

ARTICULO ORIGINAL

Nuevo índice de desempeño global de una prueba diagnóstica: el índice T

New global performance index of a diagnostic test: the T index

Eduardo de Jesús Torregroza-Diazgranados 

Médico especialista en Cirugía general y Cirugía de mama y tejidos blandos, Clínica de Cirugía de Mama, Bogotá, D.C., Colombia.

Resumen

Introducción. En el presente artículo se propone un nuevo índice de desempeño global de una prueba diagnóstica: el índice T, que corresponde a la proporción de aciertos (Pa) menos la proporción de error (Pe) dividido entre la proporción de aciertos (Pa).

$$\text{Índice T} = \frac{[(a+d) / N] - [(c+b) / N]}{[(a+d) / N]} \times 100$$

Métodos. Describir los fundamentos teóricos y prácticos del índice T. Comparar el índice T frente al índice de Youden para poder establecer si dos o más modalidades diagnósticas difieren en su capacidad de discriminar, correctamente, entre los individuos verdaderamente enfermos de los sanos. Comparar el índice T frente al índice de Youden para poder establecer si el índice T puede ser utilizado para escoger “el mejor punto de corte” cuando la prueba diagnóstica de interés expresa sus resultados en forma cuantitativa continua.

Resultados. Comparado frente al índice de Youden, el índice T permite distinguir entre las pruebas diagnósticas con alto desempeño diagnóstico de aquellas con pobre desempeño diagnóstico. Además, el índice T, de manera confiable, permite escoger “el mejor punto de corte” cuando una prueba diagnóstica utiliza una variable cuantitativa continua para expresar sus resultados.

Conclusiones. Al contrastar los resultados del índice T versus el índice de Youden y los resultados del índice T versus la curva COR, podemos afirmar que el índice T es un índice exacto y eficiente.

Palabras claves: índice de Torregroza; proporción de acierto; proporción de error; índice de Youden; curva COR; diagnóstico.

Fecha de recibido: 1/06/2021 - Fecha de aceptación: 30/09/2021 - Publicación en línea: 29/10/2021

Correspondencia: Eduardo de Jesús Torregroza-Diazgranados, Calle 111 No. 13-03, Bogotá, D.C., Colombia. Celular: 3175730414.

Correo electrónico: torregrozad@gmail.com

Citar como: Torregroza-Diazgranados EJ. Nuevo índice de desempeño global de una prueba diagnóstica: el índice T

Rev Colomb Cir. 2022;37:33-42. <https://doi.org/10.30944/20117582.944>

Este es un artículo de acceso abierto bajo una Licencia Creative Commons - BY-NC-ND <https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>

Abstract

Introduction. In this article, a new global performance index of a diagnostic test is proposed: the T Index, which corresponds to the proportion of the hit ratio (Hr) minus the error ratio (Er) divided by the hit ratio (Hr).

$$T \text{ Index} = \frac{[(a+d) / N] - [(c+b) / N]}{[(a+d) / N]} \times 100$$

Methods. To describe the theoretical and practical foundations of the T index; to compare the T index against the Youden index in order to establish whether two or more diagnostic modalities differ in their ability to correctly discriminate between truly ill and healthy individuals; and to compare the T index with the Youden index in order to establish whether the T index can be used to choose the “best cut-off point” when the diagnostic test of interest expresses its results in a continuous quantitative way.

Results. Compared with the Youden index, the T index allows us to differentiate between diagnostic tests with high diagnostic performance and those with poor diagnostic performance. In addition, the T index, reliably, allows choosing “the best cut-off point” when a diagnostic test uses a continuous quantitative variable to express its results.

Conclusions. By contrasting the results of the T index versus the Youden index and the results of the T index versus the COR curve, we can affirm that the T index is an accurate and efficient index.

Keywords: Torregroza index; hit ratio; error ratio; Youden index; ROC curve; diagnosis.

Introducción

Los resultados de los estudios de una prueba diagnóstica se analizan en una tabla de contingencia. La tabla de contingencia contiene dos columnas, que corresponden al resultado dicotómico de presencia o ausencia de enfermedad, determinada por la prueba diagnóstica estándar. Las filas, por su parte, corresponden a los resultados de la prueba de interés a evaluar. Las dos columnas y las dos filas forman cuatro celdas (tabla 1) que se designan, por convención, de izquierda a derecha y de arriba abajo con una letra: a, b, c, d.

