

## Valor predictivo de signos y síntomas respiratorios en el diagnóstico de atopia

Saúl Oswaldo Lugo Reyes,<sup>\*,\*\*</sup> María de la Luz García Cruz,<sup>\*\*</sup> Luis Manuel Terán Juárez<sup>\*\*</sup>

### RESUMEN

**Antecedentes:** En la evaluación del paciente con molestias respiratorias, tradicionalmente se considera sugestiva de origen alérgico una serie de signos y síntomas. De la presentación de estos signos y síntomas depende en buena medida el envío a un Servicio de Alergología, la realización de pruebas cutáneas y estudios de laboratorio, y el inicio de un tratamiento específico, que incluye la administración de esteroides, la evitación de alérgenos comunes, y en algunos casos inmunoterapia específica. Hacer un diagnóstico clínico rápido y confiable puede ser esencial para el manejo inicial del paciente, especialmente cuando se quiere distinguir entre infección y alergia, en escenarios de alta demanda de atención y escasez de recursos. **Objetivo:** Evaluar la utilidad de los signos y síntomas respiratorios comúnmente considerados sugestivos de alergia. Explorar la combinación de signos y síntomas que mejor nos predice un diagnóstico de atopia. **Método:** Una entrevista que incluyó interrogatorio y exploración física de signos y síntomas respiratorios comunes, se aplicó a 50 pacientes consecutivos de entre 2 y 55 años, que acudieron por primera vez al Servicio de Inmunología Clínica en un hospital de tercer nivel de atención. De manera independiente y ciega se realizó a todos los pacientes pruebas cutáneas de alergia. Usando tablas de dos por dos, se calculó el valor predictivo positivo y razón de verosimilitud de cada componente de la evaluación, con la prueba cutánea positiva como estándar de referencia. **Resultados:** Se obtuvieron resultados completos de 50 pacientes (26 masculinos). La media de edad fue 13.57 años  $\pm$  12.49 (rango 2 a 51 años). Sesenta y cuatro por ciento de las pruebas cutáneas fueron positivas para al menos un alérgeno. Los componentes de la historia clínica que tuvieron un LR+ más alto fueron: antecedente materno o paterno (LR+ 2.25), estornudos en salva (LR+ 1.77), desencadenante conocido (LR+ 2.33), pliegue de Dennie-Morgan (LR+ 3.00) y sibilancias a la auscultación (LR+ 4.03). La combinación de tres componentes: antecedente materno o paterno, estornudos en salva y doble pliegue palpebral inferior, dio un LR+ de 5.18. **Discusión:** La mayoría de los signos y síntomas clínicos evaluados no demostraron ser útiles en forma aislada y por sí mismos. Sin embargo, cuando se evalúan combinados o en secuencia, los signos y síntomas que clásicamente se consideran sugestivos de alergia, pueden de manera aceptable predecir el resultado de las pruebas cutáneas.

**Palabras clave:** Atopia, valor predictivo, síntomas respiratorios.

### ABSTRACT

**Background:** In the clinical assessment of the patient with respiratory complaints, a series of signs and symptoms are traditionally considered suggestive of allergic origin. The presence of those «classical»

\* Unidad de Investigación en Inmunodeficiencias del Instituto Nacional de Pediatría.

\*\* Departamento de Investigación en Inmunogenética y Alergia del Instituto Nacional de Enfermedades Respiratorias. Ciudad de México, México.

