

Evaluación de Test Diagnóstico. Utilidad de adenosin deaminasa (ADA) en el diagnóstico de peritonitis tuberculosa

Felipe Salech M.⁽¹⁾, Sebastián Valderrama Ch.⁽¹⁾,
Manuel Espinoza S.⁽²⁾, Ignacio Salas S.⁽²⁾, Enrique Veas R.⁽²⁾,
Francisco Barrera M.⁽²⁾, Juan Ignacio Quinteros B.⁽²⁾, Alejandro Soza R.^(2,5),
Luz María Letelier S.^(3,4) y Arnoldo Riquelme P.^(2,3,4).

EVIDENCE BASED MEDICINE IN HEPATOLOGY III-A. CRITICAL ANALYSIS OF DIAGNOSTIC TEST: EFFECTIVENESS OF ADA FOR DIAGNOSING TUBERCULOUS PERITONITIS

The present article continues the series of publications on clinical decision making about Hepatology, using Evidence-Based Medicine (EBM) criteria. In a clinical scenario of a patient with ascites, the effectiveness of adenosin deaminase (ADA) in the diagnosis peritoneal tuberculosis (PTB) is appraised. Using the clinical question based search model, seventeen studies are identified. One of these articles is critically appraised, including validity criteria (comparison with an accepted gold standard, spectrum of patients, reproducibility of the assessments), analysis of the results and a discussion of its applicability. The selected article was a prospective study that included consecutive patients. The diagnosis of peritoneal tuberculosis was confirmed by laparoscopy on all patients, and ADA was determined by Giusti method. The results showed that ADA is a discriminating test for diagnosing PTB with an optimal cut-off value of 32 IU/L. We conclude that measurement of ADA level in ascitic fluid is an accurate test for diagnosing PTB.

Key words: Evidence Based Medicine, Diagnostic test, Adenosine Deaminase, Tuberculous peritonitis, ROC curve.

Introducción

Los tests diagnósticos son utilizados de manera rutinaria en medicina para realizar *screening*, diagnosticar, clasificar y monitorizar la progresión de una enfermedad. Esta información proviene de distintas fuentes, que incluyen imágenes, exámenes bioquímicos, patológicos y psi-

cológicos, así como de síntomas y signos obtenidos de la historia y el examen físico¹. Cada uno de estas fuentes constituye un test diagnóstico, cuyos resultados pueden ser comparados entre sí y contra el Gold Standard definido para cada situación clínica.

Los clínicos deben decidir qué test diagnóstico utilizar y cómo interpretar sus resultados.

⁽¹⁾ Internos, Escuela de Medicina, Pontificia Universidad Católica de Chile.

⁽²⁾ Unidad Docente Asociada de Medicina Hospital de Urgencia Asistencia Pública.

⁽³⁾ Unidad de Medicina Basada en Evidencia, Escuela de Medicina, Pontificia Universidad Católica de Chile.

⁽⁴⁾ Departamento de Medicina Interna, Pontificia Universidad Católica de Chile.

⁽⁵⁾ Departamento de Gastroenterología, Pontificia Universidad Católica de Chile.

Recibida y aceptada: 21 julio 2005

Para esto es necesario conocer elementos estadísticos y metodológicos que nos permitan validar los resultados de la prueba.

Además, se requiere conocer el contexto clínico en que fue validado el test diagnóstico y ver si corresponde a nuestro escenario en particular.

Para validar la adenosin deaminasa (ADA) como test diagnóstico para peritonitis tuberculosa (PTB), necesitaremos responder las siguientes preguntas:

- ¿Cómo analizar un artículo referente a test diagnóstico?
- ¿Son válidos los resultados de este estudio?
- ¿Me ayudarán los resultados a hacer el diagnóstico y tomar conductas en mi paciente?

Además, discutiremos los siguientes conceptos:

- Gold Standard
- Sensibilidad (S) y especificidad (E).
- Valor predictivo positivo (VPP) y valor predictivo negativo (VPN)
- Certeza y precisión
- Razón de probabilidad (likelihood ratio)
- Prevalencia
- Pre test odds y post test odds
- Probabilidad pre y post-test

Escenario clínico

Usted es residente medicina interna y se encuentra de turno. Ingresa un paciente de 65 años con historia de 3 meses de fiebre nocturna, asociado a un aumento difuso del volumen abdominal. Al examen físico destaca la presencia de líquido ascítico (LA) de moderada cuantía, en ausencia de estigmas de daño hepático crónico. El paciente niega el consumo de alcohol y refiere como antecedente una tuberculosis (TB) pulmonar tratada hace 15 años.