Tabla 1. Estructura básica de una tabla de contingencia.

Resultado según patrón de referencia			
Resultado prueba diagnóstica	Enfermos	No enfermos	Total
Positiva	a	b	a + b
Negativa	c	d	c + d
Total	a + c	b + d	N

La celda a incluye a los pacientes enfermos en quienes la prueba diagnóstica, correctamente, resultó positiva; denominados verdaderos positivos. La celda b incluye a los pacientes sanos en quienes la prueba, equivocadamente, catalogó como positivos; denominados falsos positivos. La celda c incluye a los pacientes con la enfermedad en que la prueba diagnóstica, equivocadamente, catalogó como negativos; denominados falsos negativos. La celda d incluye a los pacientes sanos en quienes la prueba diagnóstica, correctamente, catalogó como negativos; denominados verdaderos negativos. N es la suma de los individuos enfermos (a + c) más los individuos sanos (b + d). La tabla 2, muestra el significado de cada celda en la tabla de contingencia.

Con todos los valores de la tabla de contingencia llenos, se calculan las medidas de utilidad o rendimiento de la prueba diagnóstica:

Tabla 2. Tabla de contingencia donde se aprecia el significado de cada celda.

Prueba o test diagnóstico	Positiva	Enfermedad		Total
		Presente	Ausente	
	Negativa	Verdaderos positivos (VP)	Falsos positivos (FP)	VP + FP
Total	Falsos negativos (FN)	Verdaderos negativos (VN)	FN + VN	
		VP + FN	FP + VN	N

1. Medidas de probabilidad ¹: Sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo.
2. Medidas de razones de probabilidad ²: Razón de probabilidad positiva, razón de probabilidad negativa

De forma alterna, es posible utilizar las medidas de resumen global de una prueba diagnóstica para establecer la utilidad de la prueba utilizando:

1. El índice de exactitud ³.
2. El índice de Youden ⁴.

El Índice de Exactitud, se calcula mediante la siguiente fórmula:

$$\frac{\text{Verdaderos positivos} + \text{Verdaderos negativos}}{\text{Verdaderos positivos} + \text{Verdaderos negativos} + \text{Falsos positivos} + \text{Falsos negativos}}$$

El Índice de Youden, se calcula mediante la siguiente fórmula:

$$\text{Índice de Youden} = (\text{Sensibilidad} + \text{Especificidad}) - 100$$

El propósito de una medida de resumen global de una prueba diagnóstica o índice de desempeño es resumir en una única medida o dato que exprese el poder de la prueba diagnóstica a discriminar, de manera correcta, entre individuos enfermos e individuos sanos ⁵. Los índices de desempeño global de una prueba diagnóstica se emplean en dos situaciones diferentes:

1. Son una medida que permiten decidir si dos pruebas diagnósticas difieren en su poder de discriminar, correctamente, entre los individuos enfermos y los individuos sanos.
2. Son una medida que permiten ser utilizadas para escoger “el mejor punto de corte” cuando la prueba diagnóstica de interés expresa sus resultados en forma cuantitativa continua. La utilidad de establecer “el mejor punto de corte” en este segundo escenario, es que se consideran como enfermos todos los individuos que tengan un valor mayor al punto de corte y como sanos aquellos individuos que tengan valores menores al punto de corte establecido; o viceversa.

El objetivo de la presente publicación es la de describir un nuevo índice de desempeño global de una prueba diagnóstica, el índice T.

Métodos

Se describieron los fundamentos teóricos y prácticos del índice T y se comparó el índice T frente al índice de Youden, para poder establecer si dos o más modalidades diagnósticas difieren en su capacidad de discriminar, correctamente, entre los individuos verdaderamente enfermos de los sanos.

Se ilustró el cálculo del intervalo de confianza (IC) del 95 % para el índice T y la manera correcta de comparar dos índices T, de dos pruebas diagnósticas diferentes, mediante la prueba “t de Student”.

