



UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
COMISIÓN DE ESTUDIOS DE POSTGRADO
POSTGRADO EN ESTADÍSTICA



**ESTUDIO COMPARATIVO DE LA PRECISIÓN DISCRIMINATORIA PARA PRUEBAS
SEROLÓGICAS APLICANDO CURVAS COR Y METODOLOGÍA ESTÁNDAR**

Autor: Lcdo. Marvin Querales
Tutores: Prof. Luis Pérez-Ybarra
Dra. Ana Rita De Lima

TRABAJO DE GRADO PRESENTADO COMO REQUISITO FINAL PARA OPTAR AL TÍTULO DE
MAGISTER SCIENTIARUM EN ESTADÍSTICA

Maracay, mayo de 2015

“Mientras más sueñas, más lejos logras llegar”

Michael Phelps.

.

AGRADECIMIENTOS

A mis maravillosos tutores: Prof. Ana Rita De Lima y Prof. Luis Pérez Ybarra. Sin ellos, el desarrollo de esta investigación no hubiese sido posible. Un fuerte abrazo para ustedes.

A mi comité consejero: Prof. Harú Martínez de Cordero y Prof. Carlos Espino, por realizar un excelente trabajo en la revisión detallada de este manuscrito.

A todo el personal docente del Postgrado de Estadística de la Facultad de Agronomía, por ser unos facilitadores de altura y parte importante de mi formación profesional.

A todo el personal administrativo que labora tanto en las oficinas del Postgrado de Estadística como en el Control de Estudios de postgrado de la Facultad de Agronomía de la Universidad Central de Venezuela. Gracias por sus consejos y ser excelentes amigas.

A mis compañeros de estudio de la Maestría en Estadística. Fueron un gran apoyo en momentos difíciles. Más que compañeros de estudio, son mis amigos del alma. Los quiero un montón.

ÍNDICE GENERAL

	Página
INTRODUCCIÓN.....	7
Hipótesis de la investigación.....	8
OBJETIVOS.....	9
REVISIÓN DE LA LITERATURA.....	10
METODOLOGIA.....	19
Poblaciones.....	19
Muestras simuladas.....	19
Estimación de umbrales de decisión aplicando la metodología estándar curva COR	21
Cálculo de los valores de sensibilidad y especificidad para las poblaciones y las muestras simuladas.....	22
Comparación de la precisión discriminatoria de las metodologías en estudio en las muestras simuladas.....	23
RESULTADOS.....	25
Cálculo de los parámetros poblacionales, aplicación de la metodología estándar y curva COR y sus valores de sensibilidad y especificidad en las poblaciones consideradas en el estudio.....	25
Muestras simuladas	29
Establecimiento de los umbrales de decisión para las muestras simuladas.....	31
Sensibilidad y especificidad de la metodología estándar y curva COR en las muestras simuladas.....	34
Comparación de la precisión discriminatoria de las metodologías en estudio en las muestras simuladas.....	52
DISCUSIÓN.....	79
CONCLUSIONES.....	83
RECOMENDACIONES.....	84
REFERENCIAS BIBLIOGRÁFICAS.....	85

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
COMISIÓN DE ESTUDIOS DE POSTGRADO
POSTGRADO EN ESTADÍSTICA

**ESTUDIO COMPARATIVO DE LA PRECISIÓN DISCRIMINATORIA PARA PRUEBAS
SEROLÓGICAS APLICANDO CURVAS COR Y METODOLOGÍA ESTÁNDAR**

Autor: Lcdo. Marvin Querales

Tutores: Prof. Luis Pérez-Ybarra

Dra. Ana Rita De Lima

RESUMEN

Las pruebas serológicas permiten la identificación de agentes infecciosos en muestras biológicas. Su habilidad para distinguir entre poblaciones de enfermos y no infectados es evaluada a través de la precisión discriminatoria, expresada en términos de sensibilidad y especificidad y que depende del nivel que se haya elegido como umbral de decisión. La metodología estándar y la curva Característica Operativa del Receptor (COR) permiten estimar estos puntos de corte, siendo indistinto el uso de las mismas en los laboratorios clínicos. Es por ello, que la finalidad de esta investigación fue realizar un estudio comparativo entre ambos procedimientos para evaluar cuál proporciona una mejor precisión discriminatoria. Para ello, se calcularon los parámetros de poblaciones correspondientes a valores de anticuerpos contra tres microorganismos, provenientes de pacientes previamente clasificados como enfermos o no infectados y a partir de los cuales se realizaron varios escenarios de simulaciones en condiciones homocedásticas y heterocedásticas. A estos datos les fue aplicado la metodología estándar y la curva COR para la estimación de los umbrales de decisión y, en cada caso, se calculó: sensibilidad, especificidad, y las proporciones de sensibilidad=100%, especificidad=100% y decisión perfecta (sensibilidad=100% y especificidad=100% de forma simultánea). Los resultados mostraron que la precisión discriminatoria de la prueba empleando la metodología estándar se encuentra desbalanceada favoreciendo la especificidad, mientras que utilizando la curva COR se obtiene un equilibrio óptimo entre sensibilidad y especificidad, siendo éstos valores dinámicos y que aumentan de forma progresiva a medida que la media de la población de enfermos se aleja de la media de la población sana; teniendo, además, una mayor proporción de decisión perfecta. Por ende, los estimadores de la curva COR muestran una mejor precisión discriminatoria y se recomienda su uso como método para la estimación de los umbrales de decisión en las pruebas serológicas.

Palabras clave: prueba serológica, precisión discriminatoria, curva COR, metodología estándar.

UNIVERSIDAD CENTRAL DE VENEZUELA
FACULTAD DE AGRONOMÍA
COMISIÓN DE ESTUDIOS DE POSTGRADO
POSTGRADO EN ESTADÍSTICA

**COMPARATIVE STUDY OF DISCRIMINATORY ACCURACY FOR SEROLOGICAL TESTS USING
ROC CURVE AND STANDARD METHODOLOGY**

Autor: Lcdo. Marvin Querales

Tutores: Prof. Luis Pérez-Ybarra

Dra. Ana Rita De Lima

ABSTRACT

Serologic tests allow the identification of infectious agents in biological samples. His ability to distinguish between populations of patients and uninfected is evaluated through the discriminatory accuracy, expressed in terms of sensitivity and specificity and depending on the level has been chosen as the decision threshold. The standard methodology and Receiver Operating Characteristic (ROC) curve allow estimating these cutoffs, being indiscriminate use them in clinical laboratories. It is therefore the purpose of this research was to conduct a comparative study between the two procedures to assess which provides better discrimination accuracy. For this, population parameters of data values corresponding to three microorganism antibodies from patients previously classified as sick or uninfected were calculated and from which various scenarios from which various scenarios simulations were performed in homoscedastic and heteroskedastic conditions. These data fared applied the standard methodology and the ROC curve for the estimation of decision thresholds and in each case was calculated: sensitivity, specificity, and the proportions of sensitivity=100%, specificity=100% and perfect decision (sensitivity=100% and specificity=100% simultaneously). The results showed that the discriminating accuracy of the test using the standard methodology is unbalanced favoring specificity, whereas, by using the ROC curve the optimum balance between sensitivity and specificity is obtained, with these dynamic values and increase progressively as the average population of patients away from the middle of the healthy population; further having a higher proportion of perfect decision. Therefore, the estimates of the ROC curve show better accuracy and discriminatory use as a method for estimating decision thresholds in serological tests are recommended.

Keywords: serological test, discriminatory accuracy, ROC curve, standard methodology.

INTRODUCCIÓN

Las pruebas serológicas forman parte de los procedimientos diagnósticos de los laboratorios clínicos y de investigación y se utilizan para la identificación de los agentes infecciosos presentes en las muestras biológicas. Se basan en reacciones inmunológicas de tipo antígeno-anticuerpo, en donde los antígenos o anticuerpos presentes en la muestra se ponen en contacto con reactivos químicos. Un reactivo que contenga anticuerpos detectará antígenos; en cambio, si contiene antígenos, detectará los anticuerpos respectivos presentes en la muestra. La identificación de los antígenos o anticuerpos permite al analista establecer el diagnóstico de una infección en curso o que se ha tenido previamente (1).

Algunas pruebas serológicas (conocidas también simplemente como pruebas diagnósticas) son cualitativas, pues sólo ponen de manifiesto la presencia o ausencia de un determinado antígeno o anticuerpo. Otras son cuantitativas ya que miden la concentración de un antígeno o anticuerpo en particular. Las pruebas cualitativas son más específicas y fiables que las cuantitativas; sin embargo, éstas últimas facilitan el seguimiento de la evolución de una enfermedad en la que aumentan o disminuyen los niveles de antígeno o anticuerpo.