Dentro de los exámenes de laboratorio iniciales destaca una albúmina de 3,9 g/dL, bilirrubina total de 1,0 mg/dL, tiempo de protrombina de 80%, hematocrito de 35%, recuento de plaquetas de 300.000 por mm³, leucocitos de 12.000 por mm³ y una radiografía de tórax con imágenes apicales compatibles con secuelas de TB pulmonar.

La pregunta

Con estos antecedentes, más los hallazgos del examen físico y de laboratorio usted se plantea la posibilidad que la ascitis sea de origen tuberculoso. Usted sabe que la baciloscopia en LA tiene un bajo rendimiento, que el cultivo de Koch demora tiempo, y quiere evitarle al paciente los riesgos inherentes a una laparotomía exploradora, y ha escuchado que la determinación de ADA en otros fluidos corporales permite hacer el diagnóstico de tuberculosis de forma rápida y poco invasiva. De esta manera surge la pregunta:

¿Cuál es el rendimiento de ADA en líquido ascítico para el diagnóstico de peritonitis tuberculosa (PTB)?

Búsqueda de evidencia

Para responder a esta pregunta, se realiza una búsqueda en MEDLINE combinando los términos "adenosine deaminase" AND "peritoneal tuberculosis" (ambos términos MeSH); revisando artículos desde 1996 hasta junio 2005. Esta búsqueda inicial permitió encontrar 17 publicaciones de distinto tipo relacionadas con el tema, que incluían cartas, revisiones, estudios clínicos prospectivos. De ellos se eligió, para ser analizado en detalle, el trabajo de *Bhargava D.K. y cols. Adenosine deaminase (ADA) in peritoneal tuberculosis: diagnostic value in ascitic fluid and serum*².

Evaluación crítica de un trabajo que reporta un test diagnóstico

I. Validez interna

Para la lectura crítica de un artículo se utilizan las preguntas propuestas en el artículo de Greenhalgh publicado en el *British Medical Journal*³, que buscan determinar la *validez interna* del trabajo, o sea, que el diseño y la metodología del estudio disminuyan la probabilidad de sesgo de sus resultados, es decir se aproximen a la verdad y por lo tanto nos permitan estimar la real eficacia de la intervención.

¿Ha sido el test comparado con el Gold Standard adecuado?

Para validar un nuevo test diagnóstico, se debe aplicar en cada paciente, además del test en evaluación, otra técnica diagnóstica que nos permita determinar si efectivamente el paciente padece o no la condición. Esta segunda técnica diagnóstica debe ser el Gold Standard, es decir, la mejor técnica disponible para certificar el diagnóstico. Es posible que para una determinada enfermedad exista más de un Gold Standard, situación que se da en el caso de la TB, donde se considera al caso como enfermo, si tiene una biopsia con granulomas caseosos, un cultivo o una baciloscopia positiva. No se debe considerar como Gold Standard la respuesta clínica al tratamiento, que otras patologías en su curso natural pueden resolverse espontáneamente mientras eran tratados, sin haber correspondido a una TB, situación evidentemente válida para cualquier patología.

El artículo escogido, en su metodología incluye una comparación entre el examen a validar ADA y un buen Gold Standard, en este caso la biopsia laparoscópica, sin asumir la evolución clínica como criterio diagnóstico.

¿La muestra de pacientes incluidos incorpora un espectro apropiado de casos en los cuales el test diagnóstico será aplicado en la práctica clínica?

Es importante que el espectro de enfermos en que se realizó el estudio, no incluya a pacientes seleccionados, dado que esto constituirá un sesgo que altera los resultados obtenidos en el estudio, inflando falsamente la sensibilidad del método. Debemos asegurarnos además, de que el espectro de pacientes sea adecuado a nuestra realidad, de tal forma que los resultados sean finalmente aplicables a nuestro caso.

Bhargava y cols midieron los niveles de ADA en un grupo de pacientes que entraron al estudio de forma prospectiva por ascitis, que finalmente resultó ser secundario a múltiples causas, frecuentes en nuestra práctica clínica. No especifica sin embargo, las características demográficas de la población estudiada, información que evidentemente resultaría de gran utilidad.