Se comparó el índice T frente al índice de Youden, para poder establecer si el índice T puede ser utilizado para escoger “el mejor punto de corte” cuando la prueba diagnóstica de interés expresa sus resultados en forma cuantitativa continua.

Resultados

Fundamentos teóricos y prácticos del índice T

La proporción de individuos correctamente clasificados como enfermos por la prueba diagnóstica de interés es $[(a / a + c)]$: verdaderos positivos.

La proporción de individuos correctamente clasificados como sanos por la prueba diagnóstica de interés es $[(d / b + d)]$: verdaderos negativos

Por lo tanto, la proporción de individuos correctamente clasificados por una prueba diagnóstica como enfermos y como sanos es $[(a / a + c)] + [(d / b + d)]$. Esta proporción, es la “proporción de aciertos” de la prueba diagnóstica.

Matemáticamente, la proporción de aciertos escrita como $[(a / a + c)] + [(d / b + d)]$ es igual a $[(a + d) / N]$. Donde N es el número total de individuos enfermos (a + c) más el total de individuos sanos (b + d).

Por otra parte, la proporción de individuos incorrectamente clasificados como sanos por la prueba diagnóstica de interés es $[(c / a + c)]$: falsos negativos.

Y la proporción de individuos incorrectamente clasificados como enfermos por la prueba diagnóstica de interés es $[(b / b + d)$: falsos positivos.

Por lo tanto, la proporción de individuos incorrectamente clasificados por la prueba diagnóstica es $[(c / a + c) + [(b / b + d)]$. Esta proporción, es la *proporción de error* de la prueba diagnóstica.

Matemáticamente, *la proporción de error* escrita como $[(c / a + c) + [(b / b + d)]$ es igual a $[(c + b) / N]$. Donde N es el número total de individuos enfermos (a + c) más el total de individuos sanos (b + d).

Entre mayor sea la proporción de aciertos y menor la proporción de error de una prueba diagnóstica, mejor es su utilidad clínica. Por el contrario, entre menor sea la proporción de aciertos y mayor sea la proporción de error de una prueba diagnóstica, menor es su utilidad clínica.

La diferencia entre la proporción de aciertos y la proporción de error cuantifica la magnitud del buen desempeño de una prueba diagnóstica. Por lo tanto, la utilidad clínica de una prueba diagnóstica se cuantifica mediante la diferencia entre la proporción de aciertos y la proporción de error.

Utilidad de una prueba diagnóstica = (Proporción de aciertos – Proporción de error).

La prueba diagnóstica ideal sería aquella que tendría una alta proporción de aciertos y una baja proporción de error.

La proporción de aciertos (*Pa*) menos la proporción de error (*Pe*) dividido entre la proporción de aciertos, es el índice T.

$$\text{Índice T} = \frac{[(a + d) / N] - [(c + b) / N]}{[(a + d) / N]} \times 100$$

Comparación del índice T frente al índice de Youden para poder establecer si dos o más modalidades diagnósticas difieren en su capacidad de discriminar, correctamente, entre los individuos verdaderamente enfermos y los sanos.

Propiedades sustanciales del índice T

El índice T (IT) tiene las siguientes características importantes:

En primer lugar, si la prueba diagnóstica de interés a evaluar no tiene proporción de error, el índice T será del 100 % de utilidad, como se muestra en la prueba diagnóstica A.

Prueba diagnóstica A. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	480	a	b	0
Negativa	0	c	d	520
Total	480			520

$$\text{Índice T} = \frac{[(480+520) \div 1000] - [(0+0) \div 1000]}{[(480+520) \div 1000]} \times 100 = 100 \%$$

El índice de Youden para la prueba diagnóstica A también es de 100%: $[(100+100) - 100] = 100\%$.

En segundo lugar, si la prueba diagnóstica de interés a evaluar tiene una proporción significativa de error, el índice T será menor del 100 % de utilidad, como se aprecia en el ejemplo de la prueba diagnóstica B.