En el caso de las pruebas cuantitativas, los resultados se expresan en una escala continua y para ayudar a determinar la presencia o ausencia de la enfermedad, se elige un punto de corte o umbral de decisión. Los resultados que se encuentran en uno de los lados de este umbral, por ejemplo por encima, se pueden considerar positivos para la enfermedad, mientras que los resultados que estén por debajo del punto de corte son considerados como negativos. Sin embargo, se sabe que en tal situación no todos los pacientes cuyos resultados estén por encima del punto de corte tendrán necesariamente enfermedad y no todos aquellos cuyos resultados estén por debajo del punto de corte estarán libres de esta. Para ello, juega un papel muy importante la precisión discriminatoria de la prueba diagnóstica, la cual se caracteriza por su sensibilidad y especificidad; valores éstos que dependen del nivel que se haya elegido como umbral de decisión para el establecimiento de la positividad de la prueba (2).

El método clásico o estándar para determinar el punto de corte es fijarlo a dos o tres desviaciones estándar del valor promedio de un conjunto de datos obtenidos a partir de un grupo de pacientes ya clasificados previamente como negativos para una determinada condición o enfermedad, bajo el supuesto de independencia (3). Sin embargo, este método es arbitrario y no es apropiado cuando se viola dicho supuesto (4). Una variante

de esta metodología es emplear la media y tres desviaciones típicas equilibradas, las cuales se calculan eliminando el 5% de los valores mínimos y el 5% de los valores máximos del conjunto de datos obtenidos en pacientes clasificados previamente como negativos.

Otra metodología utilizada es la curva Característica Operativa del Receptor (COR). La misma es ampliamente aceptada como un método para seleccionar un punto de corte óptimo para una prueba serológica y para comparar la precisión de varias pruebas diagnósticas (5,6). La curva es generada por el trazado de la *sensibilidad* de todos los posibles puntos de corte para la prueba en el eje de las ordenadas y como una función de *1-especificidad* en el eje de las abscisas. Idealmente, el mejor punto de corte proporciona tanto la sensibilidad como la especificidad más alta y fácilmente se ubica en la curva en el punto más alto en el eje vertical y el más a la izquierda en el eje horizontal.

Si bien la curva COR aporta ventajas en el establecimiento de los umbrales de decisión (5,7), su uso, al igual que el criterio de la metodología clásica basado en las desviaciones estándar de los datos, se realiza de forma indistinta debido a que el establecimiento de los puntos de corte no está completamente estandarizado. Es por ello que, la finalidad de esta investigación fue realizar un estudio comparativo entre ambos procedimientos para evaluar cuál proporciona una mayor sensibilidad y especificidad en el diagnóstico serológico de enfermedades infecciosas y así contribuir con la optimización de las pruebas inmunológicas en los laboratorios clínicos y de investigación.

Hipótesis de la investigación

Los umbrales de decisión obtenidos mediante la curva COR generan valores de sensibilidad y especificidad mejores y más equilibrados que los puntos de corte estimados por medio de la metodología estándar. Esto debido a que involucra para su cálculo poblaciones tanto de pacientes no infectados como con presencia del agente infeccioso, ya previamente clasificados como tales, mientras que la metodología estándar sólo considera la población de pacientes negativos a la infección.

OBJETIVOS

Objetivo General

Evaluar la precisión discriminatoria de las pruebas serológicas utilizando curvas COR y metodología estándar.

Objetivos Específicos

- Comparar los umbrales de decisión obtenidos al aplicar las curvas COR y metodología estándar en poblaciones constituidas por resultados de pruebas serológicas.
- Contrastar la sensibilidad y especificidad de las pruebas serológicas aplicando curvas COR y metodología estándar en poblaciones constituidas por resultados de pruebas serológicas.
- Comparar los umbrales de decisión obtenidos al aplicar las curvas COR y metodología estándar en muestras simuladas de resultados de pruebas serológicas bajo condiciones de homocedasticidad y heterocedasticidad.
- Contrastar la sensibilidad y especificidad de las pruebas serológicas aplicando curvas COR y metodología estándar en muestras simuladas de resultados de pruebas serológicas bajo condiciones de homocedasticidad y heterocedasticidad.
- Evaluar la precisión discriminatoria de ambas metodologías en muestras simuladas de resultados de pruebas serológicas bajo condiciones de homocedasticidad y heterocedasticidad.

REVISIÓN DE LA LITERATURA

En una revisión bibliográfica realizada en el año 2002 se exponen los conceptos básicos relacionados con las pruebas de diagnóstico en general (8). Los autores señalan que dichas pruebas se utilizan para clasificar a los pacientes, determinar la gravedad de una determinada enfermedad, pronosticar, evaluar la respuesta a tratamientos y, lo más importante, ayudar a establecer diagnósticos. Señalan que la premisa básica de la prueba de diagnóstico es que existen dos poblaciones de personas (las que tienen la enfermedad en cuestión y las que no) que difieren en al menos una variable comprobable. Por ejemplo, los pacientes con neumonía presentan infiltrados en la radiografía, mientras que aquellos sin neumonía no los manifiestan. En la práctica, sin embargo, este análisis no es tan simple: no todas las personas con neumonía tiene un infiltrado y no todos los pacientes con un infiltrado tienen neumonía. Esta variabilidad en el estado real del paciente y el resultado de la prueba diagnóstica origina una superposición entre los resultados para las poblaciones enfermas y normales para prácticamente todas las pruebas diagnósticas, la cual es mostrada esquemáticamente por los autores en la figura 1.

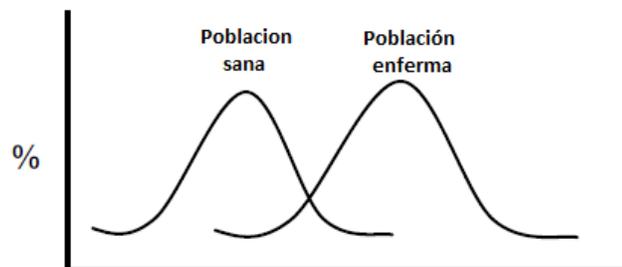


Figura 1. Distribución hipotética de un marcador biológico tanto en poblaciones con ausencia y presencia de una determinada enfermedad. **Fuente:** Worster y col (8).

La figura 1 muestra que algunos niveles de la variable que se está evaluando son compatibles tanto para la población sana como para la población que tenga presente la enfermedad. Esto pone en evidencia que los resultados no siempre pueden ser aceptados por su valor nominal, pues siempre está asociado a un margen de clasificación erróneo. La precisión del diagnóstico de una prueba serológica es el componente fundamental para su aplicación clínica. La habilidad de una prueba serológica para distinguir entre las poblaciones enferma y aquella que esté libre de la enfermedad comúnmente es evaluada a través de la precisión discriminatoria de la prueba diagnóstica. La misma es

comúnmente evaluada a través de que tan bien la técnica clasifica correctamente a un grupo de n_1 individuos conocidos previamente como enfermos ($D = 1$) y a un grupo n_0 de sujetos identificados con ausencia de la enfermedad ($D = 0$) (9). Si la prueba diagnóstica proporciona resultados dicotómicos, positivos ($T = 1$) o negativos ($T = 0$), los resultados pueden ser resumidos en una tabla de contingencia 2×2 que muestre los resultados de la prueba versus el estado real del paciente, ver tabla 1.

Tabla 1. Tabla de contingencia 2×2 que resume los resultados obtenidos de un estudio dicotómico de precisión diagnóstica.

		Estado real de la enfermedad en el paciente		
		Presente ($D = 1$)	Ausente ($D = 0$)	
Resultado de la prueba	Positivo ($T = 1$)	Verdaderos positivos (VP)	Falsos positivos (FP)	m_1
	Negativo ($T = 0$)	Falsos negativos (FN)	Verdaderos negativos (VN)	m_0
Total		n_1	n_0	N

Fuente: Shapiro (9).

En la literatura médica, la precisión discriminatoria es comúnmente medida a través de la probabilidad condicional de clasificar correctamente a un paciente enfermo cuando presenta la enfermedad (10-12), es decir:

$$P(T = 1/D = 1) = VP/n_1$$

lo que se conoce como razón de verdaderos positivos (TPR) o sensibilidad (Se).

También se utiliza la probabilidad condicional de clasificar a un paciente como sano cuando éste no presente la enfermedad:

$$P(T = 0/D = 0) = VN/n_0$$

conocida como proporción de verdaderos negativos (TNR) o especificidad (Sp).

