

# Introducción a la neuroinvestigación

<sup>1</sup>Carlos Valencia Calderón, <sup>1</sup>Amadeo Muntané Sánchez, <sup>1</sup>Cristina Masuet Augmantell,  
<sup>2</sup>Ana Calderón Valdiviezo, <sup>3</sup>María Luisa Moro Esteban

## RESUMEN

Las publicaciones científicas desempeñan un papel fundamental en las distintas etapas de la actividad investigadora. El desarrollo de cualquier proyecto de investigación requiere de un conocimiento elemental sobre estadística y epidemiología, áreas del conocimiento científico un poco alejadas de la práctica clínica habitual del neurólogo y de muchos otros especialistas en neurociencias, por lo que el enfoque de un tema tan árido como éste exige recurrir a muchos y constantes ejemplos para su comprensión y para que el trabajo tenga el beneficio esperado. Las neurociencias, como una de las más importantes ramas de la medicina, siempre se han caracterizado por un denso cuerpo de conocimientos clínicos y semiológicos, ahora apoyados por el desarrollo de nuevas técnicas diagnósticas. Estas pruebas tecnológicas no dejan de ser unos mecanismos artificiales para modificar la probabilidad diagnóstica de determinada enfermedad, de ahí la importancia de conocer la auténtica correspondencia entre la validez interna de una prueba diagnóstica y la realidad patológica, ya que no toda prueba diagnóstica positiva significa padecer la enfermedad en estudio, ni toda prueba negativa la descarta. Muy pocas pruebas diagnósticas, quizás ninguna, identifican con total certeza si el paciente tiene o no la enfermedad. El presente artículo pretende revisar las propiedades operativas de las pruebas diagnósticas para que sean consideradas con más frecuencia, y de modo explícito, en nuestra

Recibido: 23 octubre 2006. Aceptado: 8 noviembre 2006.

<sup>1</sup>Hospital Universitario de Bellvitge, <sup>2</sup>Sardomus, <sup>3</sup>CAP «Les Corts», Barcelona. Correspondencia: Carlos Valencia Calderón. Servicio de Neurocirugía, Hospital Universitario de Bellvitge. Feixa Llarga s/n. 08907. L'Hospitalet de Llobregat. Barcelona. Email: cvalencia@csub.scs.es

práctica clínica habitual.

**Palabras claves:** epidemiología, neuroinvestigación, prevalencia, sensibilidad.

## SCIENTIFIC PUBLICATIONS REPRESENT AN IMPORTANT PART OF THE RESEARCH

## ABSTRACT

The development of any research project requires a knowledge of statistics and epidemiology that are not always available to a neurologist or other neurosciences investigators and sometimes he has to use many examples in order to make his presentation more explicit overall that recent developments in diagnostic tests that not always give a complete or correct diagnosis. In this paper we make a critical analysis of these tests to evaluate their real value in the diagnosis.

**Key words:** epidemiology, neuroinvestigation, prevalence, sensibility of tests.

Las publicaciones científicas desempeñan un papel fundamental en las distintas etapas de la actividad investigadora. Constituyen el punto de partida de cualquier investigación, pues aportan el estado del conocimiento en un tema determinado, y son el canal preferente por el que se darán a conocer los resultados de una investigación ya concluida<sup>1</sup>.

El desarrollo de cualquier proyecto de investigación requiere de un conocimiento elemental sobre estadística y epidemiología, áreas del conocimiento científico un poco alejadas de la práctica clínica habitual del neurólogo y de muchos otros especialistas en neurociencias, por lo que el enfoque de un tema tan árido como éste exige recurrir a muchos y constantes ejemplos para su comprensión y para que el trabajo

tenga el beneficio esperado<sup>2</sup>.

Las neurociencias, como una de las más importantes ramas de la medicina, siempre se han caracterizado por un denso cuerpo de conocimientos clínicos y semiológicos, ahora apoyados por el desarrollo de nuevas técnicas diagnosticas. Estás pruebas tecnológicas no dejan de ser mecanismos artificiales para modificar la probabilidad diagnostica de determinada enfermedad, de ahí la importancia de conocer la auténtica correspondencia entre la validez interna de una prueba diagnostica y la realidad patológica, ya que no toda prueba diagnostica positiva significa padecer la enfermedad en estudio, ni toda prueba negativa la descarta. Muy pocas pruebas diagnosticas, quizá ninguna, identifican con total certeza si el paciente tiene o no la enfermedad.