Prueba diagnóstica B. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	400	a	b	100
Negativa	80	c	d	420
Total	480			520

$$\text{Índice T} = \frac{[(400+420) \div 1000] - [(80+100) \div 1000]}{[(400+420) \div 1000]} \times 100 = 78 \%$$

El índice de Youden para la prueba diagnóstica B también es menor del 100%: $[(83.3+80.7) - 100] = 64\%$

En tercer lugar, el índice T tendrá un valor del 0% si la prueba diagnóstica de interés tiene una proporción de aciertos y de proporción de error iguales, según se muestra en la prueba diagnóstica C.

Prueba diagnóstica C. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	240	a	b	260
Negativa	240	c	d	260
Total	480			520

$$\text{Índice T} = \frac{[(240+260) \div 1000] - [(240+260) \div 1000]}{[(240+260) \div 1000]} \times 100 = 0\%$$

El índice de Youden para la prueba diagnóstica C también es del 0%: $[(50 + 50) - 100] = 0\%$.

En cuarto lugar, si la prueba diagnóstica de interés a evaluar tiene una proporción de error mayor que la proporción de aciertos el índice T tomará valores negativos, y por lo tanto, dicha prueba será una prueba patentemente inútil, como se muestra en la prueba diagnóstica D.

Prueba diagnóstica D. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos	Sanos	
Positiva	230	a	b 320
Negativa	250	c	d 200
Total	480		520

$$\text{Índice T} = \frac{[(230+200) \div 1000] - [(250+320) \div 1000]}{[(230+200) \div 1000]} \times 100 = -32\%$$

Al igual que el índice T, esto mismo sucede con el índice de Youden. El índice de Youden tomará valores negativos cuando la magnitud de error supere la magnitud de aciertos. Calculemos el índice de Youden para la prueba diagnóstica D. La sensibilidad de esta prueba es 47,9 % y su especificidad es de 38,4 %.

$$\text{Índice de Youden} = [(47,9 + 38,4) - 100] = -13,7\%$$

En quinto lugar, acorde a lo expuesto anteriormente, el índice T tiene un rango de valores que van desde el 0 % al 100 % de utilidad y desde 0 hasta ∞ de inutilidad.

De esta forma, podemos concluir que el índice T cumple con el primer objetivo de un índice de desempeño: permite decidir si dos o más pruebas diagnósticas difieren en su poder de discriminar, correctamente, entre los individuos enfermos y los individuos sanos. La prueba diagnóstica A es

la que presenta mejor desempeño (100 % de utilidad); la prueba B es la prueba con segundo mejor desempeño diagnóstico (78 % de utilidad). Las pruebas C y D son las que presenta un peor desempeño, son pruebas diagnósticas indubitadamente inútiles (0 % de utilidad y 32 % de inutilidad).

En sexto lugar, el índice T es independiente del número total de pacientes estudiados, como se puede apreciar en los siguientes ejemplos.

Prueba diagnóstica E. N = 175 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos	Sanos	
Positiva	70	a	b 4
Negativa	1	c	d 100
Total	71		104

$$\text{Índice T} = \frac{[(70+100) \div 175] - [(1+4) \div 175]}{[(70+100) \div 175]} \times 100 = 97,1\%$$

Prueba diagnóstica F. N = 1750 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos	Sanos	
Positiva	700	a	b 40
Negativa	10	c	d 1000
Total	710		1040

$$\text{Índice T} = \frac{[(700+100) \div 1750] - [(10+40) \div 1750]}{[(700+1000) \div 1750]} \times 100 = 97,1\%$$

Al igual que el índice T el índice de Youden es independiente del número total de pacientes estudiados.

El índice de Youden para la prueba diagnóstica E es 94,6 %.

$$\text{Índice de Youden} = [(98,5 + 96,1) - 100] = 94,6\%$$

El índice de Youden para la prueba diagnóstica F es 94,6 %.

$$\text{Índice de Youden} = [(98,5 + 96,1) - 100] = 94,6\%$$

En séptimo lugar, considere las dos pruebas diagnósticas siguientes. La primera prueba tiene una sensibilidad del 90 % y una especificidad del 80 %. La segunda prueba tiene una sensibilidad del 80 % y una especificidad del 90 %. En tal situación, ambas pruebas diagnósticas tendrán un mismo índice T.