symptoms and signs can trigger patient referral, allergy skin test, laboratory workup and specific treatment. A quick and reliable clinical examination can thus be essential for the initial management, especially to distinguish between chronic infection and allergy, in the context of a high demand for health care services and a scarcity of resources. **Objective:** We aimed to assess the diagnostic usefulness of respiratory signs and symptoms typically considered suggestive of allergy, and to identify a combination of signs and symptoms that best predict a final diagnosis of atopy. **Methods:** We clinically assessed 50 consecutive patients who came for a first-time allergy consultation. Family history, patient symptoms and clinical examination were registered in a data sheet; each classical symptom, sign or background was considered a diagnostic test, and its usefulness to predict a final diagnosis of atopy was assessed using contingency tables, with the allergy skin test result as the standard of reference. Predictive values and likelihood ratios (LR) were calculated for each component of the clinical evaluation. **Results:** Mean age was 13.6 years (range 2 to 51), 52% male sex, and 64% positive allergy skin tests. The components from the clinical evaluation with the highest positive LR (LR+) were: parental antecedent (LR+ 2.25), sneezing spells (LR+ 1.77), known trigger (LR+ 2.33), and double inferior palpebral fold («Dennie Morgan signs», LR+ 3.00), as well as wheezing on auscultation (LR+ 4.03). The combination of three components (a construct): positive family history, sneezing spells, and double inferior palpebral fold, resulted in a LR+ of 5.18. **Discussion:** Most signs and symptoms did not prove to be, in and by themselves, useful to predict a final diagnosis of atopy. When considered sequentially or combined, however, the clinical components classically considered of allergic origin, can acceptably predict a positive skin allergy test.

**Key words:** Atopy, predictive value, respiratory symptoms.

## INTRODUCCIÓN

El diagnóstico clínico depende de nuestra habilidad para interpretar resultados de pruebas diagnósticas. En realidad, cualquier característica de un paciente que varía en una enfermedad dada podría calificar como una prueba diagnóstica, incluyendo prácticamente todos los elementos de la historia clínica y el examen físico. Intuitivamente, los clínicos aplicamos un razonamiento bayesiano al evaluar posibles diagnósticos diferenciales.<sup>1</sup>

### Probabilidad bayesiana

El teorema de Bayes implica que la probabilidad pre-prueba de que una hipótesis sea cierta, multiplicada por el peso de nueva evidencia (likelihood ratio), genera probabilidades post-prueba de que la hipótesis sea cierta.

La razón de verosimilitud o likelihood ratio (LR) resume las características operativas de una prueba diagnóstica, como la razón de pacientes con una enfermedad, a aquéllos sin dicha enfermedad, entre los que tienen un resultado positivo o negativo de la prueba y se deriva de la sensibilidad y especificidad de dicha prueba, de acuerdo a las siguientes fórmulas:

Para un resultado positivo:

$$LR+ = \text{sensibilidad} / (1 - \text{especificidad})$$

Para un resultado negativo

$$LR- = (1 - \text{sensibilidad}) / \text{especificidad}$$

Esto responde a la pregunta bayesiana, «¿Cuál es la probabilidad de que el paciente tenga la enfermedad dado que tiene un resultado anormal?». Un atributo im-

portante del teorema de Bayes es que los momios post-prueba para una enfermedad, luego de realizar una primera prueba, se convierten en los momios pre-prueba para la siguiente prueba, asumiendo que las pruebas son diferentes entre sí.

El único resultado verdaderamente inútil para una prueba diagnóstica, es el de un LR de 1.0 (sensibilidad y especificidad de 50%), ya que multiplicar los momios pre-prueba por 1.0 no cambia nada. Arbitrariamente, 2.0 y 0.5 se han propuesto como valores mínimamente útiles para LR's positivos y negativos, respectivamente.

Cada pregunta y examen físico durante una entrevista clínica constituye una prueba diagnóstica con su respectivo LR. Lo que distingue a estas pruebas de otras más formales, es que rara vez conocemos su sensibilidad, especificidad y LR's, por lo que la mayor parte del tiempo sólo trabajamos con impresiones generales de su utilidad.

Es sencillo medir las características operativas de una respuesta, un síntoma o un signo, como se hace con cualquier otra prueba diagnóstica. Cuando se evalúa cada una por separado, su sensibilidad y especificidad no suelen ser impresionantes, pero como sabemos, su resultado debe ser interpretado en el contexto de la historia de cada paciente. Intuitivamente, los clínicos no consideramos cada respuesta durante el interrogatorio en forma aislada, sino que sumamos y vamos actualizando nuestra estimación de la probabilidad de que el paciente padezca tal o cual enfermedad, conforme avanzamos en la entrevista.