El presente artículo pretende ser el primero de una serie de cuatro entregas que tienen como principal objetivo dar un repaso a los aspectos más importantes de las propiedades operativas de las pruebas diagnosticas para que sean consideradas con más frecuencia, y de modo explícito, en nuestra práctica clínica habitual.

### 1a. Pruebas para validar un test diagnóstico

La validez de una prueba diagnostica depende de su capacidad para detectar correctamente la presencia o ausencia de la enfermedad que se estudia, considerando como validez interna la sensibilidad y especificidad; y validez externa como valor predictivo positivo (VPP) y valor predictivo negativo (VPN).

Estos índices se obtienen a partir del análisis de una serie de pacientes a los que se les realizó una prueba diagnostica (prueba problema), comparándose los resultados con los de una prueba de superior rendimiento diagnóstico (prueba de referencia o *gold standard* en inglés). Los resultados obtenidos se expresan de la siguiente manera (tabla 1).

**Tabla 1.** Índices de validez de una prueba diagnostica.

	ENFERMO	SANO	
Resultado de la prueba diagnóstica que evaluamos	Positivo VP a	FP b	Total positivos a+b
	Negativo FN c	VN d	Total negativos c+d
	Total enfermos = prevalencia a+c	Total de sanos b+d	Total individuos a+b+c+d

VP. Verdaderos positivos: resultados positivos en sujetos enfermos

VN. verdaderos negativos: resultados negativos en sujetos sanos

FP. Falsos positivos: resultados positivos en sujetos sanos

FN. Falsos negativos: resultados negativos en sujetos enfermos

Sensibilidad:  $a/(a+c)$  Razón de probabilidad positiva S/1-E Valor predictivo positivo  $a/(a+b)$   
Especificidad:  $d/(b+d)$  Razón de probabilidad negativa E/1-S Valor predictivo negativo  $d/(c+d)$

**Sensibilidad:** es la probabilidad de que un individuo enfermo tenga un *test* positivo. La sensibilidad

indica la proporción del total de enfermos que el *test* es capaz de detectar. Es la proporción de verdaderos positivos. Enfermos con prueba positiva de entre todos los enfermos. Una prueba sensible es en especial útil cuando su resultado es negativo, es decir, inversamente relacionado con el valor de falsos negativos (FN). La fórmula es:  $S = a/(a+c)$ .

**Ejemplo 1:** ¿cuál es la sensibilidad de los signos de Kernig y de Brudzinski en el diagnóstico de la meningitis?: es del 5% en ambos signos<sup>3</sup>. Así, la ausencia de signos meníngeos en la exploración física no excluye la probabilidad de que un paciente tenga una meningitis.

**Ejemplo 2:** ¿cuál es la sensibilidad del TAC craneal en los servicios de urgencias hospitalarios?. Revisamos 500 TAC craneales consecutivos de pacientes con patología intracraneal de menos de 24 horas, encontrándose que sólo 36 de ellos presentaban anomalías estructurales, lo cual, según la tabla uno, representaría una sensibilidad del 7.2%. Lógicamente, este resultado se encuentra sesgado, ya que el TAC puede ser normal en las primeras 24 horas en la patología cerebrovascular isquémica aguda<sup>4</sup>.

**Ejemplo 3:** supongamos que disponemos de una nueva prueba diagnostica de *miastenia gravis*. Si la aplicamos a 100 miasténicos conocidos y es negativa en tres de ellos, ¿cuál es la sensibilidad de la prueba? R/ es del 0,97 (97%).

**Especificidad:** es la probabilidad de que un individuo sano tenga un resultado negativo en la prueba diagnostica; es decir, la proporción de verdaderos negativos, es decir, inversamente relacionado con el valor de falsos positivos (FP). La especificidad indica la proporción de individuos sanos confirmados como tales por el resultado negativo del *test* (tabla 1). La fórmula es:  $E = d/(b+d)$ .