Prueba diagnóstica G. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	432	a	b	104
Negativa	48	c	d	416
Total	480			520

$$\text{Índice T} = \frac{[(432+416) \div 1000] - [(48+104) \div 1000]}{[(432+416) \div 1000]} \times 100 = 82\%$$

Prueba diagnóstica H. N = 1000 pacientes.

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	384	a	b	52
Negativa	96	c	d	468
Total	480			520

$$\text{Índice T} = \frac{[(384+468) \div 1000] - [(96+52) \div 1000]}{[(384+468) \div 1000]} \times 100 = 82\%$$

Analizando en detalle estas dos pruebas, podemos darnos cuenta que ambas tienen la misma proporción total de error: 30 % para la prueba diagnóstica G y 30 % para la prueba diagnóstica H. Es por esta razón, que ambas pruebas diagnósticas tienen un mismo índice T.

Los dos porcentajes de error de la prueba G son:

1. En el grupo de enfermos es $(48 \div 480) \times 100 = 10\%$.
2. En el grupo de sanos es $(104 \div 520) \times 100 = 20\%$

La proporción total de error para la prueba G es: $10\% + 20\% = 30\%$

Los dos porcentajes de error de la prueba H son:

1. En el grupo de enfermos es $(96 \div 480) \times 100 = 20\%$
2. En el grupo de sanos es $(52 \div 520) \times 100 = 10\%$

La proporción total de error para la prueba H es $20\% + 10\% = 30\%$.

De nuevo, en este escenario, el índice de Youden se comporta de la misma manera que el índice T.

El índice de Youden para la prueba diagnóstica G es $(90\% + 80\%) - 100 = 70\%$.

El índice de Youden para la prueba diagnóstica H es $(80\% + 90\%) - 100 = 70\%$.

Cálculo del intervalo de confianza (IC) del 95 % para el índice T.

Podemos calcular el intervalo de confianza (IC) del 95 % para el índice T⁶.

IC del 95 % Índice T =

$$\frac{(Pa - Pe) \pm 1.96 \times \sqrt{\frac{PVP \times (1-PVP)}{Ne} + \frac{PVN \times (1-PVN)}{Ns}}}{Pa}$$

Donde PVP es la proporción de verdaderos positivos. PVN es la proporción de verdaderos negativos. Ne es el número total de enfermos catalogados como tal por la prueba estándar. Ns es el número de total de sanos catalogados como tal por la prueba estándar.

El IC del 95 % para el índice T de la prueba diagnóstica B, con un índice T del 78 % es:

Prueba Estándar

Prueba diagnóstica	Enfermos		Sanos	
Positiva	400	a	b	100
Negativa	80	c	d	420
Total	480			520

IC del 95 % Índice T =

$$\frac{(0,82 - 0,18) \pm 1,96 \times \sqrt{\frac{0,83 \times (0,27)}{480} + \frac{0,80 \times (0,20)}{520}}}{0,82}$$

IC del 95 % Índice T = $\frac{(0,64) \pm (0,0545)}{0,82}$

IC del 95 % Índice T = $\frac{(0,64) + (0,0545)}{0,82} = 0,84$ o 84 %

IC del 95 % Índice T = $\frac{(0,64) - (0,0545)}{0,82} = 0,71$ o 71 %

El intervalo de confianza del 95 % para el índice T de la prueba B es del 71 % al 84 %. El Índice T de la prueba B es de 78 %.

Puede observarse que el intervalo de confianza del intervalo T es bastante estrecho, debido a que el número de pacientes de la prueba B es grande (1000 pacientes).

Comparación de dos índices T

La manera correcta de comparar dos índices T, de dos pruebas diagnósticas diferentes, es mediante la prueba “t de Student”⁷.

$$t = \frac{\text{Diferencia entre los índices T}}{\text{Error estándar de sus diferencias}}$$

El Error Estándar de sus diferencias se calcula de la siguiente manera:

$$E.E \text{ diferencias} = \sqrt{(E.E_1)^2 + (E.E_2)^2}$$

El Error Estándar del índice T de una prueba diagnóstica se calcula de la siguiente manera:

$$E. E. \text{ Índice T} = \sqrt{\frac{PVP \times (1 - PVP)}{N_e} + \frac{PVN \times (1 - PVN)}{N_s}}$$

Donde PVP es la proporción de verdaderos positivos. PVN es la proporción de verdaderos negativos. Ne es el número total de enfermos

catalogados como tal por la prueba estándar. Ns es el número de total de sanos catalogados como tal por la prueba estándar.