A cada nuevo hallazgo, la probabilidad de un diagnóstico sospechoso incrementa, a pesar de que ninguna de las «pruebas» por sí solas (adenomegalia cervical

Cuadro I. Tabla de 2 x 2.

Prueba	Enfermedad	
	Presente	Ausente
Positiva	a	B
Negativa	c	D

anterior, moco hialino) es ni remotamente específica para infección o alergia, por ejemplo. Su pobre desempeño individual no disminuye su gran importancia cuando se combinan en una secuencia lógica. Así, la impresión cualitativa que teníamos al principio puede transitar de «posible» a «muy probable», lo que dicta las pruebas diagnósticas formales a las que el paciente ha de someterse, y el manejo inicial.

$$\begin{aligned} \text{Sensibilidad} &= a/(a + c) & \text{Especificidad} &= d/(b + d) \\ \text{Valor predictivo positivo} &= a/(a + b) & \text{Valor predictivo negativo} &= d/(c + d) \\ \text{Momios Pre-prueba} &= \text{prevalencia}/(1 - \text{prevalencia}) \\ \text{Likelihood ratio para una prueba positiva} &= \text{sensibilidad}/(1 - \text{especificidad}) \\ \text{Likelihood ratio para una prueba negativa} &= (1 - \text{sensibilidad})/\text{especificidad} \\ \text{Probabilidad Pre-prueba (prevalencia)} &= (a + c)/(a + b + c + d) \end{aligned}$$

#### Razones de verosimilitud (*likelihood ratios*)

Los LR son estadísticos inferenciales de probabilidad condicional, que resumen la exactitud o utilidad diagnóstica de una prueba, y tienen varias propiedades particulares poderosas que los hacen más útiles clínicamente.<sup>2</sup> Cada resultado de una prueba tiene su propio LR, que indica cuántas veces más (o menos) probable es que los pacientes enfermos tengan ese particular resultado, comparado con pacientes sin la enfermedad. Más formalmente, es la razón entre la probabilidad de obtener el resultado específico en gente con la enfermedad y la probabilidad de obtenerlo en gente sin la enfermedad.<sup>3</sup>

Un LR mayor de 1 indica que el resultado de la prueba se asocia con la presencia de enfermedad, mientras que un LR menor de 1 indica que el resultado de la prueba se asocia con la ausencia de enfermedad. Entre más se aleje el LR de 1 más fuerte es la evidencia para decretar presencia o ausencia de enfermedad. Se considera que LR por arriba de 10 y por debajo de 0.1 proveen evidencia fuerte para confirmar o descartar diagnósticos.

En la práctica clínica es importante conocer cómo una prueba en particular predice el riesgo de una anomalía. Esto no lo hacen la sensibilidad y especificidad: ellas describen cómo la anomalía predice un

resultado particular de una prueba.<sup>5</sup> Los valores predictivos positivo y negativo, por otro lado, sí predicen la probabilidad de anomalía dado que se obtuvo un resultado particular, pero dependen de la prevalencia de anomalía en la muestra estudiada, lo que hace muy difícil generalizar los valores predictivos arrojados por un estudio diagnóstico, a no ser que se trate de un estudio poblacional.<sup>6</sup> Los LR se pueden emplear para calcular la probabilidad de anomalía, al tiempo que se adaptan a probabilidades pre-prueba variables y en diferentes contextos.<sup>3</sup>

Nos propusimos evaluar la utilidad diagnóstica de signos y síntomas respiratorios comúnmente considerados sugestivos de alergia, y explorar una combinación de signos y síntomas que mejor prediga un diagnóstico final de atopia, utilizando como estándar de referencia las pruebas cutáneas.