**Ejemplo 4:** para estudiar la utilidad de una nueva prueba diagnostica para la enfermedad "E", se aplica a 100 enfermos con "E" y resulta positiva en 95 y negativa en 5 y, luego, a 100 sujetos sanos, sin "E", y resulta positiva en 3 y negativa en 97. ¿Qué podemos decir al respecto? R/ la sensibilidad de la prueba es de 0.95 y su especificidad es de 0.97. Hay que comparar estas cifras con las de otras pruebas antes de introducir la nueva. Así como el costo de la prueba diagnostica.

**Razón de probabilidad positiva:** compara la probabilidad de que un paciente que realmente esté enfermo presente un resultado positivo en una prueba diagnostica comparado con la probabilidad de que el mismo resultado positivo se presente en un individuo sano. Indica por cuánto se multiplica la

probabilidad *pre-test* (o prevalencia) una vez que el *test* resulte positivo. La fórmula es: RPP = S/1-E

*Ejemplo 5:* si una prueba diagnóstica tiene una sensibilidad del 80% y una especificidad del 90%, ¿cuál es la RPP? R/ S/1-E: 0.8/1-0.9 = 8.

*Razón de probabilidad negativa:* compara la probabilidad de que un paciente que realmente esté enfermo presente un resultado negativo positivo en una prueba diagnóstica comparado con la probabilidad de que el resultado negativo se presente en un individuo sano. Indica por cuánto se multiplica la probabilidad de estar libre de la enfermedad una vez que el *test* resulte negativo. La fórmula es: RP = E/1-S.

*Valor predictivo positivo (VPP):* se trata de la proporción de verdaderos positivos entre aquellos que han sido identificados como positivos en una prueba de *test*. Probabilidad de padecer la enfermedad cuando el resultado de la prueba es positivo. La fórmula es: VPP = a/(a+b).

*Ejemplo 6:* usted dispone de una prueba diagnóstica de neurosífilis y decide hacer un despistaje de la enfermedad en la población que atiende. Sabemos que la sensibilidad de la prueba es de 0.95 y la especificidad de 0.95. Si la prevalencia de la sífilis en su población es 1/1000 aproximadamente, ¿cuál será el valor predictivo positivo? (tabla 2). Así, según la fórmula del VPP (a/(a+b)) = 0.95/50 = 0.01866 = 1.87%.

*Valor predictivo negativo (VPN):* se trata de la proporción de verdaderos negativos entre aquellos que han sido identificados como negativos en un *test*. Probabilidad de no padecer la enfermedad cuando el resultado de la prueba es negativo. La fórmula es d/(c+d).

Tabla 2. Cálculo del valor predictivo positivo.

		REALIDAD		
		ENFERMO	SANO	
Resultado de la prueba diagnóstica que evaluamos	Positivo	VP a = 0.95	FP b = 49.05	Total positivos a+b = 50
	Negativo	FN c = 0.05	VN d = 949.95	Total negativos c+d = 950
		Total enfermos = prevalencia a+c = 1	Total sanos b+d = 999	Total individuos a+b+c+d = 1000

*Ejemplo 7:* la probabilidad de padecer estenosis carotídea-coronaria en hombres de más de 65 hasta 75 años, con amaurosis fugax-angina de pecho de esfuerzo típica, es mayor del 70 al 90%. A un paciente de estas características se le práctica un Doppler de troncos supra-aórticos-ECG de esfuerzo que se informa como negativo. La sensibilidad es del 80% y su especificidad del 90%. R/65,85%

¿Qué se podría deducir al respecto?:

No se puede excluir la estenosis carotídea-coronaria porque el valor predictivo negativo de la prueba es bajo (el lector puede utilizar la tabla uno y calcular el VPN).

*Ejemplo 8:* está usted evaluando una prueba diagnóstica y se encuentra con la siguiente tabla (tabla 3). A partir de ella y redondeando decimales calcule el VPP y el VPN. R/ El VPP es del 44% y el VPN es del 98%.

Tabla 3. Cálculo del VPP y del VPN.

