

# Evaluación de los procedimientos diagnósticos

## Aspectos metodológicos

**RICARDO H. FESCINA**

ESPECIALISTA EN SALUD PÚBLICA

EX SECRETARIO ACADÉMICO DE FLASOG

### Resumen

Se observó que el informe de un estudio radiológico enviado a diferentes expertos no coincidía en sus apreciaciones y así nació la necesidad de evaluar la capacidad que tiene la prueba de diagnosticar los verdaderos enfermos y los verdaderos sanos, descartando los enfermos y los sanos falsos.

### Palabras clave

Especificidad; sensibilidad; valores predictivo positivo y negativo.

## Introducción

El gran desarrollo tecnológico de los últimos decenios involucró a las ciencias médicas y a la perinatología en particular. Se han incorporado, con frecuencia, a la práctica clínica nuevos procedimientos diagnósticos, sin que se abandonaran los usados anteriormente. La simultaneidad de exámenes, a veces molestos y riesgosos para la misma finalidad, habitualmente arroja resultados confusos, cuando no contradictorios, a la vez que se multiplican los costos asistenciales.

Considerando la natural tendencia humana a creer más en los procedimientos que en los criterios, puede ser beneficioso recordar la metodología que debe seguirse en la evaluación de cualquier procedimiento diagnóstico.

Diagnosticar significa *corroborar la existencia de un estado mórbido particular*. En la práctica clínica se llega al diagnóstico recorriendo dos etapas diferenciadas. En la primera se establece una presunción, sospecha o hipótesis de existencia de la enfermedad, a la que se arriba en base a la jerarquización de síntomas y signos clínicos y a su asociación lógica con patologías conocidas. La segunda etapa, se dirige a verificar si esa presunción, sospecha o hipótesis, corresponde a la verdad. Con tal fin, se procede a discriminar mediante:

- pruebas y exámenes, que si son *positivos* indican confirmación de la enfermedad;
- pruebas y exámenes que si son *negativos* descartan la presencia de enfermedades de sintomatología parecida.

Por usos y costumbres, existe consenso en asociar el resultado positivo de un examen con la presencia de enfermedad y el resultado negativo con ausencia.

Los datos del examen clínico deben constituir el pre-requisito básico en la orientación de cuáles pruebas o procedimientos corresponde indicar y en la interpretación de resultados. De esta manera, la información adicional da mayor seguridad y acrecienta el conocimiento clínico, pero no lo sustituye.

## Desarrollo de un procedimiento diagnóstico

En 1947, Yerushalmi, se sorprendió de los muy variados y antagónicos diagnósticos a que arribaban radiólogos muy experimentados al examinar las mismas radiografías de tórax y advirtió que no existía medida de la veracidad o error de los exámenes.

		Patología confirmada		Sentido del análisis
		SI	NO	
Resultado del procedimiento diagnóstico	+	a verdaderos positivos	b falsos positivos	↓
	-	c falsos negativos	d verdaderos negativos	

a) Casos con diagnóstico Positivo en Presencia de Enfermedad (Verdaderos Positivos)  
 b) Casos con diagnóstico Positivo en Ausencia de Enfermedad (Falsos Positivos)  
 c) Casos con diagnóstico Negativo en Presencia de Enfermedad (Falsos Negativos)  
 d) Casos con diagnóstico Negativo en Ausencia de Enfermedad (Verdaderos Negativos)

FIGURA 1. TABLA DE DOBLE ENTRADA. ANÁLISIS EN SENTIDO VERTICAL

Con esta idea, registró los diagnósticos y siguió los casos hasta confirmar la presencia o ausencia de la enfermedad por la anatomía

patológica obtenida en intervenciones quirúrgicas o en autopsias. Mediante esta metodología, definió los lineamientos para la evaluación del grado de certeza de los procedimientos o pruebas diagnósticas. Al realizar pruebas o exámenes, surgen de inmediato dos preguntas:

- si la enfermedad está *presente*, ¿cuál es la probabilidad de que el resultado sea positivo?, y
- si la enfermedad está *ausente*, ¿cuál es la probabilidad de que el resultado sea negativo?

La contestación al primer interrogante determina la *sensibilidad* y al segundo, la *especificidad*.

**Sensibilidad** es la capacidad del procedimiento de efectuar diagnósticos correctos de enfermedad cuando ésta está presente (verdaderos positivos o enfermos).

**Especificidad** es la capacidad del procedimiento de efectuar diagnósticos correctos de ausencia de una enfermedad cuando ésta no está presente (verdaderos negativos o sin esa patología).