¿Los resultados de los estudios que están siendo evaluados influyen en la decisión de realizar el Gold Standard?

Es importante que el Gold Standard se practique a todos los pacientes y no sólo a aquellos que tienen el test diagnóstico positivo.

En el estudio, a todos los pacientes se les realizaron ambas técnicas diagnósticas de forma prospectiva. En el caso de exámenes que requieran una interpretación subjetiva es fundamental que, además de realizar ambas técnicas diagnósticas a todos los pacientes, éstas se realicen de forma ciega, puesto que el conocer previamente el resultado de un examen puede alterar la interpretación del segundo, situación conocida como sesgo de expectación.

¿Son los métodos descritos con suficiente detalle para permitir su reproducción?

Es importante que se describa con suficiente detalle el test diagnóstico de tal forma de hacerlo reproducible en nuestro paciente.

Bhargava y cols utilizan el método de Giusti para realizar la medición de ADA. Técnica ampliamente descrita en la literatura y disponible en muchos laboratorios de mediana complejidad⁴.

II. Resultados

Evaluación crítica de los resultados de un trabajo que reporta un test diagnóstico

Es posible que un estudio determinado cumpla con todas las características recién señaladas, sin embargo, si los resultados que entrega no son discriminadores, el test no tendrá aplicabilidad alguna. A continuación analizaremos cómo valorar cuantitativamente la capacidad de discriminación de un test diagnóstico.

Propiedades de un Test Diagnóstico

Teorema de Bayes

Permite estimar el valor predictivo de un test, es decir estimar las probabilidades de predecir correctamente (valor predictivo) una patología basado en la validez del test efectuado (sensibilidad, especificidad, probabilidad de falsos positivos y falsos negativos).

En otras palabras, nos permite conociendo la prevalencia de una enfermedad en la población a la que pertenece un individuo y los valores de sensibilidad y especificidad de la prueba, calcular la probabilidad de que un sujeto que ha dado positivo en el test, verdaderamente tenga esa enfermedad.

En la práctica, en el teorema de Bayes se tabulan en forma cruzada las 4 combinaciones posibles:

- a) Gold Standard (+) y test (+)
- b) Gold Standard (-) y test (+)
- c) Gold Standard (+) y test (-)
- d) Gold Standard (-) y test (-).

La información así ordenada nos permite el cálculo de las propiedades del test, que nos posibilita objetivar el poder discriminador de un test diagnóstico.

Estas propiedades, que a continuación se detallan, se encuentran resumidos en la Figura 1, y se presenta el cálculo de cada uno de ellos para el artículo de Bhargava y cols.

1. Sensibilidad

Se define como la razón de los verdaderos positivos (test positivos en enfermedad) o proporción de individuos con la enfermedad que presentan un resultado positivo. Indica qué tan