Resultado de la prueba diagnóstica que evaluamos	REALIDAD		Total positivos a+b = 50	Total negativos c+d = 950
	ENFERMO	SANO		
Positivo	VP a = 0,95	FP b = 49,05		
Negativo	FN c = 0,05	VN d = 949,95		
		Total enfermos = prevalencia a+c = 1	Total de sanos b+d = 999	TOTAL INDIVIDUOS a+b+c+d = 1000

*Valor global (eficiencia) del test:* indica la proporción de resultados válidos entre el conjunto de resultados. En general el término eficiencia se utiliza para valorar eficacia o efectividad vs coste.

#### 1b. Curvas de características operativas para el receptor (COR).

El resultado de un *test* puede ser continuo (p. ej. niveles de glucemia en mg/dl) y entonces hay que decidir cuál se considerará como resultado positivo, hay que elegir un punto de corte.

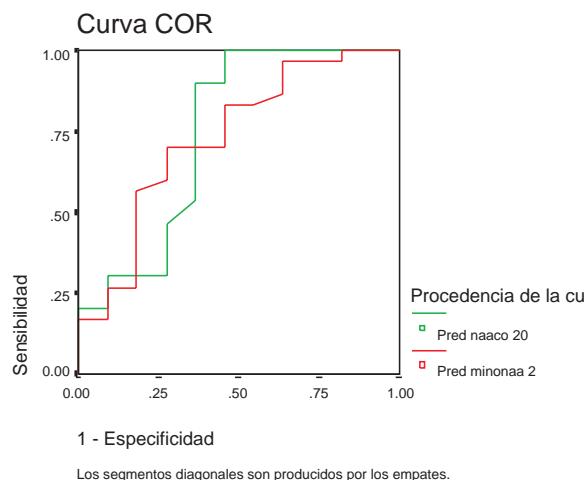
El punto de corte escogido determinará la sensibilidad y especificidad de la prueba (en el caso de la glucemia, si escogemos 70, la prueba será muy sensible y poco específica; si escogemos 140, será poco sensible y muy específica).

Para determinar el punto de corte se pueden utilizar las curvas de características operativas para el receptor (COR). Se seleccionan varios puntos de corte y se estima la sensibilidad y especificidad para cada uno de ellos. Después, se representa gráficamente la sensibilidad en función de (1-Especificidad). La prueba ideal se sitúa en el ángulo superior izquierdo (S y E = 1).

Una prueba sin ningún valor sigue la diagonal que va del ángulo inferior izquierdo al ángulo superior derecho (cada incremento de la sensibilidad se asocia a una pérdida de igual magnitud de especificidad).

De esta forma, la sensibilidad y la especificidad son valores interdependientes, de forma que si aumenta la sensibilidad disminuye la especificidad y viceversa. Si se adoptan criterios de diagnóstico muy

estrictos disminuye la sensibilidad (hay menos enfermos que cumplen estos criterios), y paralelamente aumenta la especificidad (pocos sanos cumplen estos criterios) figura 1<sup>5</sup>.



**Figura 1.** Curva de características operativas para el receptor (sólo a modo de ejemplo). Tomando de referencia 5: Valencia C. Análisis comparativo de la leucoaraiosis en la enfermedad de Alzheimer, la enfermedad de Binswanger y en los infartos lacunares. Tesis doctoral. Universidad Autónoma de Barcelona. 2002.

**Ejemplo 9:** la escala de ictus del *National Institute of Health* (NIHSS) es la escala más utilizada para valorar el estado neurológico de pacientes con enfermedad cerebrovascular isquémica. Esta escala puntuá del 0 al 42: una función normal puntúa como 0 y en un ictus grave la puntuación supera los 20 puntos<sup>6</sup>. Unos autores encontraron que al construir una curva COR, un punto de corte de 15 en la escala NIHSS en un paciente predecía con una sensibilidad del 90% y con una especificidad del 88% la presencia de una oclusión de la arteria cerebral media<sup>7</sup>.

#### 1c. Prevalencia y su relación con los índices de probabilidad.

La prevalencia es la proporción de sujetos de una colectividad que padecen un determinado fenómeno en un momento dado. No hace distinción entre casos antiguos y nuevos nos describe la situación en un momento determinado de tiempo. Es mayor a mayor incidencia y mayor duración de la enfermedad (tabla 1).