Como estos datos son cualitativos, ya que indican presencia o ausencia confirmada de la enfermedad y resultado positivo o negativo de un procedimiento diagnóstico, se usan las tablas de doble entrada (matriz de decisión) para su presentación numérica (**FIGURA 1**).

La sensibilidad se mide por la *proporción de individuos en que el procedimiento resultó positivo, en relación al total de los que tienen la enfermedad*. Es el análisis de la primera columna de la **FIGURA 1** (sentido vertical).

$$\text{Sensibilidad} = \frac{\text{Verdaderos Positivos}}{\text{Total de Enfermos}} \times 100 = \frac{a}{a+c} \times 100$$

La especificidad se mide por la *proporción de individuos con diagnóstico negativo y sin la patología, en relación al total de los que no tienen esa enfermedad* (análisis de la segunda columna de la FIGURA 1).

$$\text{Especificidad} = \frac{\text{Verdaderos Negativos}}{\text{Total sin patología}} \times 100 = \frac{d}{b+d} \times 100$$

La sensibilidad considera los aciertos positivos, en tanto que la especificidad considera los aciertos negativos.

Como medida de resumen para evaluar la eficacia de las pruebas se puede usar el índice de Youden. Este no es afectado por la prevalencia de la enfermedad en la muestra estudiada, ya que en su cálculo interviene la sensibilidad y especificidad. Tiene el inconveniente de que al combinar el resultado en un valor único, se deja de conocer si el test es bueno en sensibilidad o especificidad. Se debe tener en cuenta que el valor de la sensibilidad y especificidad considerado para el cálculo, no se expresa en porcentaje sino en frecuencia relativa (ejemplo: 80 % = 0,8).

$$Y = (S + E) - 1$$

El valor oscila entre 0 y 1, siendo una prueba sin error de clasificación.

La misma matriz de decisión permite evaluar los errores en uno y otro sentido. Así, desde el punto de vista de la prueba, son llamados *falsos negativos* cuando el procedimiento resulta negativo pero la enfermedad está presente. Este tipo de error se mide mediante la relación de falsos negativos.

$$\text{Relación de Falsos Negativos} = \frac{\text{Falsos Negativos}}{\text{Total de Enfermos}} \times 100 = \frac{c}{a+c} \times 100$$

Dicho de otra manera, esta relación constituye el complemento de la sensibilidad para llegar a 100 %. Por ello, se puede calcular:

$$100 - \text{Sensibilidad} = \text{Relación de Falsos Negativos}$$

Se llaman *falsos positivos* cuando el procedimiento resulta positivo pero la enfermedad está ausente.

Este tipo de error del procedimiento se mide mediante la relación de falsos positivos b.

$$\text{Relación de Falsos Positivos} = \frac{\text{Falsos Positivos}}{\text{Total Sin Patología}} \times 100 = \frac{c}{b+d} \times 100$$

Representa el complemento a 100 % de la especificidad, por lo que se puede calcular:

$$100 - \text{Especificidad} = \text{Relación de Falsos Positivos}$$

Cuanto más próximo a 100 % sea la sensibilidad y a 100 % la especificidad, mayor es la capacidad de la prueba para discriminar los que tienen de los que no tienen la enfermedad. Con excepción de ejemplos

muy forzados (diagnóstico de muerte por registro prolongado del EEG), siempre debe considerarse la existencia de errores en ambos sentidos.

La **FIGURA 2** ilustra lo esperable de una prueba ideal, lo que ocurriría si las distribuciones con y sin patología no superponen sus valores, hecho que no ocurre en la gran mayoría de las pruebas (**FIGURA 3**).