Equivalentemente, la probabilidad de cometer el error tipo I

$$P(T = 1/D = 0) = \frac{FP}{n_0} = 1 - \text{especificidad}$$

se conoce como falso positivo o proporción de falsos positivos (FPR); y el error tipo II

$$P(T = 0/D = 1) = \frac{FN}{n_1} = 1 - \text{sensibilidad}$$

como falso negativo o razón de falsos negativos (FNR).

Es necesario mencionar que la evaluación de la precisión discriminatoria se centra en el rendimiento de la prueba diagnóstica cuando se conoce el estado de la enfermedad; sin embargo, al interpretar el resultado de la misma para un paciente individual, por lo general el estado real de la enfermedad es desconocido, razón por la cual el resultado de la prueba serológica se emplea para estimar probabilidades posteriores de padecer la enfermedad. En el caso de resultados dicotómicos, las mismas corresponden a la probabilidad posterior de padecer la enfermedad dado un resultado positivo de la prueba

$$P(T = 1/D = 1) = VP/m_1$$

conocida como valor predictivo positivo (PPV). De forma relativa, la probabilidad posterior de no padecer la enfermedad dado un resultado negativo

$$P(T = 0/D = 0) = VN/m_0$$

es conocida como valor predictivo negativo (NPV). Ambos, son de mucha importancia desde el punto de vista clínico de abordaje al paciente.

Muchas pruebas de diagnóstico de laboratorio dan un resultado cuantitativo x en lugar de un resultado dicotómico. El grado de superposición en las distribuciones de los resultados de la prueba para los sujetos enfermos $[f(x|D = 1)]$ y sujetos sanos $[f(x|D = 0)]$ determina la capacidad de la prueba para distinguir entre ambas poblaciones, ver figura 2. En situaciones clínicas diagnósticas donde se requiera separar a los pacientes en dos grupos (sanos o negativos y enfermos o positivos) es común definir umbrales de decisión o puntos de corte γ y clasificar a los pacientes como enfermos o positivos si $x > \gamma$ y sano o negativo si $x < \gamma$.

Cada umbral de decisión genera una tabla 2x2 diferente para los resultados de la prueba dicotomizada versus el estado real del paciente, por lo que la sensibilidad y especificidad pueden ser estimadas para cada punto de corte γ . Sin embargo, cuando decrece el valor de γ , aumenta la sensibilidad pero la especificidad disminuye, lo que indica que hay una relación inversa que debe ser tomada en cuenta cuando se evalúa la precisión discriminatoria de una prueba serológica.

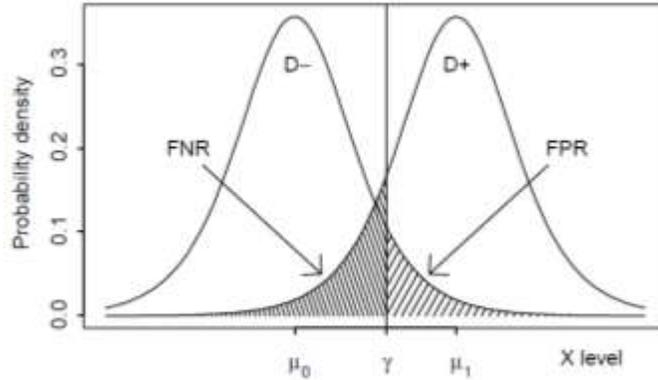


Figura 2. Distribución hipotética de los resultados x de una prueba diagnóstica para sujetos con presencia y ausencia de enfermedad. La línea vertical $x = \gamma$ indica el umbral de decisión para una prueba positiva. El área sombreada a la derecha de γ corresponde a la razón de falsos positivos (FPR). El área sombreada a la izquierda de γ corresponde a la razón de falsos negativos (FNR). **Fuente:** Shapiro (9).

Un resumen útil de precisión discriminatoria lo constituye un gráfico que representa, sobre un sistema de coordenadas, la sensibilidad y la especificidad, conocido como curva Característica Operativa del Receptor (COR) (13). En el mismo, se tiene una medida determinada X realizada sobre una población de positivos X_P , y otra de negativos X_N , con funciones de distribución acumulativa G y F , respectivamente. Suponiendo que $E(X_N) \leq E(X_P)$, para clasificar a los individuos en uno u otro grupo, se debe fijar un criterio, punto de corte, a partir del cual un individuo será considerado positivo. Por tanto, fijado un punto de corte t , la sensibilidad de la prueba vendrá determinada por $1 - G(t)$, siendo $F(t)$ su especificidad y quedando por tanto determinada la curva COR por las coordenadas del vector $(1 - F(t), 1 - G(t))$, o, equivalentemente, por la función

$$1 - G(F^{-1}(t)), t \in [0,1]$$

A modo de ejemplo, en la figura 3 se presenta un modelo de tres curvas COR.

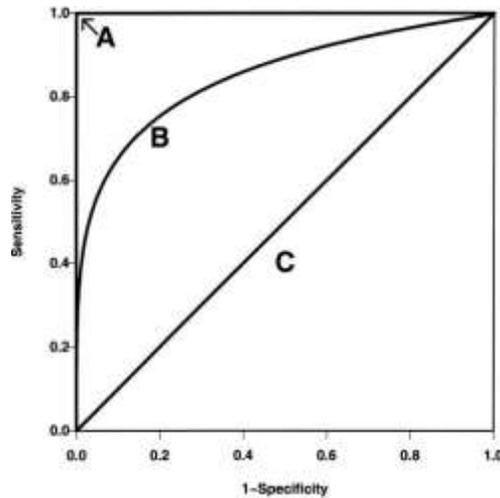


Figura 3. Representación gráfica de 3 curvas COR hipotéticas. La esquina superior izquierda (A) de la gráfica representa la discriminación perfecta (TPR = 1, FPR = 0), mientras que la línea diagonal (C) donde TPR=FPR representa una discriminación muy baja. La línea (B), por otra parte, muestra el comportamiento de una prueba con una discriminación intermedia. **Fuente:** Grzybowski y Younger (14).

Son dos los métodos comúnmente utilizados para identificar los umbrales de decisión óptimos empleando la sensibilidad y especificidad de una curva COR (15, 16). El primer método, mínima distancia cuadrática (MDC), asume que el mejor umbral de decisión para optimizar y mantener un equilibrio entre la sensibilidad y especificidad de una prueba es el punto sobre la curva cercano al punto (0,1). En este método, la sensibilidad y especificidad óptima están definidas como aquellos puntos que generen el valor mínimo para

$$(1 - sensibilidad)^2 + (1 - especificidad)^2$$

El segundo método usado es el índice de Youden (J), el cual es definido como la distancia vertical máxima (dentro de un rango de 0 y 1) entre la curva COR y la línea diagonal y es calculado como:

$$J = \text{máximo}\{sensibilidad + especificidad - 1\}$$

Usando esta medida, el valor del umbral de decisión sobre la curva COR corresponderá aquel valor que maximice la expresión de J . Una interpretación intuitiva de J sería aquel valor que coincida con el punto sobre la curva más alejado de la diagonal, ver figura 4.

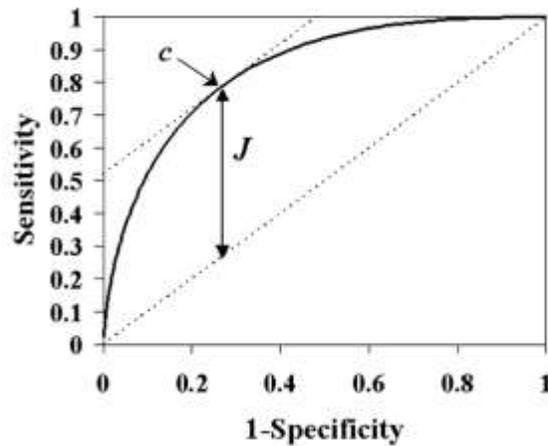


Figura 4. Curva Característica Operativa del Receptor (COR), de un conjunto de datos correspondiente a valores de Interleuquina-6 que representa el índice de Youden (J) y el umbral de decisión óptimo (c). **Fuente:** Schisterman y col (17).

El punto de corte correspondiente a estos valores de sensibilidad y especificidad corresponden al mejor umbral de decisión para diferenciar personas con alguna enfermedad de los individuos sin esa condición según uno y otro método.