#### MATERIAL Y MÉTODOS

En pacientes de entre 2 y 55 años que acudieron a consulta por primera vez en el Servicio de Inmunología Clínica del INER, de enero a julio de 2008, se realizó una historia clínica completa, que incluyó interrogatorio de antecedentes familiares y personales, molestia principal y semiología de los síntomas respiratorios, así como exploración física dirigida a piel, cabeza y aparato respiratorio. Las respuestas y observaciones del entrevistador se registraron en una hoja de recolección. En otra visita, sin conocer el resultado del examen clínico se realizaron pruebas cutáneas por punción (*prick test*) con lanceta de plástico bifurcada, para 23 alérgenos comunes, con histamina como control positivo y solución salina normal como control negativo. La prueba se consideró positiva si se registraba una reacción de roncha y eritema mayor al control negativo y mayor o igual de 3 mm en niños menores de 8 años, o  $\geq 5$  mm en adultos y niños de 8 años en adelante. Además se solicitó a cada paciente determinación de eosinófilos en moco nasal y en sangre periférica y niveles séricos de IgE total.

Se consideró a cada componente de la historia clínica una prueba diagnóstica, positiva si la semiología era sugestiva de un diagnóstico de atopia y negativa si por el contrario sugería otra etiología (ej. estornudos en salva vs aislados, moco hialino vs purulento). Y se comparó el resultado de cada prueba por separado, con el resultado de las pruebas cutáneas, que sirvió como estándar de referencia (*gold standard*). Uno de los autores realizó la mayoría de las evaluaciones clínicas y un técnico realizó casi la totalidad de las pruebas cutáneas, cada uno de manera independiente y a ciegas. Cuando había duda, se consultaba el resultado o la impresión clínica entre los investigadores. No se realizó una prueba formal de concordancia o reproducibilidad.

## ANÁLISIS ESTADÍSTICO

Con ayuda del paquete estadístico SPSS versión 16.0 para Windows, se construyeron tablas de 2 x 2 para cada elemento de la historia clínica y se calculó sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y likelihood ratios (LR) positivos y negativos para cada uno, comparado con el resultado de las pruebas cutáneas como referencia. Se calcularon intervalos de 95% de confianza (95%CI) para cada LR. Después se eligieron los componentes de la historia clínica con un LR+ mayor de 1.5, para elegir una combinación de tres a cinco elementos que pudiera predecir el resultado positivo en las pruebas cutáneas. Se exploró mediante análisis multivariado la secuencia de tres elementos que mejor pudiera predecir el resultado, positivo o negativo, de las pruebas cutáneas. Se consideró significativa una *p* menor de 0.001.

## RESULTADOS

Se obtuvieron resultados completos de 50 pacientes (26 masculinos). La media de edad fue 13.57 años  $\pm$  12.49 (rango 2 a 51 años). Sesenta y cuatro por ciento de las pruebas cutáneas fueron positivas para al menos un alérgeno, lo que representa una prevalencia o probabilidad pre-prueba de 0.64. Los componentes de la historia clínica que tuvieron un LR+ más alto fueron: antecedente materno o paterno (LR+ 2.25), estornudos en salva (LR+ 1.77), pliegue de Dennie-Morgan (LR+ 3.00), y sibilancias a la auscultación (LR+ 4.03).

La combinación de tres componentes: antecedente materno o paterno, estornudos en salva y doble plie-

gue palpebral inferior, dio un LR+ de 5.18, con un valor predictivo positivo (VPP) de 90% y una especificidad de 97.4%, con respecto a las pruebas cutáneas como estándar de referencia. En el *cuadro II* se presentan los LR+ e intervalos de confianza al 95% para cada componente de la historia clínica evaluado.

## DISCUSIÓN

La mayoría de los signos y síntomas clínicos evaluados no demostraron ser útiles en forma aislada y por sí mismos. Sin embargo, cuando se evalúan combinados o en secuencia, los signos y síntomas que clásicamente se consideran sugestivos de alergia, pueden de manera aceptable predecir el resultado de las pruebas cutáneas. Como se puede observar, los intervalos de confianza al 95% son demasiado amplios, lo que probablemente refleja un tamaño de muestra insuficiente.<sup>7</sup> Ésta es la principal limitación del presente trabajo en su versión actual. Se tiene contemplado continuar el estudio hasta tener al menos el doble de sujetos para tener márgenes más estrechos y mayor confianza en las estimaciones.