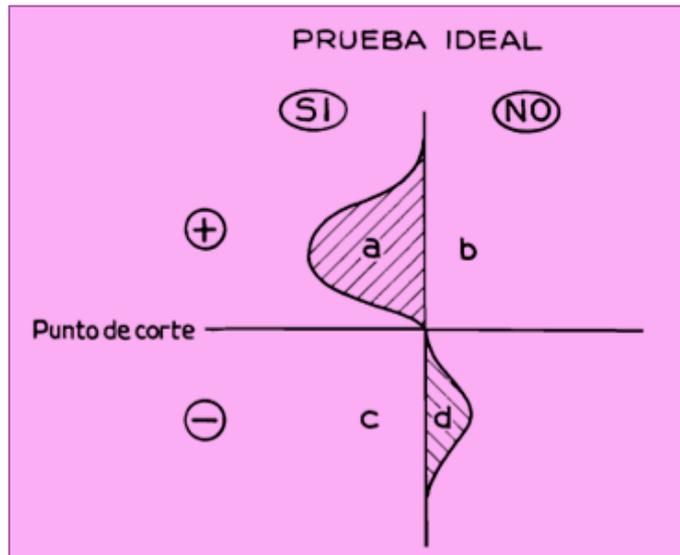


FIGURA 2. PRUEBA IDEAL. SENSIBILIDAD Y ESPECIFICIDAD 100 %, SÓLO CASOS CORRECTOS DE PATOLOGÍA PRESENTE Y AUSENTE

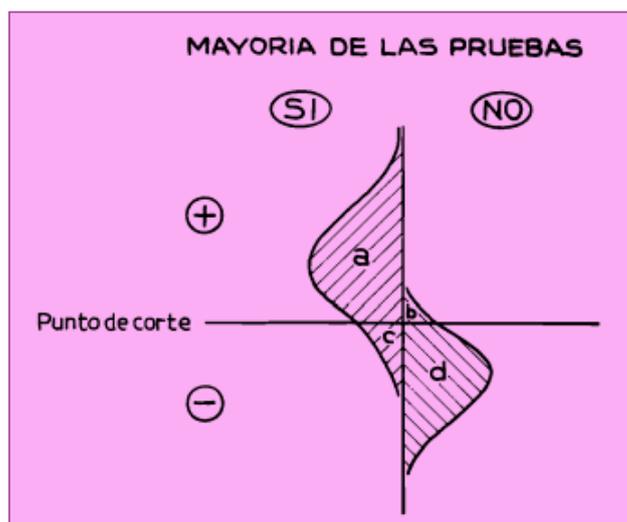


FIGURA 3. MAYORÍA DE LAS PRUEBAS. SENSIBILIDAD Y ESPECIFICIDAD MENOR DE 100 %. HAY ERRORES DIAGNÓSTICOS, FALSOS POSITIVOS Y NEGATIVOS

Por ejemplo, si interesa determinar el umbral o nivel de glicemia posprandial que discrimine a las gestantes en diabéticas o no, es posible calcular la sensibilidad y especificidad para distintos valores de corte, de acuerdo a la conveniencia del investigador (FIGURA 4).

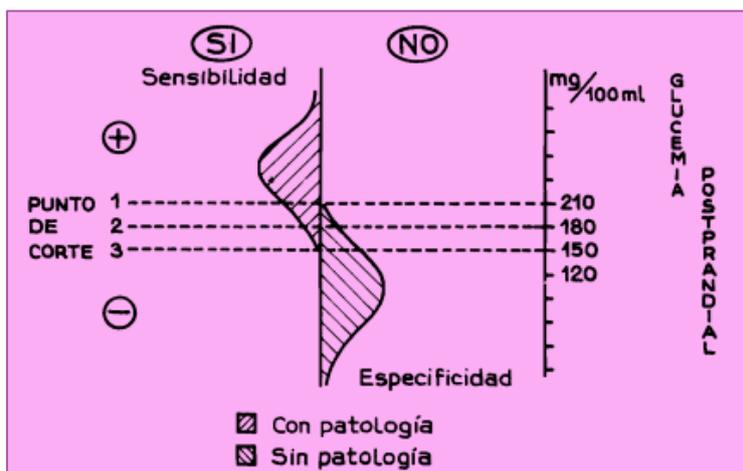


FIGURA 4. ELECCIÓN DEL PUNTO DE CORTE. ES UNA REGLA QUE, A MEDIDA QUE SE QUIERE AUMENTAR LA SENSIBILIDAD, SE PIERDE EN ESPECIFICIDAD Y VICEVERSA, POR LO QUE AMBOS VALORES DEBEN EXAMINARSE EN CONJUNTO

Así, seleccionando el nivel 1, la especificidad es próxima a 100 % (se separa bien a las no diabéticas) pero la sensibilidad es baja, ya que escapan al diagnóstico cierta proporción de enfermos (hay falsos negativos).

Eligiendo el nivel 3, la sensibilidad es próxima a 100 %, se captan prácticamente todos los enfermos, pero se discriminan mal los no diabéticos (especificidad baja).

En esta etapa el investigador que conoce quiénes tienen la enfermedad, determina a su conveniencia la sensibilidad y especificidad, seleccionando el nivel o valor de corte.

La sensibilidad y especificidad no varían con la prevalencia de la enfermedad, pero sí con el valor de corte o umbral.