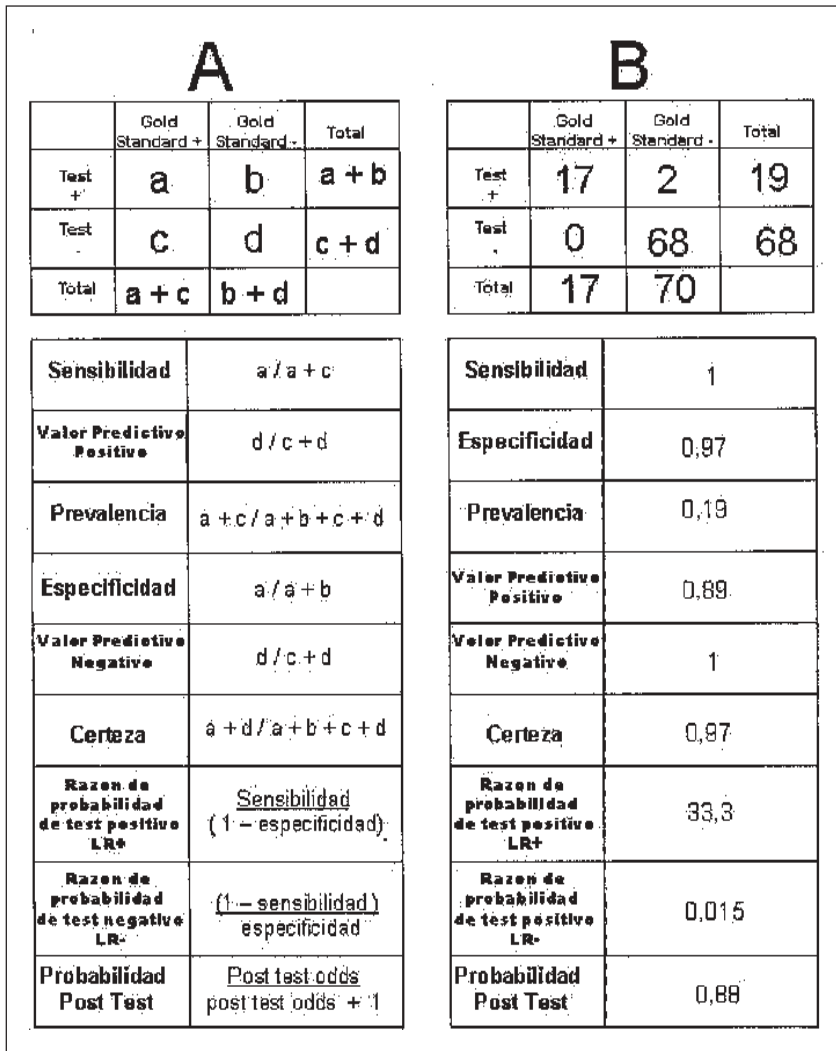


Figura 1. En la Tabla A, se resume el teorema de Bayes, y se muestran las fórmulas para calcular las propiedades de un test diagnóstico. En la Tabla B, se encuentran aplicados estos conceptos al artículo revisado.

buena es una prueba para identificar a las personas enfermas.

2. Especificidad

Definida como la razón de verdaderos negativos (test negativos en los sin enfermedad) es decir, la proporción de individuos sin la enfermedad que presentan un resultado negativo. Indica cuán buena es la prueba para identificar a los individuos que no tienen la enfermedad.

3. Prevalencia

Frecuencia con la que se presenta una determinada enfermedad en una población, puede ser estimada por los clínicos de diferentes maneras: por experiencia personal, por estadísticas de salud.

En otras palabras, es la frecuencia (expresada en porcentaje) en la cual se encuentra una condición determinada en el total de la población y pretende responder a la pregunta:

¿Cuán frecuente es esta condición en mi población?

Corresponde a la Probabilidad pre Test.

La sensibilidad y la especificidad son dos parámetros independientes. De ellos y de la prevalencia de la enfermedad, dependen los valores predictivos positivo (VPP) y negativo (VPN).

4. Valor predictivo positivo

Si existe una persona con un test positivo, cuál es la probabilidad de que él o ella tengan la condición en estudio.

5. Valor predictivo negativo

Indica cuál es probabilidad de que una persona con un examen negativo no tenga la condición en estudio.

6. Certeza

Corresponde a los positivos verdaderos y negativos verdaderos como una proporción de todos los resultados. Responde a la pregunta:

¿Qué proporción de todos los test han dado el resultado correcto?

7. Razón de probabilidad

Razón de probabilidad o LR: es la propie-

dad más importante de un test, pues nos permite movernos de una probabilidad pre-test a una probabilidad post-test. Se define como la razón entre la probabilidad de tener un resultado del test en la población con la condición versus tener el mismo resultado en la población sin la condición.

8. Probabilidad post test

Éste es el marcador más importante para nuestro paciente ya que si aplicamos una combinación entre la prevalencia de la enfermedad, la probabilidad pre-test del paciente y el LR + del test obtendremos la probabilidad de que tenga la enfermedad o condición luego de aplicado el test.

Resultados del Trabajo

El trabajo escogido entrega los valores de ADA para cada paciente, e informa sus resultados como sensibilidad y especificidad para distintos puntos de corte, arbitrariamente escogidos por el autor, que van desde 29 hasta 36 UI/L. El autor selecciona este último como punto de corte, por mostrar la mejor relación de sensibilidad y especificidad (100% y de 97,1% respectivamente).

En la Figura 2 se muestra de forma gráfica como se distribuyen los valores de ADA en pacientes con ascitis de origen tuberculoso y no tuberculoso. Esta gráfica también nos da una idea del punto de corte óptimo para el test.