Con esta información podemos comparar el índice T de la prueba diagnóstica B (índice T = 78 %) versus el índice T de la prueba diagnóstica F (índice T = 97,1 %.)

$$t = \frac{\text{Diferencia entre los índices T}}{\text{Error estándar de sus diferencias}}$$

El Error Estándar del índice T de la prueba diagnóstica B es:

$$\sqrt{\frac{0,83 \times (0,27)}{480} + \frac{0,80 \times (0,20)}{520}} = 0,0278$$

El Error Estándar del índice T de la prueba diagnóstica F es:

$$\sqrt{\frac{0,98 \times (0,02)}{710} + \frac{0,96 \times (0,04)}{1040}} = 0,00803$$

El Error Estándar de la diferencia de la prueba diagnóstica B y F es

$$\sqrt{(0,0278)^2 + (0,00803)^2} = 0,0289$$

$$t = \frac{0,971 - 0,78}{0,0289} = 6,6 \quad p = 0,0001$$

La fórmula para el cálculo del error estándar del índice T es adecuada siempre y cuando el número de sujetos en un grupo sea al menos de 30 individuos⁷.

Uso del índice T para escoger “el mejor punto de corte” cuando la prueba diagnóstica de interés expresa sus resultados en forma cuantitativa continua

La tabla 3, muestra los resultados de los niveles de creatina quinasa (CK) para el diagnóstico del infarto del miocardio⁸.

Tabla 3. Resultados de creatina quinasa (CK) para el diagnóstico de infarto del miocardio.**Infarto del miocardio**

Prueba de CK	Presente	Ausente
≥ 280 U/L	97	1
80 a 279 U/L	118	15
40 a 79 U/L	13	26
1 a 39 U/L	2	88
Total	230	130

La tabla 4 muestra la sensibilidad, la especificidad, el índice de Torregroza y el índice de Youden para cada punto de corte de creatina quinasa (CK)

Tabla 4. Reporte detallado de Sensibilidad (S), Especificidad (E), Índice de Torregroza (IT) y el Índice de Youden (IY) para cada punto de corte de creatina quinasa (CK).

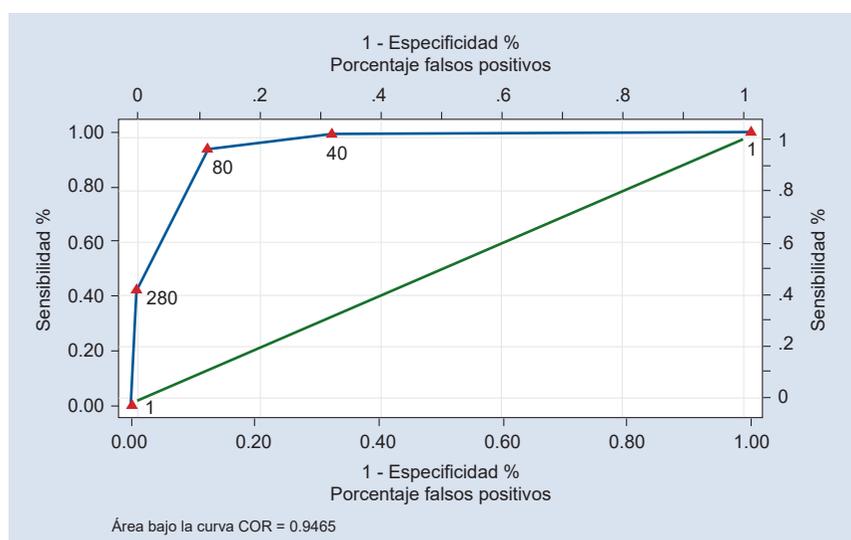
Punto de corte de CK	Sensibilidad	Especificidad	Índice de Torregroza	Índice de Youden
≥ 1	100 %	0 %	43,2 %	0 %
≥ 40	99,13 %	67,69 %	85,9 %	66,82 %
≥ 80	93,48 %	87,69 %	90,4 %	81,17 %
≥280	42,17%	99,23 %	40,5 %	41,4 %

Acorde a esta tabla, el “mejor punto de corte” de CK, determinado por el índice T y el índice de

Youden, es un valor de CK igual o mayor a 80 U/L, el cual tiene un índice T del 90,4 % y un índice de Youden de 81,17 %. De esta manera, unos niveles de CK mayores a 80 U/L predicen la presencia de infarto del miocardio, mientras que unos niveles de CK menores a 80 U/L, descartan la presencia de infarto del miocardio.