Para el desarrollo de un procedimiento diagnóstico, es imprescindible contar con la confirmación final, fuera de toda duda, de la existencia o no de la patología y se requiere del cálculo de la sensibilidad y especificidad. La investigación original está dirigida primordialmente a emplear el procedimiento en individuos en que se puede conocer finalmente el diagnóstico correcto.

En la práctica, estos procedimientos se usan de manera muy distinta de cómo se desarrollan. El médico los utiliza en pacientes con diagnósticos no conocidos y el propósito de la indicación es aproximarse a la verdadera condición del paciente. Se trabaja en condiciones de incertidumbre y las pruebas se utilizan para predecir la presencia o ausencia de enfermedad, por lo que existe siempre un margen de error. Para afirmar si existe o no enfermedad, empleando un procedimiento diagnóstico, hay que considerar:

		Patología confirmada	
		SI	NO
Resultado del procedimiento diagnóstico	+	a	b
	-	c	d

→ Sentido del análisis

FIGURA 5. TABLA DE DOBLE ENTRADA. ANÁLISIS EN SENTIDO HORIZONTAL

- Si el procedimiento resulta positivo, ¿cuál es la probabilidad de que la enfermedad esté verdaderamente presente (de que se confirme la enfermedad)?
- Si resulta negativo, ¿cuál es la probabilidad de que la enfermedad esté verdaderamente ausente (de que se descarte la enfermedad)?

La misma matriz de decisión ya elaborada permite la contestación de ambas preguntas pero, en este caso, el análisis se realiza por filas (sentido horizontal) (FIGURA 5).

### Valor predictivo de la prueba positiva (VPPP)

Es la probabilidad de que la enfermedad esté presente cuando el resultado del procedimiento es positivo (análisis de la primera fila, sentido horizontal). Se expresa como el porcentaje de verdaderos enfermos de todos los que tuvieron pruebas positivas.

$$\text{V.P.P.} = \frac{\text{Verdaderos Positivos}}{\text{Total de las Pruebas Positivas}} \times 100 = \frac{a}{a+b} \times 100$$

### Valor predictivo de la prueba negativa (VPPN)

Es la probabilidad de que la enfermedad no esté presente cuando el resultado del procedimiento es negativo (análisis de la segunda fila). Se expresa como el porcentaje de los sin patología de todos los que tuvieron pruebas negativas.

$$\text{V.P.N.} = \frac{\text{Verdaderos Negativos}}{\text{Total de las Pruebas Negativas}} \times 100 = \frac{a}{c+d} \times 100$$

Se observa que hay casos en que la prueba indica como enfermos y que no lo son (falsos positivos de la predicción) y casos en que la prueba es negativa, pero la enfermedad está verdaderamente presente (falsos negativos de la predicción). Ambos tipos de errores se miden en las relaciones:

### **Falsos positivos de la prueba positiva (FPPP)**

Son los individuos sin la enfermedad que el procedimiento clasifica como enfermos (b) (análisis de la primera línea horizontal), en relación con el total de pruebas positivas.

$$\text{F.P.P.} = \frac{\text{Falsos Positivos}}{\text{Total de las Pruebas Positivas}} \times 100 = \frac{d}{c+d} \times 100$$

Es el complemento a 100 % del valor predictivo de la prueba positiva (VPPP).

### **Falsos negativos de la prueba negativa (FNPN)**

Son los individuos enfermos que el procedimiento clasificó sin la enfermedad (e) (análisis de la segunda línea horizontal) en relación al total de pruebas negativas.

$$\text{F.N.P.N.} = \frac{\text{Falsos Negativos}}{\text{Total de las Pruebas Negativas}} \times 100 = \frac{c}{c+d} \times 100$$

Es el complemento a 100 % del valor predictivo de la prueba negativa (VPPN).

También se emplea la exactitud que es un índice que considera, en forma conjunta, las predicciones correctas del procedimiento, positivas y negativas.

$$\text{Exactitud} = \frac{\text{Verdad positivos} + \text{Verdad negativos}}{\text{total indiv. con y sin la enfermedad}} \times 100 = \frac{a+d}{a+b+c+d} \times 100$$

Este índice permite comparar diferentes procedimientos entre sí, pero tiene la desventaja de adjudicar igual valor a los falsos positivos y negativos, errores que pueden tener muy diferente repercusión en la clínica.

Es importante destacar el valor predictivo de una prueba y la exactitud, que se modifican sustancialmente al variar la *prevalencia*(\*).

Con alta prevalencia cabe esperar alta predicción (ejemplo: epidemia); con baja prevalencia disminuye la predicción, aunque se mantenga la misma sensibilidad y especificidad.