Sin embargo, la mejor forma de estimar el punto de corte óptimo, es a través de la utilización de una curva ROC. Se construye graficando la sensibilidad (eje Y = ordenada) v/s 1-especificidad (eje X = abscisa). Su utilidad radica en que permite definir el mejor punto de corte para un determinado test diagnóstico, es decir, el que tiene la mejor sensibilidad y especificidad simultáneamente⁵. La curva ROC no fue construida por los autores, sin embargo, dado que entregan los valores de ADA para cada paciente, y su diagnóstico final por el Gold Standard, es posible construirla. Para ello se utilizó el software Med-Calc versión 8.0.0.0.

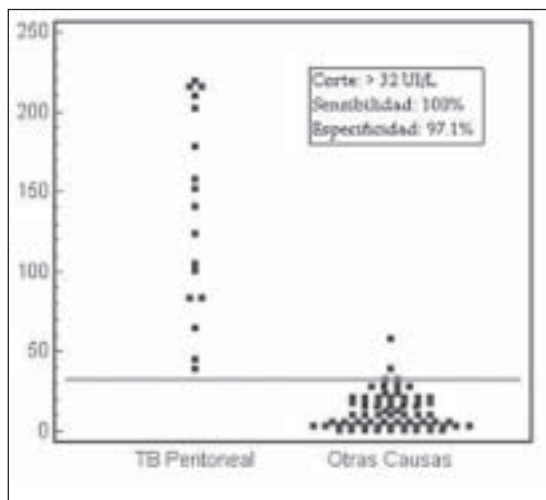


Figura 2. En esta figura se muestran los valores de ADA de cada paciente, ordenados según la causa de ascitis diagnosticada por el Gold Standard. Este tipo de gráficas nos entregan una aproximación visual del poder discriminador de un test diagnóstico, y del mejor punto de corte.

Con la curva ROC se determina que la mejor relación de sensibilidad y especificidad se da estableciendo como punto de corte valores de ADA mayores a 32 UI/L (Figura 3).

La discordancia entre puntos de corte señalado por el autor (> 36 UI/L), y la mostrada por la curva ROC (> 32 UI/L), en este caso no se tradujo en un cambio en la sensibilidad ni en la especificidad del estudio, dado que no habían pacientes con ADA entre estos valores. Sin embargo, esta diferencia puede ser de gran importancia en estudios con más pacientes, o en aquellos donde los valores sean menos dispersos.

III. Aplicabilidad

¿Los resultados me ayudarán en el cuidado de mi paciente?

Finalmente, es fundamental aclarar si el test analizado tiene aplicabilidad real en el caso.

Las técnicas actualmente disponibles para el diagnóstico de tuberculosis peritoneal son muy invasivas, y requieren de largo tiempo de espera. Por otro lado, el tratamiento de la tuberculosis es por período prolongado, no exento de complicaciones. En este contexto, contar con un test poco invasivo y altamente discrimi-

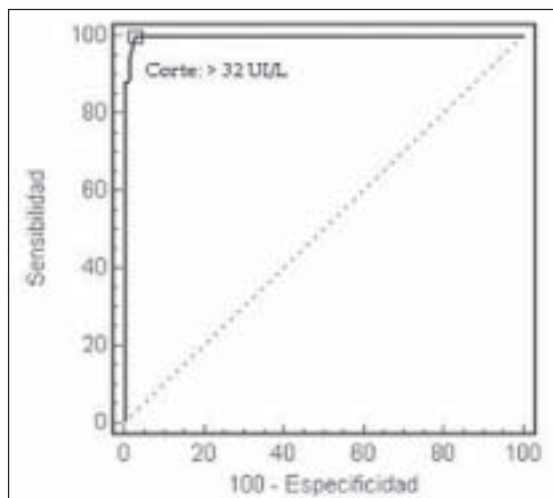


Figura 3. Curva ROC construida con los datos del trabajo. En ella se observa que el punto donde se encuentra la mejor relación entre sensibilidad y especificidad corresponde al punto de corte 32 UI/L.

minador para diagnosticar TBP resulta de gran utilidad, por lo que ADA se transforma una herramienta de gran utilidad en el estudio del paciente con ascitis.

Resolución del escenario clínico

Volviendo al caso que motivó nuestra discusión, al paciente se le realizó una paracentesis y se envió la muestra de líquido ascítico a análisis de ADA, el que mostró un valor de 58 UI/L.

Considerando que el paciente tiene una probabilidad pre-test del 23% de presentar una TB peritoneal; y el LR al que se asocian los valores de ADA > a 32, ya calculado de 33,3, la probabilidad pos test de que el paciente tenga una TB peritoneal es de 88%.