Por otro lado, existen parámetros que pueden ser evaluados en una curva COR para estudiar la precisión diagnóstica de una prueba serológica. Entre ellos, el área bajo la curva (AUC) de una curva COR, representa la mejor medida del poder discriminatorio de una prueba diagnóstica (18). AUC se define como:

$$AUC = \int_0^1 (1 - G(F^{-1}(1 - t))) dt$$

Nótese que sólo haciendo un cambio de variable se obtiene:

$$\begin{aligned} AUC &= 1 - \int_0^1 G(F^{-1}(t)) dt \\ &= 1 - \int G(S) dF(S) = 1 - \iint_{-\infty}^y dG(X) dF(Y) = P\{X_P > X_N\} = \theta \end{aligned}$$

Este valor de θ se puede interpretar como la probabilidad de que en un par aleatorio de individuos sanos y enfermos se obtengan valores más elevados para los individuos enfermos (19). Valores de AUC cercanos a 1,0 indican que el marcador evaluado presenta una elevada precisión diagnóstica, mientras que un AUC de 0,5 (aquella donde la curva se

superpone sobre la línea diagonal), la prueba tiene una escasa precisión diagnóstica. Hay varias escalas para la interpretación del AUC pero, en general, curvas COR con una AUC $\leq 0,75$ no son clínicamente usadas, mientras que un AUC de 0,97 o más tiene un elevado valor clínico (7).

Los orígenes de las curvas COR se remontan a la Segunda Guerra Mundial, en donde se desarrollaron como herramientas para el estudio de detección e interpretación de señales de radar (20). El objetivo de los operadores de radar era distinguir las verdaderas señales del ruido de fondo, debido a que al variar la amplitud de la señal definida para considerar un pico como misil, varían la sensibilidad y la especificidad del sistema de recepción. Para ello, establecieron un umbral de decisión k que representaba el valor de intensidad de sonido por encima del cual la señal era considerada como verdadera. La selección de este punto de corte está sujeta al valor que mejor optimice el resultado de sensibilidad y especificidad, considerando de antemano que se produce un solapamiento entre las verdaderas señales y el ruido. De ahí que dicha metodología sea aplicada a las pruebas diagnósticas en donde existe un solapamiento entre los resultados de los pacientes con una condición particular y los de aquellos que no la tienen (21).

Las aplicaciones de las curvas COR se han extendido a múltiples sistemas diagnósticos, incluidas las técnicas radiológicas (22) y las pruebas de laboratorio (23), y su denominación se está sustituyendo cada vez más por «*Performance Characteristic Curves*» o curvas de rendimiento diagnóstico. Estas curvas proporcionan un criterio unificador en el proceso de evaluación de una prueba diagnóstica serológica debido a sus diversas aplicaciones y varias ventajas (24), entre las que destacan:

1. Son una representación fácilmente comprensible de la capacidad de discriminación de la prueba en todo el rango de puntos de corte.
2. Son simples, gráficas y fáciles de interpretar visualmente.
3. No requieren un nivel de decisión particular porque está incluido todo el espectro de puntos de corte.
4. Son independientes de la prevalencia, ya que la sensibilidad y la especificidad se obtienen en distintos subgrupos. Por tanto, no es necesario tener cuidado para obtener muestras con prevalencia representativa de la población.
5. Proporcionan una comparación visual directa entre pruebas en una escala común, mientras que otro tipo de gráficos, como los diagramas de puntos o los histogramas de frecuencias, requieren diferentes gráficos cuando difieren las escalas.
6. La especificidad y la sensibilidad son accesibles en el gráfico, en contraste con los diagramas de puntos y los histogramas.

Sin embargo, las curvas COR presentan ciertas desventajas, entre las cuales se pueden citar:

1. Los puntos de corte, aunque son conocidos y se han usado para generar el gráfico, no aparecen en él, a diferencia de los diagramas de puntos y los histogramas. Al observar un punto de la curva se observan la sensibilidad y la especificidad que tiene asociadas, pero no se conoce el valor concreto de dicho punto.
2. No se muestra el número de sujetos, y a medida que el tamaño de la muestra decrece, la representación gráfica tiende a volverse progresivamente mellada y desigual.
3. La generación de la figura y el cálculo de los parámetros es difícilmente manejable sin programas de computación, los que no están ampliamente disponibles.
4. No tienen aplicación cuando los resultados de la prueba son medidos en una escala dicotómica.

Las curvas COR han tenido un incremento considerable en su aplicación en diversas áreas de la medicina como una herramienta en las técnicas diagnósticas. No obstante, su uso no es generalizado, pues aún no se cuenta con un criterio universal para la selección de los puntos de corte que permitan distinguir a las personas con una determinada enfermedad de las que no la padezcan. De hecho, Ridge y Vizard (4), refieren que los valores de umbrales de decisión establecidos para las pruebas diagnósticas en la mayoría de los casos han sido determinados por métodos arbitrarios que no toman en cuenta las consecuencias económicas y sociales que implica un mal diagnóstico. Estos autores hicieron una revisión de 21 artículos de investigación publicados en la revista *Journal of Clinical Microbiology* en el año 1991 en los cuales describieron el desarrollo, evaluación o implementación de la técnica serológica ELISA (inmunoensayo enzimático), encontrando que 12 de los 21 artículos establecieron el valor de punto de corte basándose en estadísticos arbitrarios, tal como la metodología estándar de calcular 3 desviaciones estándar sobre la media de los resultados obtenidos en los controles negativos. Este procedimiento, según los autores, le añade subjetividad al análisis de los resultados, no teniendo en cuenta a su vez la especificidad y sensibilidad de la prueba y el número o los costos de diagnósticos erróneos producidos por los ajustes de la metodología. Destacan la importancia de mejorar el cálculo de los umbrales de decisión para clasificar correctamente a un individuo si presenta o no una determinada enfermedad.

Aún y cuando, la metodología estándar es la más comúnmente utilizada en el laboratorio clínico para la determinación de los umbrales de decisión, son escasas las investigaciones que comparen su eficiencia con otras metodologías, tal como lo hicieron Kim, Choi, Chung, Rha y Kim (25). Estos investigadores partieron de la premisa de que la metodología

estándar para determinar el valor de los umbrales de decisión puede verse afectado cuando no se cumple con el supuesto de independencia. Razón por la cual, propusieron un método alternativo para el cálculo de puntos de corte en caso de no cumplimiento de este supuesto utilizando un modelo lineal de efectos mixtos. Para una tasa de error de tipo fijo calcularon la probabilidad condicional de que la prueba es positiva dado que el sujeto tiene la enfermedad y luego compararon la sensibilidad de este método con la metodología clásica, observando que el primero mostraba mayor eficiencia. Esta investigación destaca las limitantes de la metodología clásica en la evaluación de datos de biomarcadores, debido a que muestran que sus valores de sensibilidad son menores a los obtenidos con el modelo lineal de efectos fijos mixtos. En vista de estos resultados, la evaluación del grado de una determinada enfermedad empleando la metodología estándar para el cálculo en el punto de corte, no posee diferencias estadísticamente significativas con un grupo control de pacientes sanos.

METODOLOGÍA

Poblaciones

Se utilizaron tres poblaciones correspondientes a los títulos de anticuerpos de tipo IgG contra los microorganismos *Trypanosoma cruzi*, *Leishmania sp* y *Toxoplasma sp*, medidos en forma de densidad óptica (DO), y suministrados por el Instituto de Biología Molecular de Parásitos (InBioMolP) y por el Departamento de Parasitología de la Universidad de Carabobo, provenientes del archivo de muestras de pacientes procesadas en ambas dependencias desde el año 1992 hasta el año 2014. Estos valores de DO incluyeron observaciones tanto de pacientes sanos como enfermos, ambos confirmados por pruebas y criterios diagnósticos preestablecidos. Entre estos criterios se encuentra que, para considerar a un paciente con la enfermedad, el mismo debe resultar positivo en por lo menos dos de tres pruebas diagnósticas, dentro de las cuales las más comunes son: ELISA, hemaglutinación y Western Blott.

Para cada una de estas poblaciones se calculó el valor de los parámetros media y varianza tanto para los resultados de las pruebas serológicas de pacientes sanos, μ_S y σ_S^2 , como de pacientes enfermos, μ_E y σ_E^2 .

Muestras simuladas

La generación de los datos simulados se hizo con el programa Excel en su versión para Windows 2010 por medio de la función generación de números aleatorios que se encuentra en el complemento “Análisis de datos” (27).

Los valores que definieron los parámetros de las poblaciones de donde se generaron las muestras simuladas se obtuvieron de la caracterización de las poblaciones anteriormente descritas.

Se definió una población de pacientes sanos P_S con media μ_S y varianza σ_S^2 , y tres conjuntos de cinco poblaciones de pacientes enfermos P_E , a saber, un conjunto con la misma varianza de la población de pacientes sanos, es decir, homocedástico, y dos conjuntos con varianzas poblacionales diferentes a la de la población de pacientes sanos, es decir, heterocedásticos, lo cual generó un total de 16 poblaciones simuladas para cada enfermedad, y dado que el trabajo contempló datos de 3 enfermedades, se simularon 48 poblaciones en total.