El ejemplo siguiente (**FIGURA 6**) muestra cómo la prevalencia hace variar el resultado predictivo de una prueba sin que se modifiquen la sensibilidad ni la especificidad.

Se observa que al disminuir la prevalencia, el VPPP disminuye en tanto que el VPPN aumenta. Se deduce claramente por qué no debe hacerse la

comparación directa de los valores predictivos de procedimientos, desarrollados en grupos de prevalencia diferente. El teorema de Bayes, de la probabilidad condicionada, permite ajustar el valor predictivo a una prevalencia dada de la enfermedad.

$$V.P.P. = \frac{\text{Prevalencia de la enfermedad} \times \text{Sensibilidad}}{\text{Preval.} \times \text{Sensib.} + (1 - \text{prevalencia}) \times (1 - \text{especificidad})}$$

\* Prevalencia: Corresponde al número de enfermos en relación al total de casos considerados

	SI	NO	
+	40	2	42
-	10	23	33
	50	25	75

Prevalencia =  $\frac{50}{75} \times 100 = 67\%$

Sensibilidad =  $\frac{40}{50} \times 100 = 80\%$

Especificidad =  $\frac{23}{25} \times 100 = 92\%$

V.P.P.R. =  $\frac{40}{42} \times 100 = 95\%$

V.P.P.N. =  $\frac{23}{33} \times 100 = 70\%$

	SI	NO	
+	40	20	60
-	10	230	240
	50	250	300

Prevalencia =  $\frac{50}{300} \times 100 = 17\%$

Sensibilidad =  $\frac{40}{50} \times 100 = 80\%$

Especificidad =  $\frac{230}{250} \times 100 = 92\%$

V.P.P.R. =  $\frac{40}{60} \times 100 = 67\%$

V.P.P.N. =  $\frac{230}{240} \times 100 = 96\%$

FIGURA 6. CAMBIO DE LOS VALORES PREDICTIVOS CON LA MODIFICACIÓN DE LA PREVALENCIA DEL EVENTO CONSIDERADO

## Medida del acuerdo entre el procedimiento diagnóstico y la presencia o ausencia de patología

Es factible diagnosticar correctamente una patología sólo por azar y erróneamente imputarle el éxito a un procedimiento diagnóstico utilizado. Para determinar si el grado de acuerdo entre el resultado obtenido por un procedimiento diagnóstico y la verdad (presencia o ausencia de patología) excede lo que se esperaría por azar, Cohen describió un índice que lo

llamó Kappa. Este índice es la razón entre la exactitud observada que sobrepasa a la esperada por azar, sobre la exactitud máxima alcanzable que sobrepasa a la esperada por azar.

$$\text{Índice Kappa} = \frac{\text{exactitud observada} - \text{exactitud esperada}}{1 - \text{exactitud esperada}}$$

La forma de calcular la exactitud observada fue explicada más arriba. Para calcular la esperada se procede igual, pero con los valores esperados por el azar en las casillas de los verdaderos positivos (a) y negativos (d) (FIGURA 7). Para los verdaderos positivos, esto se logra multiplicando el total de la primera fila (a + b) con el total de la primera columna (a + e) y se lo divide por el número total de casos (N): N.

$$\frac{(a+b)(a+c)}{N}$$

Para los verdaderos negativos (d) se multiplica el total de la segunda fila (e + d) con el total de la segunda columna (b + d) y se divide por el número total de casos (N):

$$\frac{(c+d)(b+d)}{N}$$

Luego de obtener los valores esperados en las casillas a y d, se calcula la exactitud esperada por azar:

$$\frac{(a \text{ esperada} + d \text{ esperada})}{N}$$

	SI	NO	
+	a	b	a + b
-	c	d	c + d
	a + c	b + d	a + b + c + d = N

FIGURA 7. NOTACIÓN USADA PARA LAS FRECUENCIAS ABSOLUTAS EN LA TABLA DE DOBLE ENTRADA

El valor máximo del índice es 1. Cuando los valores oscilan entre 0.75 y 1, se denomina *acuerdo excelente*; si el resultado está entre 0.4 y 0.74, reflejan de *regular a buen acuerdo*; los valores menores de 0.4 sugieren que la concordancia es debida exclusivamente al azar.

Igual que los valores predictivos, este índice es influenciado por la prevalencia del fenómeno.