El análisis recién realizado se visualiza de forma más simple al revisar el nomograma creado por Fagan⁶, donde se grafican los cambios de la probabilidad pre-test al aplicar un LR positivo o negativo sobre él.

Para el caso recién presentado se muestra el nomograma de Fagan, con los trazados correspondientes a valores de ADA por sobre y bajo el punto de corte analizado (Figura 4).

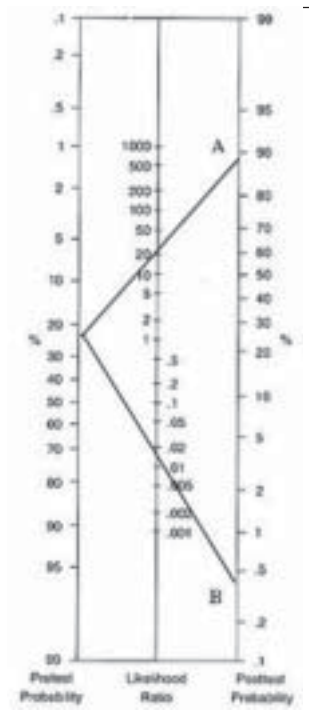


Figura 4. Nomograma para interpretar resultados de test diagnóstico, adaptado por Fagan⁶. Los trazados A y B muestran la Probabilidad Post-test para pacientes con valores de ADA mayores y menores de 32 UI/L respectivamente.

usando los pasos de la Medicina basada en Evidencia. Se analizó la utilidad de la adenosin deaminasa (ADA) en el diagnóstico de tuberculosis peritoneal (TBP) en un paciente con ascitis. Usando un modelo de búsqueda basado en una pregunta clínica, se identificaron 17 artículos relacionados. Uno de ellos fue extensa y críticamente analizado, incluyendo una revisión de su validez (comparación con gold standard, espectro de pacientes analizados, y reproducibilidad del trabajo), análisis de sus resultados, y discusión de su aplicabilidad. El artículo seleccionado fue prospectivo, incluía a pacientes consecutivos, confirmó el diagnóstico de tuberculosis con una biopsia laparoscópica en todos los pacientes, y realizó la determinación de ADA utilizando el método de Giusti. Los Resultados mostraron que ADA es un test discriminativo para el diagnóstico de TBP, con un punto de corte óptimo de 32 UI/L. Concluimos que la medición de niveles de ADA en líquido ascítico, es una forma efectiva de diagnosticar TBP.

Palabras clave: Medicina Basada en Evidencia, Test diagnósticos, Adenosin Deaminasa, Peritonitis tuberculosa, Curva ROC.

Conclusión

El trabajo de Bhargava y cols, cumple con los criterios de validación para un estudio de test diagnóstico, y muestra para ADA resultados ampliamente discriminadores en el diagnóstico de tuberculosis peritoneal.

Según el análisis desarrollado en esta revisión, es posible afirmar que medir el valor de ADA en pacientes con ascitis es una forma rápida, simple, y de alta sensibilidad y especificidad para el diagnóstico de TBP, que no requiere de técnicas sofisticadas ni gran implementación de laboratorio, siendo de gran aplicabilidad en la práctica clínica cotidiana.

Resumen

El siguiente estudio continúa la serie de publicaciones sobre decisiones clínicas en Hepatología,

Bibliografía

- 1.- Sackett D, Haynes R, Guayatt G, Tugwell P. Clinical epidemiology: a basic science for clinical medicine, 2° ed. Boston: Little, Brown, 1991.
- 2.- Bhargava D, Gupta M, Nijhawan S, Dasarathy S, Kushwaha A. Adenosine deaminase (ADA) in peritoneal tuberculosis: diagnostic value in ascitic fluid and serum. Tubercle 1990; 71: 121-6
- 3.- Greenhalgh T. How to read a paper: Papers that report diagnostic or screening test. BMJ 1997; 315: 540-3.
- 4.- Giusti G. Adenosine deaminase. In: Bergmeyer HV, ed. Methods of enzymatic analysis. New York: Academic Press 1974: 1092-9.
- 5.- Hanley J, McNeil B. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982; 143: 29-36.
- 6.- Fagan T. Nomogram for Bayes theorem. N Engl J Med 1975; 293: 257.