Para la condición de homocedasticidad, se consideró como varianza de las poblaciones simuladas de pacientes enfermos la obtenida de la caracterización de la población de pacientes sanos σ_S^2 . En el primero de los conjuntos heterocedásticos se fijó como varianza de las poblaciones simuladas de pacientes enfermos la resultante de la caracterización de

la población pacientes enfermos σ_E^2 , mientras que en el segundo de los conjuntos heterocedásticos la varianza correspondió a la varianza conjunta entre poblaciones sanas y enfermas σ_c^2 .

Independientemente de la condición de homocedasticidad o heterocedasticidad de las poblaciones simuladas, las medias de las poblaciones simuladas de pacientes enfermos se definieron como función de la media de los pacientes sanos μ_S y la desviación estándar conjunta σ_c , por cuanto que estas combinaciones de media y desviación estándar presentaron el comportamiento más equilibrado. Así las medias para los pacientes enfermos quedaron definidas por P_{E1} con $\mu_{E1} = \mu_S + 0,5\sigma_c$; P_{E2} con $\mu_{E2} = \mu_S + \sigma_c$; P_{E3} con $\mu_{E3} = \mu_S + 2\sigma_c$; P_{E4} con $\mu_{E4} = \mu_S + 3\sigma_c$; P_{E5} con $\mu_{E5} = \mu_S + 4\sigma_c$; esto con el fin de considerar poblaciones de pacientes enfermos con medias cada vez más alejadas del valor promedio de la población sana.

Para cada población de pacientes sanos P_S y pacientes enfermos P_E , se generaron un total de $n^* = 1000$ muestras aleatorias, cada una de las cuales constó de n_S observaciones simuladas provenientes de P_S , y n_E observaciones simuladas provenientes de las P_{Ei} , $i = 1,2,\dots,5$ para cada una de las poblaciones de enfermos consideradas. Por simplicidad y comodidad el tamaño de n_S y n_E se fijó en 30 observaciones. Esto obedece a que este es el tamaño de muestra más utilizado en las pruebas con datos reales para calcular la sensibilidad y especificidad de una prueba serológica a partir de una prueba de oro.

El esquema de la simulación conducida en este trabajo se resume la tabla 2.

Tabla 2. Esquema de la simulación realizada en la investigación.

Población	Condición de igualdad de varianzas					
	Homocedásticas		Heterocedásticas 1		Heterocedásticas 2	
	Media	Varianza	Media	Varianza	Media	Varianza
Sanos	μ_S	σ_S^2	μ_S	σ_S^2	μ_S	σ_S^2
Enfermos 1	$\mu_S + 0,5\sigma_c$		$\mu_S + 0,5\sigma_c$	σ_E^2	$\mu_S + 0,5\sigma_c$	σ_c^2
Enfermos 2	$\mu_S + \sigma_c$		$\mu_S + \sigma_c$		$\mu_S + \sigma_c$	
Enfermos 3	$\mu_S + 2\sigma_c$		$\mu_S + 2\sigma_c$		$\mu_S + 2\sigma_c$	
Enfermos 4	$\mu_S + 3\sigma_c$		$\mu_S + 3\sigma_c$		$\mu_S + 3\sigma_c$	
Enfermos 5	$\mu_S + 4\sigma_c$		$\mu_S + 4\sigma_c$		$\mu_S + 4\sigma_c$	

Se denominó escenario al conjunto formado por los resultados tanto de pacientes sanos como de pacientes enfermos. Su creación fue necesaria por cuanto que para el cálculo de la sensibilidad y especificidad se necesitan datos provenientes de ambas poblaciones, sanos y enfermos, por lo cual, para cada una de las condiciones de igualdad de varianza se construyeron cinco escenarios, a saber

Escenario	Poblaciones
I	P_S y P_{E1}
II	P_S y P_{E2}
III	P_S y P_{E3}
IV	P_S y P_{E4}
V	P_S y P_{E5}

Lo cual generó 15 escenarios con datos simulados para cada enfermedad, y un total de 45 escenarios para el trabajo.

Estimación de umbrales de decisión aplicando la metodología estándar y curva COR

Tanto para las poblaciones como para los escenarios simulados, se le aplicó la metodología estándar y curva COR para obtener los respectivos valores de umbrales de decisión o valores críticos (V^c). En el caso de la metodología estándar se usaron cuatro criterios de establecimiento:

Metodología estándar 1 (Est1):	<i>media + dos desviaciones estándar</i>
Metodología estándar 2 (Est2):	<i>media + tres desviaciones estándar</i>
Metodología estándar 3 (Est3):	<i>media equilibrada + dos desviaciones estándar equilibrada</i>
Metodología estándar 4 (Est4):	<i>media equilibrada + tres desviaciones estándar equilibrada</i>

Esto debido a que las cuatro metodologías son utilizadas comúnmente y de forma indistinta en los laboratorios clínicos. La media equilibrada y desviación estándar equilibrada se calcularon tomando los valores de densidad óptica una vez eliminados el 5% de los valores mínimos y el 5% de los valores máximos, previamente ordenados.

Los umbrales de decisión empleando curvas COR se calcularon empleando tanto la mínima distancia cuadrática (MDC) como el índice de Youden.

Los umbrales de decisión obtenidos por la prueba p , $p = 1, 2, \dots, 6$, (4 empleando metodología estándar y 2 empleando curvas COR), en el escenario i y la muestra j fueron expresados en tablas para su comparación.

Cálculo de los valores de sensibilidad y especificidad para las poblaciones y las muestras simuladas

Para las poblaciones consideradas y para todos los escenarios contemplados en la simulación, una observación fue declarada como sana por la prueba p , $p = 1, 2, \dots, 6$, (4 empleando metodología estándar y 2 empleando curvas COR), si $y_{ijk} \leq V_{ijp}^c$, y como enferma en caso contrario, es decir, si $y_{ijk} > V_{ijp}^c$, siendo V_{ijp}^c el valor crítico de la prueba p , en el escenario i y la muestra j .

Con los datos ya clasificados se realizó el cálculo de sensibilidad y especificidad para las poblaciones consideradas y para cada una de las muestras simuladas de acuerdo a las siguientes ecuaciones, dadas por la tabla 1:

$$\%sensibilidad = \frac{\text{verdaderos positivos}}{\text{total de enfermos}} \times 100$$

$$\%especificidad = \frac{\text{verdaderos negativos}}{\text{total de sanos}} \times 100$$

Siendo los verdaderos positivos aquellos resultados de la población de pacientes enfermos declarados positivos por la prueba p , y verdaderos negativos aquellos resultados de la población de pacientes sanos declarados negativos por la prueba p .

Para las poblaciones consideradas, los resultados de sensibilidad y especificidad para cada enfermedad fueron expresados en tablas para su comparación. Por su parte, para establecer comparaciones entre las metodologías aplicadas en las poblaciones simuladas se realizaron gráficos en los cuales se representaron los valores del estimador de sensibilidad y especificidad obtenidos por ambas metodologías en las n^* muestras simuladas; siendo éste definido como el promedio de dichos valores sobre las n^* muestras simuladas, es decir

$$\widehat{\%especificidad} = \frac{1}{n^*} \sum_{j=1}^{n^*} \%especificidad_{ijp}$$

$$\widehat{\%sensibilidad} = \frac{1}{n^*} \sum_{j=1}^{n^*} \%sensibilidad_{ijp}$$

Comparación de la precisión discriminatoria de las metodologías en estudio en las muestras simuladas

Para comparar la precisión discriminatoria entre las metodologías en estudio, además de la estimación de la sensibilidad y especificidad, para cada uno de los escenarios y pruebas consideradas, se contó en cuantas muestras las pruebas presentaron sensibilidad y especificidad igual al 100%, y se estimaron las proporciones de tales resultados. En este sentido, para el par j -ésimo de muestras $\left\{ \{y_{0jp}\}; \{y_{ijp}\}_{i \geq 1} \right\}$, que constituye el conjunto de pacientes sanos y enfermos simulados en el escenario i , sobre la cual se aplica la prueba p , se tiene

$\forall \left\{ \{y_{0jp}\}; \{y_{ijp}\}_{i \geq 1} \right\}$	$Z_{Sijp} = 1; \text{ si } \% \text{especificidad}_{ijp} = 100\%$
	$Z_{Sijp} = 0; \text{ en otro caso}$
	$Z_{Eijp} = 1; \text{ si } \% \text{sensibilidad}_{ijp} = 100\%$
	$Z_{Eijp} = 0; \text{ en otro caso}$