El valor así obtenido es puntual e importa conocer su intervalo de confianza a 95 % para darle más peso al resultado hallado. Para ello, se calcula el error estándar del índice de Kappa y luego se lo multiplica por 1.96 (valor de Z para probabilidad de 0,025); este resultado se suma y resta al índice de Kappa previamente obtenido.

$$S.e.K = \frac{1}{(1-p_e)\sqrt{n}} \sqrt{p_e + p_e^2 - \sum_{i=1}^K p_i \cdot p_i \cdot (p_i + p_i)}$$

En casos de tablas de 2 x 2 esta fórmula se puede expresar de la siguiente manera:

$$S.e.K = \frac{1}{(1-\rho_e) \sqrt{n}} \sqrt{\rho_e + \rho_e^2 - \rho_1 \cdot \rho_1 (\rho_1 + \rho_1) + \rho_2 \cdot \rho_2 (\rho_2 + \rho_2)}$$

Se debe tener presente que en este caso lo que entra en las celdas a, b, e y d es la proporción de los sujetos (FIGURA 8) y no el número absoluto como en los ejemplos anteriores. Es decir:

$$P_{11} = \frac{a}{N}; P_{12} = \frac{b}{N}, \text{ y así sucesivamente.}$$

## Prueba del procedimiento diagnóstico

Ya desarrollado el procedimiento diagnóstico en su parte teórica, deberá probarse con sentido prospectivo. Es importante que se consideren grupos de diferente prevalencia de la enfermedad y que se verifique si los resultados son replicables por otros grupos de investigación en diferentes lugares (FIGURA 9). De esta manera, se podrán conocer alcances y limitaciones y estudiar la confiabilidad y precisión en distintas circunstancias (complejidad asistencial, calidad del personal técnico, distintos aparatos, asociación con otras patologías, etc.).

	<b>SI</b>	<b>NO</b>	
<b>+</b>	P <sub>11</sub>	P <sub>12</sub>	p <sub>1</sub>
<b>-</b>	P <sub>21</sub>	P <sub>22</sub>	p <sub>2</sub>
	p.1	p.2	1

FIGURA 8. NOTACIÓN USADA PARA LAS PROBABILIDADES EN LA TABLA DE DOBLE ENTRADA

Etapas de desarrollo de procedimientos diagnósticos			
Etapa	Desarrollo	Evaluación	Observaciones
1ª	Mejoramiento técnico en las medidas	Exactitud y replicabilidad de las medidas	Variación intra e interobservaciones
2ª	Definición de prueba anormal	Eficacia: sensibil. especif. V.P.P. + y -	Asociación de prueba + con malos resultados
3ª	Intervención clínica	cambio de los resultados	No asociación de prueba + con malos resultados
4ª	Difusión	Efectividad: el cambio es mejor que lo que había	Costo personal Aspectos éticos

FIGURA 9. ETAPAS QUE DEBE SEGUIR UN NUEVO PROCEDIMIENTO DIAGNÓSTICO ANTES DE SU MASIFICACIÓN

Es conveniente que el procedimiento sea también replicado por el mismo grupo que lo desarrolló y que exista intercambio de información entre los equipos de investigación. No resulta conveniente la difusión y el uso extensivo sin haber cumplido satisfactoriamente esta fase de la evaluación.

## Utilización del procedimiento diagnóstico

Mediante la aplicación de un procedimiento diagnóstico nuevo, se pretende mejorar los resultados en mortalidad y morbilidad atribuibles a la enfermedad (FIGURA 9). El principal requerimiento que se solicita es la capacidad para cambiar los resultados finales: un diagnóstico correcto y oportuno lleva más fácilmente a una intervención exitosa. Para evaluar la

eficacia del uso del procedimiento, se considerarán factores diferentes a las etapas anteriores: características de infraestructura, disponibilidad de equipos y personal que permitan su factibilidad; el costo financiero y humano (agresividad, dolor, etc.) y los aspectos éticos en la distribución de sus beneficios (para todos, para quienes puedan financiarlo, para quienes lo necesiten, etc.).

En esta etapa de evaluación del uso del procedimiento, se tendrían que contestar las siguientes preguntas:

- El procedimiento desarrollado, ¿es mejor que los anteriores?
- ¿Es más fácil, más rápido, más sencillo, etc.?
- ¿Mejora el manejo y el tratamiento de los pacientes?
- ¿Es el tratamiento más oportuno, más específico, menos doloroso, etc.?
- ¿Permite una recuperación más rápida?
- ¿Con menos días de enfermedad, menos costo, etc.?
- ¿Contribuye a mejorar la salud de la población?

## Propósitos de los procedimientos diagnósticos

Estas pruebas se emplean especialmente con dos motivos: a) detectar la enfermedad; b) confirmarla.

