [here](#)

<http://www.espanol.equator-network.org/home/>

<http://www.equator-network.org/resource-centre/library-of-health-research-reporting/>

- CONSORT statement - *Required for all randomised controlled trials*
- PRISMA statement - *Required for all systematic reviews*
- EVEREST statement - *Required for all economic evaluations*
- STARD statement - *Required for all diagnostic research papers*
- STROBE statement - *Required for all observational studies*
- MOOSE statement - *Required for all meta-analyses of observational studies*