El que un síntoma aislado se considere poco útil como prueba diagnóstica no debe sorprender y coincide con lo reportado por otros grupos, que han evaluado cada componente de la evaluación clínica inicial para predecir un diagnóstico final. La explicación para este fenómeno es que cuando evaluamos un paciente en nuestro consultorio no consideramos cada signo o síntoma aisladamente, sino varios datos (demográficos, antecedentes, síntomas, signos) en conjunto y cada respuesta positiva aumenta o ajusta nuestra estimación de la probabilidad de que ese paciente dado padezca determinada enfermedad. De tal modo que, en una secuencia de preguntas durante un interrogatorio, cada respuesta conlleva una razón de verosimilitud (LR), lo que nos da una probabilidad post-prueba, que se convierte en la probabilidad pre-prueba actualizada de la siguiente pregunta.

Así, una combinación o secuencia de respuestas y signos positivos, se puede sumar para darnos una razón de verosimilitud (LR) tan importante que nos permita prácticamente confirmar o descartar un diagnóstico. La utilidad de estudios como el presente, además de poner a prueba el conocimiento tradicional de una manera objetiva, es ubicar aquellos componentes de la historia clínica que pueden ser más útiles para decidir la probabilidad aproximada de un diagnóstico y seleccionar a los pacientes que sean mejores candidatos para una prueba diagnóstica formal.

## BIBLIOGRAFÍA

- Gill JG, Sabin L, Schmid CH. Why clinicians are natural bayesians. *BMJ* 2005; 330: 1080-3.

**Cuadro II. Razones de verosimilitud (LRs) e intervalos de confianza (IC) al 95%.**

Signos y síntomas	LR-	LR+	IC95%
Pliegue Dennie-Morgan		3.00	(1.01-8.9)
Antecedente materno/paterno		2.25	(0.53-9.48)
Estornudos en salva		1.77	(0.95-3.3)
IgE elevada	0.11	5.43	(0.90-32.7)
Sibilancias a la auscultación		4.03	(0.22-73.9)
Antecedente + pliegue + estornudos en salva		5.182	(0.3-91.1)
Prurito ocular o nasal		1.12	(0.85-1.49)
Desencadenante conocido		2.33	(0.92-5.93)
Moco hialino	0.42	1.46	(0.94-2.28)
Predominio matutino	0.9	1.12	(0.6-2.1)
Alergia a alimentos/meds.	0.9	1.31	(0.39-4.45)
Cornetes hipertróficos	0.45	1.21	(0.88-1.66)
Moco hialino	0.82	1.29	(0.66-2.53)
Mucosa nasal pálida	0.79	1.27	(0.7-2.3)
Rudeza/inspiración prolong.	0.95	1.41	(0.3-6.5)

2. Sackett DL, Straus S, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach EBM. 2nd ed. Edinburgh: Churchill Livingstone, 2000: p. 67-93.
3. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. *BMJ* 2004; 329: 168-9.
4. Fagan TJ. Letter: nomogram for Bayes' theorem. *N Engl J Med* 1975; 23: 257.
5. Altman DG, Bland JM. Diagnostic tests 1: sensitivity and specificity. *BMJ* 1994; 308: 1552.
6. Altman DG, Bland JM. Diagnostic tests 2: predictive values. *BMJ* 1994; 309: 102.
7. Faul F, Erdfelder E, Lang AG, Buchner A. G\*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 2007; 39: 175-191.
8. Delamothe T. Diagnosis-the next frontier. *BMJ* 2006; 333: 0-f.
9. Sackett DL, Haynes RB. Evidence base of clinical diagnosis: The architecture of diagnostic research. *BMJ* 2002; 324: 539-41.
10. Call SA, Vollenweider MA, Hornung CA, Simel DL, McKinney WP. The rational clinical examination. Does this patient have influenza? *JAMA* 2005; s.l. 293, 8: 987-997.

Dirección para correspondencia:  
Saúl Oswaldo Lugo Reyes  
Instituto Nacional de Pediatría  
Torre de Investigación  
E-mail: dr.logo.reyes@gmail.com