La prevalencia es la probabilidad de que un individuo de una población tenga una enfermedad en un momento dado. Este concepto se denomina también probabilidad *pre-test* o probabilidad *a priori*, es decir, la probabilidad de que un individuo con un determinado perfil presente una enfermedad determinada antes de realizar una prueba diagnóstica<sup>6,8</sup>.

**Ejemplo 10:** una mujer joven con historia de abortos a repetición que es trasladada a urgencias por una enfermedad cerebrovascular isquémica nos hará sospechar en un síndrome antifosfolípido o un *lupus*.

**Ejemplo 11:** no será la misma probabilidad de que un paciente con una lesión quística intracraneal tenga una neurocisticercosis, condicionada al hecho de que sea sudamericano, a de que sea oriundo de países mediterráneos (hidatidosis).

**Ejemplo 12:** no será la misma probabilidad de que un paciente con una lesión expansiva única intracraneal tenga un linfoma cerebral, condicionada al hecho de tener SIDA o a la de no tener SIDA.

Es importante recordar que los valores predictivos de un *test* son variables y dependen de la prevalencia de la enfermedad en la población. La sensibilidad y la especificidad son características propias del *test* y no se modifican con cambios en la prevalencia.

Si la prevalencia de la enfermedad aumenta, aumenta el valor predictivo positivo, mientras que disminuye el valor predictivo negativo.

Si la prevalencia de la enfermedad disminuye, aumenta el valor predictivo negativo y disminuye el valor predictivo positivo.

#### 1d. Aceptabilidad de un método diagnóstico

No existe un parámetro guía útil para todas las situaciones. La aceptabilidad de un *test* depende de la patología estudiada y de las condiciones reales en el medio y en la colectividad.

Si lo que interesa es detectar el mayor número posible de enfermos, se debe usar un *test* con alta sensibilidad. Así se nos escaparán pocos, aunque al precio de bastantes "falsos positivos". Elegiremos un *test* sensible cuando:

- La enfermedad sea grave y no pueda pasar desapercibida.
- La enfermedad sea tratable.
- Los resultados falsamente positivos no supongan un traumatismo psicológico en los individuos examinados.

Si lo que se quiere es "asegurar" el diagnóstico, debes usar un *test* cuya especificidad sea máxima.

*Utilizaremos un test lo más específico posible cuando:*

- La enfermedad sea importante, pero difícil de curar o incurable.
- Los resultados falsamente positivos puedan suponer un trauma psicológico para el individuo examinado.

- El tratamiento de los falsos positivos pudiera tener graves consecuencias  
1e. *Screening*/prueba de despistaje o cribado.  
Estrategia de detección precoz de la enfermedad que hace referencia a prevención secundaria y supone poder clasificar la población en bajo y alto riesgo. Lo ideal es aplicar primero un *test* muy sensible (detecta todos los casos posibles de enfermedad → se obtienen muchos FP y pocos FN) y en una segunda fase aplicar un *test* muy específico (se confirma el diagnóstico de esos posibles enfermos → se obtienen muy pocos FP).

En la puesta en marcha de un programa de screening deben tenerse en cuenta criterios dependientes de la enfermedad, del *test* y de la población diana.

### 1. Criterios dependientes de la enfermedad

- La enfermedad debe ser común y grave.
- Debe conocerse la historia natural de la enfermedad.
- El tratamiento, en el estado presintomático, debe reducir la morbilidad en mayor medida que el tratamiento después de la aparición de los síntomas.

### 2. Criterios dependientes del *test*

- De fácil aplicación.
- Coste razonable dentro del presupuesto de salud o que minimice los costes.
- Inocua para la población o exenta de efectos adversos.
- Confiabilidad o capacidad de repetición.
- Validez. Se refiere a la capacidad del *test* de medir lo que realmente deseamos medir. El concepto de validez incluye los de sensibilidad, especificidad y valor predictivo.

### 3. Criterios dependientes de la población diana

- El riesgo de ser afectado por la enfermedad debe ser alto. Prevalencia alta de la enfermedad.
- La información demográfica debe estar disponible en la comunidad.
- La comunidad debe sentir la necesidad de dicho problema de salud pública.