Por lo tanto, las proporciones de muestras para los diferentes escenarios considerados en las cuales las pruebas presentaron especificidad y sensibilidad igual a 100%, denotadas por $P_{\{\% \text{especificidad}_{ijp}=100\%\}}$ y $P_{\{\% \text{sensibilidad}_{ijp}=100\%\}}$, respectivamente, vienen dadas por

$$P_{\{\% \text{especificidad}_{ijp}=100\%\}} = \frac{1}{n^*} \sum_{j=1}^{n^*} Z_{Sijp}$$

$$P_{\{\% \text{sensibilidad}_{ijp}=100\%\}} = \frac{1}{n^*} \sum_{j=1}^{n^*} Z_{Eijp}$$

Por último, para cada uno de los escenarios y pruebas consideradas, se contó en cuantas muestras las pruebas presentaron sensibilidad y especificidad igual a 100% simultáneamente; resultado este que puede denominarse decisión perfecta, y se estimaron las proporciones de tales resultados de la siguiente manera

$$W_{ijp} = 1; \text{ si } \%especificidad_{ijp} = 100\% \wedge \%sensibilidad_{ijp} = 100\%$$

$$W_{ijp} = 1; \text{ en otro caso}$$

Por lo tanto, las proporciones de muestras para los diferentes escenarios en las cuales las pruebas presentaron decisión perfecta, denotadas por $P_{\{decisión\ perfecta_{ip}\}}$, vienen dadas por

$$P_{\{decisión\ perfecta_{ip}\}} = \frac{1}{n^*} \sum_{j=1}^{n^*} W_{ijp}$$

RESULTADOS

Cálculo de los parámetros poblacionales, aplicación de la metodología estándar y curva COR y sus valores de sensibilidad y especificidad en las poblaciones consideradas en el estudio.

Infección por Trypanosoma cruzi

El tamaño de la población de los pacientes sanos fue $N_s = 901$, teniendo como parámetros $\mu_s = 0,1226$ y $\sigma_s = 0,0531$. Por su parte, el tamaño de la población de enfermos fue $N_E = 342$, con media $\mu_E = 0,4093$ y desviación estándar $\sigma_E = 0,2234$. Además, la desviación estándar conjunta entre los valores de absorbancia de las poblaciones sanos y enfermos fue $\sigma_c = 0,1255$. Asimismo, en la figura 5, se observa que las distribuciones de frecuencias absolutas presentan una intersección no vacía o región de respuesta común para los grupos de pacientes sanos y enfermos, lo cual indica que la selección del umbral de decisión debe ser tal que minimice el error cometido en la clasificación de los pacientes sanos y enfermos. En este sentido, se encontró que un total de 589 datos (589/1243; 47,39%) pertenecen a esta intersección.

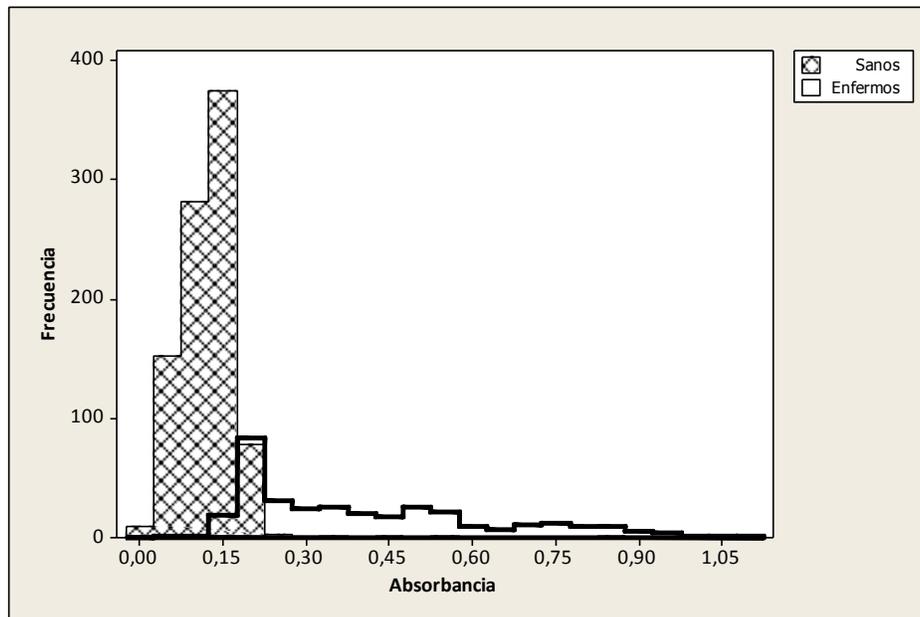


Figura 5. Población de datos constituidos por valores de absorbancia asociados a la concentración de anticuerpos tipo IgG contra *Trypanosoma cruzi*, provenientes de pacientes sanos y enfermos.

La tabla 3 muestra los umbrales de decisión obtenidos al aplicar las metodologías en estudio.

Tabla 3. Aplicación de la metodología estándar y curva COR a una población de datos correspondiente a la concentración de anticuerpos tipo IgG contra *Trypanosoma cruzi*.

Metodología	Umbral de decisión	Sensibilidad	Especificidad
Curva COR			
Índice de Youden	0,186	96,44	90,35
MDC	<u>0,182</u>	<u>94,56</u>	<u>92,11</u>
Estándar			
Est1	0,229	67,84	99,22
Est2	0,282	59,65	99,56
Est3	0,194	87,14	98,22
Est4	0,230	67,25	99,22

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

Se observa que el umbral de decisión para las metodologías de curvas COR fueron similares y menores a todos los umbrales de las metodologías estándar. Asimismo, se observó que los umbrales de las metodologías COR se ubicaron hacia el centro de la región común de ambas distribuciones, mientras que los umbrales de las metodologías clásicas tendieron a ubicarse hacia el extremo derecho de la región común (ver figura 5), favoreciendo de esta manera más la especificidad que la sensibilidad de la prueba. De igual manera, se puede apreciar que los valores de sensibilidad y especificidad fueron más equilibrados para las curvas COR que para las metodologías clásicas. De éstas últimas, la que muestra un mejor comportamiento del par sensibilidad/especificidad es Est3, aunque, igualmente, desbalanceado hacia la especificidad.

Infección por Leishmania sp.

En cuanto al cálculo de los parámetros de los valores de absorbancia asociados a la concentración de anticuerpos tipo IgG contra *Leishmania sp.*, se obtuvo para el grupo de los sanos ($N_s = 500$), un valor de $\mu_s = 0,1141$ y $\sigma_s = 0,053$, mientras que para el grupo de los pacientes enfermos ($N_E = 62$), estos valores fueron $\mu_E = 0,3281$ y $\sigma_E = 0,1486$, teniendo un valor de desviación estándar conjunta de $\sigma_c = 0,070$. En cuanto a las distribuciones de frecuencias absolutas, mostradas en la figura 6, se observa que la región de intersección no vacía entre las poblaciones sanas y enfermas es menor en términos relativos a la obtenida en la infección por *Trypanosoma cruzi*, teniendo que el 6,95% de los datos (39/561) se ubicaron en esta región común.

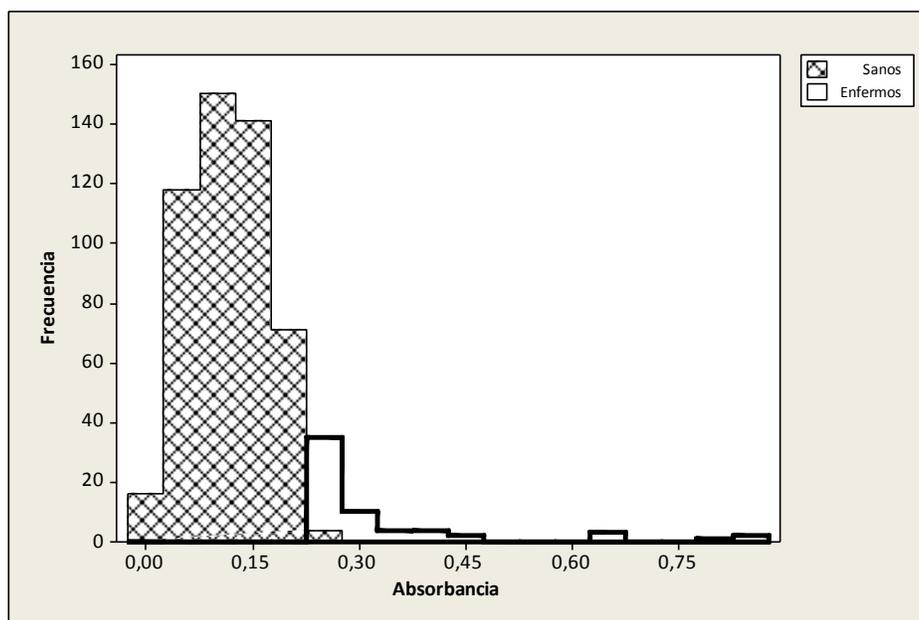


Figura 6. Población de datos constituidos por valores de absorbancia asociados a la concentración de anticuerpos tipo IgG contra *Leishmania sp.*, provenientes de pacientes sanos y enfermos.