Es importante resaltar que tanto el índice T como el índice de Youden identifican el punto de corte que determina la sensibilidad y especificidad más alta conjuntamente; sin embargo, dicho punto de corte no necesariamente determina la sensibilidad ni la especificidad más alta que podría alcanzar la prueba diagnóstica. Obsérvese, que el punto de corte con más alta sensibilidad es el punto de corte ≥ 1 y que el punto de corte con más alta especificidad es el punto de corte ≥ 280 . No obstante, ninguno de estos puntos de corte tiene en conjunto mayor sensibilidad y especificidad. El punto de corte que tiene mayor sensibilidad y especificidad en forma conjunta es el punto de corte de ≥ 80 .

Igualmente, podemos utilizar las propiedades de la curva COR (acrónimo de Característica Operativa del Receptor) para escoger el mejor punto de corte de CK que tenga mayor sensibilidad y especificidad en conjunto. La figura 1 muestra la curva COR para los valores de CK en infarto del miocardio.

**Figura 1.** Curva COR para valores de CK en infarto del miocardio.

La curva COR para una prueba diagnóstica determinada, se configura al ubicar el valor de sensibilidad y el porcentaje de falsos positivos para cada punto de corte de la prueba ⁹. El punto de corte que más se acerque al ángulo superior izquierdo será el mejor punto de corte dado que esta esquina indica una prueba diagnóstica perfecta: una sensibilidad del 100 % y un porcentaje de falsos positivos del 0 % ¹⁰.

Por consiguiente, el punto de corte ≥ 80 U/L es el mejor “punto de corte” para la prueba de CK, dado que se ubica en el ángulo superior izquierdo, es decir, tiene mayor sensibilidad y especificidad en forma conjunta. De esta manera unos niveles de CK mayores a 80 U/L predicen la presencia de infarto del miocardio, mientras que unos niveles de CK menores a 80 U/L, descartan la presencia de infarto del miocardio.

Discusión

El propósito de todo índice de desempeño global de una prueba diagnóstica es el de resumir en una única medida o dato que exprese de manera sensible el poder de la prueba diagnóstica en discriminar, de manera correcta, entre individuos enfermos y los individuos sanos.

El índice T como medida del rendimiento global de una prueba diagnóstica combina las fortalezas del cálculo de la proporción de aciertos (sensibilidad y la especificidad), así como la proporción de error (falsos negativos y falsos positivos), lo cual le permite ser en un indicador global de rendimiento independiente de la prevalencia de la enfermedad.

El índice T tiene un rango de valores que van desde el 0 % al 100 % de utilidad y desde 0 hasta ∞ de inutilidad.

Siempre que dos pruebas diagnósticas tengan las mismas proporciones de error, el índice T y el índice de Youden tendrán el mismo puntaje. Esta propiedad, tanto del índice T como del índice de Youden, tienen un tremendo impacto en la práctica clínica usual. Ante dos pruebas diagnósticas con un mismo índice T y un mismo índice de Youden, cuando se deba decidir cuál prueba a utilizar, es imperativo examinar las propiedades de sensi-

bilidad y especificidad, valor predictivo positivo y valor predictivo negativo de cada prueba por separado. En esta situación en particular, el índice T como el índice de Youden no distingue entre estos dos tipos de errores diagnósticos.

Existe una advertencia universal en la utilidad de todos los indicadores de rendimiento global de las pruebas diagnósticas: su magnitud depende del espectro de la enfermedad, así como de la preselección de los individuos a estudiar mediante el uso de otras pruebas diagnósticas previas.