### Prueba para detección de la enfermedad (tamizaje)

Se trata de emplear el procedimiento diagnóstico en la población aparentemente sana, que no presenta síntomas ni signos clínicos evidentes de enfermedad, para despistar casos en estadios iniciales. Son ejemplos conocidos la investigación de glucemia en ayunas para detectar

la diabetes clínica, el examen de Papanicolau para detectar cáncer genital. Para estos rastreos (*screening*), se requieren pruebas que tengan alta sensibilidad (que capten el mayor número posible de enfermos) y que sean fáciles de realizar, rápidas, no invasivas y de bajo costo. Se acepta cierto número de errores o de falsos positivos de la prueba. Por ello, los procedimientos de *screening* cuando son positivos, requieren siempre de otras pruebas para confirmar la enfermedad. Los casos con resultado positivo no deben considerarse directamente como enfermos, sino como individuos con más probabilidad de tener la enfermedad. Debe existir mesura al solicitar las pruebas confirmatorias que irán de lo más simple a lo más complejo (FIGURA 10).

### **Pruebas para confirmar la enfermedad**

Se emplean en situaciones donde existe fuerte sospecha clínica de presencia de la enfermedad. Su propósito será verificarla. Son procedimientos habitualmente más molestos o agresivos, como por ejemplo la broncoscopía o la biopsia de cérvix uterino, en caso de que se sospeche neoplasia. Estas pruebas deberán tener muy alta especificidad, de tal manera que no existan casi falsos positivos de las pruebas, por lo que se requiere alto valor predictivo positivo. En general, resultan procedimientos costosos, manejados por especialistas, se usan cuando existe cierta presunción diagnóstica y se pretende una mayor confirmación para realizar procedimientos quirúrgicos de riesgo o tratamientos agresivos (FIGURA 10).

	<b>TAMIZAJE (Screening)</b>	<b>CONFIRMACION (Diagnósticas)</b>
<b>OBJETIVOS</b>	Separar grupos de alto riesgo	Separar normales de anormales.
<b>CARACTERÍSTICAS</b>	Sencilla	Más compleja.
	Sin peligros	Más efectos secundarios.
	Aceptable por el usuario.	Aceptable para los de riesgo elevado.
	Alta sensibilidad	Alta sensibilidad.
	Baja especificidad	Gran especificidad.
	No invasiva	Invasiva.
	Barata	Cara.
	Iniciativa del equipo de salud.	Iniciativa del individuo enfermo.

FIGURA 10. OBJETIVOS Y CARACTERÍSTICAS DE LAS DISTINTAS PRUEBAS

## Combinación de pruebas

En la práctica, habitualmente, se usa más de un procedimiento, aunque esto dependa de cada situación particular y de la disponibilidad de pruebas. En general, el proceso diagnóstico se realiza por etapas y, según los resultados obtenidos con unos procedimientos, se indicarán otros y aplicándolos en forma sucesiva, lo que se llama «serie» o «tandem» de pruebas, en tanto no se arribe a un diagnóstico.

Es infrecuente que una sola prueba diagnóstica cumpla satisfactoriamente el propósito de detección y confirmación.

Se emplean pruebas de alta sensibilidad para hallar a los probables enfermos y los casos con resultado positivo son seguidos con pruebas de alta especificidad para confirmar la existencia de enfermedad, obviando los falsos positivos. Los buenos resultados dependen de una correcta y oportuna secuencia de las pruebas.

En términos generales, cuando se aplican las pruebas diagnósticas de detección en serie, se gana en especificidad y cuando se usan en paralelo (todas a todos los casos), se gana en sensibilidad.

Como en la actualidad proliferan las pruebas paraclínicas, es aconsejable que, para la aplicación rutinaria de un procedimiento, se analicen sus características de especificidad, sensibilidad, valor predictivo, etc. Es de importancia conocer los trabajos originales y las características de las poblaciones estudiadas. Conviene establecer cierta normalización para la aplicación lógica de una «batería de pruebas» de acuerdo a las situaciones planteadas, bien sea detectar la patología en pacientes que aparentemente no la tienen, confirmarla o excluirla cuando existen sospechas de que esté presente. En estos casos, hay que seleccionar los procedimientos menos agresivos para el paciente y que, a la vez, aporten mayor seguridad en la información. Estas selecciones no deben ser decisiones personales, sino que estarán previstas y normatizadas por el equipo de salud, lo que contribuye a su evaluación. Se debe evitar la proliferación de pruebas para un mismo propósito y que se difundan sin haber completado su evaluación.

Para una mejor interpretación del uso de esta metodología, se presenta el siguiente ejercicio. Trate de resolverlo y de interpretar los resultados.