La tabla 4, muestra los valores de los umbrales de decisión obtenidos con las metodologías aplicadas, así como sus resultados de sensibilidad y especificidad.

Tabla 4. Aplicación de la metodología estándar y curva COR a una población de datos correspondiente a la concentración de anticuerpos tipo IgG contra *Leishmania sp.*

Metodología	Umbral de decisión	Sensibilidad	Especificidad
Curva COR			
Índice de Youden	<u>0,228</u>	<u>99,8</u>	<u>100</u>
MDC	<u>0,228</u>	<u>99,8</u>	<u>100</u>
Estándar			
Est1	0,220	100	98,8
Est2	0,273	42,62	100
Est3	0,205	100	96,6
Est4	0,251	70,49	100

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

La tabla 4 muestra que los umbrales de decisión obtenidos con los estimadores de la curva COR así como los de las metodologías Est1 y Est3 presentaron valores similares, ubicándose hacia el centro de la región de no discriminación y arrojando un buen equilibrio entre sensibilidad y especificidad. Esta buena discriminación entre las poblaciones sanas y enfermas está influenciada, principalmente, por la baja cantidad de observaciones superpuestas en la región común. Por su parte, Est2 y Est4 arrojaron valores de puntos de corte elevados ubicados hacia la zona superior derecha de la región de no discriminación, favoreciendo la especificidad pero afectando la sensibilidad de la técnica.

Infeción por *Toxoplasma sp.*

En lo que respecta a los valores parametrales relacionados con la concentración de anticuerpos tipo IgG contra *Toxoplasma sp.*, se obtuvo para el grupo de pacientes sanos ($N_s = 36$), un valor de $\mu_s = 0,1071$ y $\sigma_s = 0,0276$, mientras que para el grupo de los pacientes enfermos ($N_E = 26$), estos valores fueron $\mu_E = 0,2128$ y $\sigma_E = 0,0705$, teniendo un valor de desviación estándar conjunta de $\sigma_c = 0,0501$. La figura 7 muestra las distribuciones de las frecuencias absolutas tanto para los pacientes sanos y enfermos, notándose una región mayor de no discriminación en relación con las infecciones por *Trypanosoma cruzi* y *Leishmania sp.* Efectivamente, 38 de las 62 observaciones en total (61,29%) se encuentran en dicha intersección.

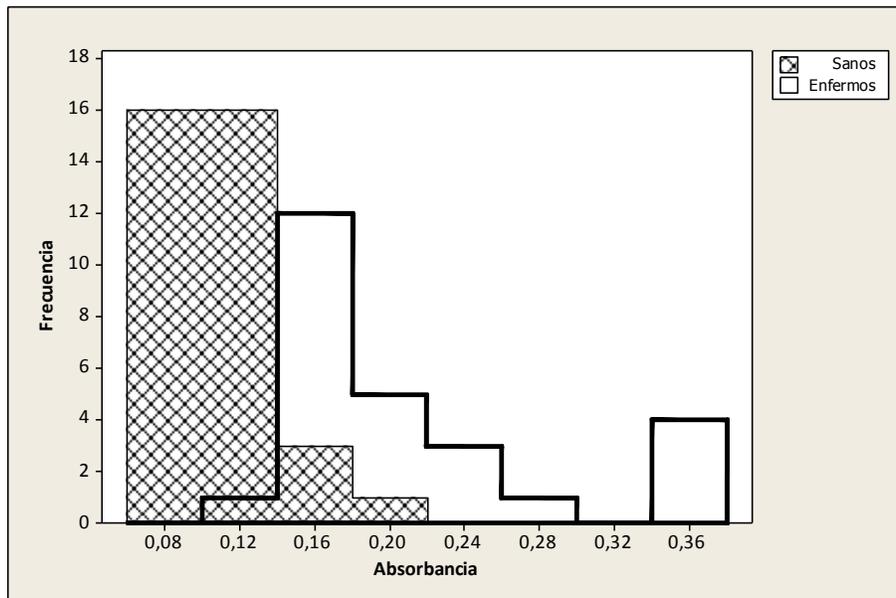


Figura 7. Población de datos constituidos por valores de absorbancia asociados a la concentración de anticuerpos tipo IgG contra *Toxoplasma sp.*, provenientes de pacientes sanos y enfermos.

Los valores de los umbrales de decisión obtenidos con las metodologías en estudio, así como sus respectivos valores de sensibilidad y especificidad, son mostrados en la tabla 5.

Tabla 5. Aplicación de la metodología estándar y curva COR a una población de datos correspondiente a la concentración de anticuerpos tipo IgG contra *Toxoplasma sp.*

Metodología	Umbral de decisión	Sensibilidad	Especificidad
Curva COR			
Índice de Youden	0,134	86,11	100
MDC	0,147	91,67	92,31
Estándar			
Est1	0,162	84,62	97,22
Est2	0,190	46,15	97,22
Est3	0,147	92,31	91,67
Est4	0,168	69,23	97,22

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

Se observa que todos los umbrales estimados se ubican dentro de la región de no discriminación, debido a la amplitud de la misma. Sin embargo, las metodologías que arrojaron puntos de corte similares y con un mejor balance entre sensibilidad y especificidad fueron MDC y Est3, siendo mejores que el índice de Youden y el resto de las pruebas clásicas, las cuales favorecen más la especificidad que la sensibilidad.

Muestras simuladas

Infección por Trypanosoma cruzi

Se generaron un total de $n^* = 1000$ muestras aleatorias para las poblaciones de pacientes sanos P_S y pacientes enfermos P_E , cada una con 30 observaciones. La población de pacientes sanos tuvo una media fijada en $\mu_S = 0,1226$. Por su parte, para los pacientes enfermos se consideraron cinco escenarios: P_{E1} con $\mu_{E1} = 0,18535$; P_{E2} con $\mu_{E2} = 0,2481$; P_{E3} con $\mu_{E3} = 0,3736$; P_{E4} con $\mu_{E4} = 0,4991$; P_{E5} con $\mu_{E5} = 0,6246$.

Para la condición de homocedasticidad, se consideró como varianza común la obtenida de la caracterización del conjunto de datos reales sanos, $\sigma^2 = (0,0531)^2$. En las condiciones

heterocedásticas, se fijó como varianza para las observaciones sanas $\sigma_s^2 = (0,0531)^2$, mientras que para las poblaciones enfermas se consideraron dos valores: varianza obtenida de la caracterización de datos reales enfermos (heterocedásticas 1), $\sigma_{E1}^2 = (0,2234)^2$ y la varianza conjunta para los grupos de pacientes sanos y enfermos resultante de dicha caracterización (heterocedásticas 2), $\sigma_{E2}^2 = (0,1255)^2$.

Infección por Leishmania sp.

Al igual que en el caso de *Trypanosoma cruzi*, fueron generadas un total de $n^* = 1000$ muestras aleatorias para las poblaciones de pacientes sanos P_S y pacientes enfermos P_E , cada una con 30 observaciones. La población de pacientes sanos tuvo una media fijada en $\mu_S = 0,1141$. Por su parte, para los pacientes enfermos se consideraron cinco escenarios: P_{E1} con $\mu_{E1} = 0,1491$; P_{E2} con $\mu_{E2} = 0,1841$; P_{E3} con $\mu_{E3} = 0,2541$; P_{E4} con $\mu_{E4} = 0,3241$; P_{E5} con $\mu_{E5} = 0,3941$.

Para la condición de homocedasticidad, se consideró como varianza común la obtenida de la caracterización del conjunto de datos reales sanos, $\sigma^2 = (0,053)^2$. En las condiciones heterocedásticas, se fijó como varianza para las observaciones sanas $\sigma_s^2 = (0,053)^2$, mientras que para las poblaciones enfermas se consideraron dos valores: varianza obtenida de la caracterización de datos reales enfermos (heterocedásticas 1), $\sigma_{E1}^2 = (0,1486)^2$ y la varianza conjunta para los grupos de pacientes sanos y enfermos resultante de dicha caracterización (heterocedásticas 2), $\sigma_{E2}^2 = (0,07)^2$.