Es importante puntualizar que el índice T tiene las siguientes ventajas sobre el índice de Youden:

1. La conceptualización del índice T es de fácil entendimiento y asimilación por parte de los clínicos.
2. El cálculo del índice T se deriva directamente de los datos de la tabla de contingencia de la prueba diagnóstica. Así pues, no es necesario calcular primero la sensibilidad y la especificidad para poder calcular el índice T, como si lo hace el índice de Youden.
3. El cálculo del índice T utiliza la información de todas las celdas de la tabla de contingencia de la prueba diagnóstica, VP, VN, FP, FN. Por el contrario, el índice de Youden utiliza solamente la información de dos celdas de la tabla de contingencia, VP y VN.

Conclusión

Al contrastar los resultados del índice T *versus* el índice de Youden y los resultados del índice T *versus* la curva COR, podemos afirmar que el desempeño del índice T es un índice exacto y eficiente que cumple con los dos importantes objetivos de un índice de desempeño, tanto para ser utilizado para decidir si dos pruebas diagnósticas difieren en su poder de discriminar entre los individuos enfermos y los individuos sanos, como para establecer “el mejor punto de corte” cuando la prueba diagnóstica expresa sus resultados en forma cuantitativa continua.

Se debe resaltar que los puntos de corte obtenidos mediante estas pruebas no deben basarse

solamente en el criterio estadístico, sino que en múltiples ocasiones, el punto de corte debe basarse más en lo que se quiere de la prueba, mayor sensibilidad o mayor especificidad.

Cumplimiento de normas éticas

Consentimiento informado: Este estudio es una revisión de la literatura, y como tal, no hay necesidad de un consentimiento informado ni de aprobación del Comité de Ética Institucional.

Conflicto de interés: Ninguno declarado por el autor.

Fuentes de financiación: Los recursos de financiación del proyecto provienen en su totalidad de aportes del autor.

Contribución del autor: El autor declara haber contribuido en la totalidad de la concepción y diseño del estudio, adquisición, análisis e interpretación de datos, redacción y revisión crítica del manuscrito.

Referencias

1. Torregroza-Diazgranados EJ. Pruebas diagnósticas: Fundamentos de los estudios diagnósticos, evaluación de la validez e interpretación clínica de sus resultados. *Rev Colomb Cir.* 2021;36:193-204. <https://doi.org/10.30944/20117582.716>
2. Torregroza-Diazgranados EJ. Pruebas diagnósticas: Razones de probabilidad. *Rev Colomb Cir.* 2021;36:403-10. <https://doi.org/10.30944/20117582.717>
3. Bravo-Grau S., Cruz JP. Estudios de exactitud diagnóstica: Herramientas para su Interpretación. *Rev Chil Radiol.* 2015;21:158-64. <https://doi.org/10.4067/S0717-93082015000400007>
4. Youden WJ. Index for rating diagnostic tests. *Cancer.* 1950;3:32-5. [https://doi.org/10.1002/1097-0142\(1950\)3:1<32::AID-CNCR2820030106>3.0.CO;2-3](https://doi.org/10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3)
5. Glas AS, Lijmer JG, Prins MH, Bossel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol* 2003;56:1129-1235. [https://doi.org/10.1016 / s0895-4356 \(03\) 00177-x](https://doi.org/10.1016 / s0895-4356 (03) 00177-x)
6. Wayne D. Bioestadística. Base para el análisis de las ciencias de la salud. Cuarta edición. México: Limusa Wiley; 2002. p.178-179.
7. Dawson-Saunders B; Trapp RG. Bioestadística médica. Segunda edición. México: Manual moderno; 1999. p. 370.
8. Smith AF. Diagnostic value of serum-creatinine-kinase in a coronary care unit. *Lancet.* 1967;2:178-82. [https://doi.org/10.1016/S0140-6736\(67\)90005-0](https://doi.org/10.1016/S0140-6736(67)90005-0)
9. Altman DG, Bland JM. Diagnostic tests 3: receiver operating characteristic plots. *Br Med J.* 1994;309:188. <https://doi.org/10.1136 / bmj.309.6948.188>
10. Akobeng A. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Pædiatr.* 2007;96:644-7. <https://doi.org/10.1111 / j.1651-2227.2006.00178.x>