Se estudiaron 163 embarazadas con amenorrea confiable y feto único a las que se midió semanalmente la altura uterina. El diagnóstico de hipocrecimiento se efectuaba cuando la evolución de la altura uterina mostraba un perfil de retardo (disminución de la velocidad de crecimiento con uno o más valores por debajo del percentil elegido). Los recién nacidos se clasificaron por la curva de peso neonatal según edad gestacional construida en el CLAP. Se obtuvo la calificación de pequeño cuando el neonato, al nacer, presentaba un peso menor que el valor del percentil 10 correspondiente a la edad.

De los 163 nacidos, 42 tuvieron diagnóstico confirmado de pequeño para la edad gestacional (PEG) y 121 de peso adecuado para la edad.

Si se elegía como nivel discriminante del método (altura uterina), el percentil 10 se diagnosticaban antenatalmente, solamente 22 retardos. De los 121 pesos adecuados se diagnosticaron como tal 111.

Con estos datos, construya la matriz de decisión y calcule la sensibilidad, especificidad, prevalencia, valor predictivo de la prueba positiva y valor predictivo de la prueba negativa (CUADRO 1).

Altura uterina	SI	NO	
Menor que P <sub>10</sub>			
Igual o mayor que P <sub>10</sub>			
Sensibilidad = Especificidad =		Prevalencia = V.P.P. Positiva = V.P.P. Negativa =	

CUADRO 1. DIAGNÓSTICO ANTENATAL DE PEG

Altura uterina	SI	NO	
Menor que P <sub>25</sub>			
Igual o mayor que P <sub>25</sub>			
Sensibilidad = Especificidad =		Prevalencia = V.P.P. Positiva = V.P.P. Negativa =	

CUADRO 2. DIAGNÓSTICO ANTENATAL DE PEG

Si el nivel de discriminación del método (altura uterina) es el P25 se diagnosticarán 29 pequeños. De los 121 nacidos de peso adecuado se diagnostican 94.

Realice los mismos cálculos que efectuó cuando usó el P10.

Compare, interprete y tome una decisión acerca de cuál de ambos percentiles de altura uterina (P10 o P25) es el que conviene elegir.

## Resultados del ejercicio

Utilizando el P10 de la altura uterina para la edad gestacional se podrá diagnosticar 1 de cada 2 RCI (sensibilidad 52 %) y afirmar con 8 % de error (especificidad 92 %), que el feto no es pequeño para la edad gestacional. La altura uterina es así un excelente método para separar al grupo no afectado por su alta especificidad.<sup>(\*)</sup>

En el uso clínico, además de la sensibilidad y especificidad de un método, es necesario conocer el valor predictivo que tiene una prueba, en nuestro caso, la medida de la altura uterina. El valor predictivo dependerá de la prevalencia de la patología, siendo más elevado este valor en poblaciones donde la prevalencia de PEG es mayor y viceversa. El valor predictivo de la prueba positiva en el presente trabajo (prevalencia de PEG 25,8 %) resulta de 69 %.

Considerando como límite discriminante el percentil 25 de la altura uterina, se mejora la sensibilidad (69 %) pero a expensas de una disminución de la especificidad y del valor predictivo positivo.

**(\*)** Datos tomados de: FESCINA RH, QUEVEDO C, MARTELL M, NIETO F, SCHWARCZ R. Altura uterina como método para predecir en crecimiento fetal. *Bol of Saint Panam*. 1984;96(5):377-386. <<https://iris.paho.org/handle/10665.2/17004>>

## Bibliografía

BENNET BM. On comparisons of sensitivity, specificity and predictive value of a number of diagnostic procedures. *Biometrics*. 1972;28(3):793-800.

BELITZKY R. Desarrollo, prueba y evaluación de los procedimientos diagnósticos. Guía de discusión. *Publicación Científica CLAP*. 1983;995.

FEINSTEIN AR. *Clinical Biostatistics*, Saint Louis: Mosby, 1977; pp. 214-226.

FESCINA RH, BELITZKY R, SIMINI F. Evaluación de los procedimientos diagnósticos. *Publicación Científica CLAP*. 1983;999.

FLEISS JL. *Statistical Methods for Rates and Proportions*. Second Ed. New York: Wiley & Sons, 1981.

MANTEL N. Evaluation of a class of diagnostic test. *Biometrics*. 1951;7:240-246.

SIMINI F. *Eficacia de los Procedimientos Diagnósticos. Curso de Metodología Científica del CLAP*. Montevideo, 1979.