#### 1f. Fiabilidad de un método diagnóstico

La fiabilidad, llamada también repetibilidad, precisión o reproducibilidad se refiere a la estabilidad de la prueba, es decir, a la propiedad por la cual la variación observada es debida a la diversidad de valores

que adopta la variable y no al error, sea sistemático o aleatorio. La fiabilidad comprende el grado de acuerdo entre diferentes observadores (fiabilidad interobservador) y el grado de acuerdo para el mismo observador en momentos diferentes (fiabilidad intraobservador o *test-retest*).

La proporción de acuerdos (número de observaciones concordantes / número total de observaciones) no tienen en cuenta el azar, por lo que uno de los métodos que corrigen este defecto es el índice o coeficiente *kappa*, que mide el grado de concordancia no aleatoria interobservador o intraobservador. Este coeficiente depende de la prevalencia, por lo que está debería acompañar a la presentación de los resultados<sup>8,9</sup>.

*Ejemplo 13:* dos observadores analizan varias resonancias magnéticas cerebrales de pacientes con esclerosis múltiples en busca de lesiones hiperintensas en las secuencias potenciadas en T2.

*Ejemplo 14:* para establecer el nivel de concordancia interobservador entre neurorradiólogos y neurólogos a la hora de interpretar los TC craneales del servicio de urgencias de un hospital universitario, analizamos de forma aleatoria 55 TC craneales; utilizamos el índice *kappa* para cuantificar la concordancia, obteniendo los siguientes resultados según las diferentes fórmulas<sup>10,11</sup> del *kappa* (tabla 4).

**Tabla 4.** Contingencia: cálculo de los índices *kappa* según diferentes autores: el resultado muestra una concordancia óptima en las pruebas examinadas por neurológos y neurorradiólogos.

		REALIDAD	
		ENFERMO	SANO
Resultado de la prueba	Positivo	80	100
	Negativo	20	800

Los valores del índice *kappa* van del -1 al +1.

- 1 K = -1: discrepan más mediciones de lo esperado por el azar.
- 2 K = 0: concordancia debida al azar.
- 3 K = >0,6: grado de acuerdo aceptado por la comunidad científica.
- 4 K = >0,7: óptimo.
- 5 K = +1: concordancia completa.

## REFERENCIAS

1. Valencia C. El factor de impacto. Un deber y un derecho. *Rev Ecuat Neurol* 2004;13:55-6.
2. Valencia C. El reto científico de las neurociencias en

- Latinoamérica. *Arch Neurocienc Mex* 2004;9(4):187-8
3. Thomas KE, Hasbun R, Jekel J. The diagnostic accuracy of Kernig's sign, Brudzinski's sign, and nuchal rigidity in adults with suspected meningitis. *Clin Infect Dis* 2002;35:46-52.
  4. Muntané A, Valencia C, Martínez L. Sensibilidad de la tomografía computadorizada craneal en el servicio de urgencias de un hospital universitario de tercer nivel. *Neurología* 2005;20(9):501.
  5. Valencia C. Análisis comparativo de la leucoaraiosis en la enfermedad de Alzheimer, la enfermedad de Binswanger y en los infartos lacunares. [Tesis doctoral]. Barcelona: Departamento de Publicaciones, Universidad Autónoma de Barcelona. 2002.
  6. Montaner J, Álvarez-Sabin J. La escala de *ictus* del *National Institute of Health* (NIHSS) y su adaptación al español. *Neurología* 2006;21(4):192-202.
  7. Molina C, Álvarez-Sabin J, Montaner J, Abilleira S, Arenillas J. Predicting location of MCA occlusion by NIHSS score in candidates for thrombolytic therapy. *Cerebrovasc Dis* 2001;11(Suppl.4):65.
  8. Gómez de la Cámara A. Todo lo que el neurólogo quiere saber sobre sensibilidad, especificidad y valores predictivos. *Neurología* 2003;18(spl 2):11-8.
  9. Muntané A, Valencia C. Análisis comparativo de los diagnósticos tomográficos en pacientes neurológicos y neuroquirúrgicos. *Neurología* 2006.
  10. Byrt T, Bishop J, Carlin JB. Bias. Prevalence and kappa. *J Clin Epidemiol* 1993;46(5):423-9.
  11. Martínez González MA, de Irala Estévez J. *Métodos en salud pública*. 1<sup>a</sup> Edición. Pamplona: Newbook ediciones; 1999.