Infección por Toxoplasma sp.

De igual forma a los casos anteriores, fueron generadas un total de $n^* = 1000$ muestras aleatorias para las poblaciones de pacientes sanos P_S y pacientes enfermos P_E , cada una con 30 observaciones. La población de pacientes sanos tuvo una media fijada en $\mu_S = 0,0276$. Por su parte, para los pacientes enfermos se consideraron cinco escenarios: P_{E1} con $\mu_{E1} = 0,13215$; P_{E2} con $\mu_{E2} = 0,1572$; P_{E3} con $\mu_{E3} = 0,2073$; P_{E4} con $\mu_{E4} = 0,2574$; P_{E5} con $\mu_{E5} = 0,3075$.

Para la condición de homocedasticidad, se consideró como varianza común la obtenida de la caracterización del conjunto de datos reales sanos, $\sigma^2 = (0,0276)^2$. En las condiciones heterocedásticas, se fijó como varianza para las observaciones sanas $\sigma_s^2 = (0,0276)^2$, mientras que para las poblaciones enfermas se consideraron dos valores: varianza obtenida de la caracterización de datos reales enfermos (heterocedásticas 1), $\sigma_{E1}^2 =$

$(0,0705)^2$ y la varianza conjunta para los grupos de pacientes sanos y enfermos resultante de dicha caracterización (heterocedásticas 2), $\sigma_{E2}^2 = (0,0501)^2$.

Establecimiento de los umbrales de decisión para las muestras simuladas

Al aplicar las metodologías clásica y curva COR a las muestras aleatorias correspondientes a cada una de las enfermedades evaluadas, se obtuvieron los respectivos valores de umbrales de decisión, cuyos promedios son presentados en las tablas 6, 7 y 8.

Tabla 6. Promedio de los umbrales de decisión obtenidos en las muestras simuladas correspondientes a la concentración de anticuerpos IgG contra *Trypanosoma cruzi*.

Condición de igualdad de varianzas	Metodología	Escenario				
		I	II	III	IV	V
Homocedásticas	Índice de Youden	0,1492	0,1804	0,2238	0,23	0,2301
	MDC	0,1526	0,1819	0,2238	0,23	0,2301
	Est1	0,2273				
	Est2	0,2796				
	Est3	0,1953				
	Est4	0,2316				
Heterocedásticas 1	Índice de Youden	0,1994	0,2006	0,2057	0,2143	0,2223
	MDC	0,1719	0,1768	0,1925	0,2093	0,2215
	Est1	0,2273				
	Est2	0,2796				
	Est3	0,1953				
	Est4	0,2316				
Heterocedásticas 2	Índice de Youden	0,1822	0,1865	0,2059	0,2226	0,2293
	MDC	0,1608	0,1739	0,2022	0,2227	0,2293
	Est1	0,2273				
	Est2	0,2796				
	Est3	0,1953				
	Est4	0,2316				

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

Tabla 7. Promedio de los umbrales de decisión obtenidos en las muestras simuladas correspondientes a la concentración de anticuerpos IgG contra *Leishmania sp.*

Condición de igualdad de varianzas	Metodología	Escenario				
		I	II	III	IV	V
Homocedásticas	Índice de Youden	0,1271	0,1459	0,1790	0,2051	0,2192
	MDC	0,1305	0,1479	0,1803	0,2054	0,2192
	Est1	0,2182				
	Est2	0,2702				
	Est3	0,1866				
	Est4	0,2228				
Heterocedásticas 1	Índice de Youden	0,1794	0,1791	0,1822	0,1909	0,2002
	MDC	0,1503	0,1552	0,1667	0,1809	0,1953
	Est1	0,2182				
	Est2	0,2702				
	Est3	0,1866				
	Est4	0,2228				
Heterocedásticas 2	Índice de Youden	0,1441	0,1532	0,1755	0,1983	0,2142
	MDC	0,1345	0,1488	0,1746	0,1987	0,2143
	Est1	0,2182				
	Est2	0,2702				
	Est3	0,1866				
	Est4	0,2228				

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

Tabla 8. Promedio de los umbrales de decisión obtenidos en las muestras simuladas correspondientes a la concentración de anticuerpos IgG contra *Toxoplasma sp.*

Condición de igualdad de varianzas	Metodología	Escenario				
		I	II	III	IV	V
Homocedásticas	Índice de Youden	0,1171	0,1300	0,1518	0,1621	0,1630
	MDC	0,1189	0,1312	0,1523	0,1621	0,1630
	Est1	0,1619				
	Est2	0,1893				
	Est3	0,1449				
	Est4	0,1637				
Heterocedásticas 1	Índice de Youden	0,1394	0,1401	0,1456	0,1529	0,1585
	MDC	0,1261	0,1305	0,1408	0,1515	0,1584
	Est1	0,1619				
	Est2	0,1893				
	Est3	0,1449				
	Est4	0,1637				
Heterocedásticas 2	Índice de Youden	0,1327	0,1351	0,1459	0,1561	0,1620
	MDC	0,1227	0,1295	0,1442	0,1560	0,1620
	Est1	0,1619				
	Est2	0,1893				
	Est3	0,1449				
	Est4	0,1637				

MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4.

Se puede observar que los valores medios obtenidos con para cada una de las pruebas clásicas representan un valor fijo para las tres condiciones de igualdad de varianzas, pues para su cálculo se utiliza sólo la información de la población de pacientes sanos. Además, es posible notar que las metodologías Est2 y Est4, generan los umbrales de decisión más elevados en comparación al resto de las metodologías. De hecho, ni en la condición más extrema para la curva COR en donde la media de los valores enfermos estuviese alejada a cuatro desviaciones estándar del promedio de observaciones sanas (escenario 5), se obtuvieron puntos de corte tan elevados. Por su parte, los umbrales promedios obtenidos con Est3 presentan una diferencia mínima con los puntos de corte de la curva COR en el tercer escenario.

Los dos estimadores que emplea la Curva COR (índice de Youden y MDC) generan puntos de corte con poca diferencia entre sí, siendo esta distancia reducida a medida que el valor medio del grupo de enfermos se aleja progresivamente de la media de los sanos. No obstante, esta diferencia es mayor cuando las poblaciones sanas y enfermas son heterocedásticas. Además, se muestra que los umbrales de decisión aumentan conforme se incrementa el valor de la media del grupo de observaciones enfermas. Esto debido a que se utiliza para su estimación la información procedente de ambas poblaciones, con el objetivo de minimizar el error cometido al clasificar una observación que se encuentre en la zona de no discriminación.

Sensibilidad y especificidad de la metodología estándar y curva COR en las muestras simuladas

Infección por Trypanosoma cruzi

La figura 8 muestra la comparación de los estimadores del valor de sensibilidad y especificidad de la metodología clásica y curva COR en el conjunto de datos simulados correspondiente al diagnóstico de *Trypanosoma cruzi* bajo condiciones de homocedasticidad. En la misma se observa que los valores promedio de especificidad son mayores empleando la metodología Est2, seguidos de los obtenidos con las pruebas Est4 y Est1. Dentro de las clásicas, la metodología que mostró promedios menores de especificidad fue Est3, sin embargo, todos los estimadores de la metodología estándar arrojaron valores de especificidad por encima del 90%.

En lo que corresponde a la curva COR, MDC e índice de Youden mostraron valores de especificidad similares, teniendo un mínimo alrededor del 75% y cuyo valor aumenta progresivamente conforme el valor promedio del grupo de pacientes enfermos se aleja del valor medio de los pacientes no infectados. De hecho, se llega a obtener valores por encima del 90% a partir del escenario 2, en donde la media de los enfermos se aleja a una desviación de la media de los pacientes sin infección y llegando a alcanzar el 100% de especificidad en el escenario 4.

Por su parte, los valores de sensibilidad muestran que en todos los escenarios, los mismos son superiores al aplicar los índices de Youden y MDC, los cuales oscilan entre el 75 y el 100%. Para el caso de las metodologías estándar, el índice que mostró un mejor comportamiento fue Est3, con un 42% de sensibilidad en el escenario 1, el cual se incrementa progresivamente y coincide con los estimadores de la curva COR a partir del escenario 3. De forma contraria, el que muestra un menor valor medio de sensibilidad es Est2, observándose un porcentaje bajo (5%) en el escenario 1, pero que al igual que el resto de las pruebas clásicas, sus resultados convergen con los de la curva COR a partir del tercer escenario.

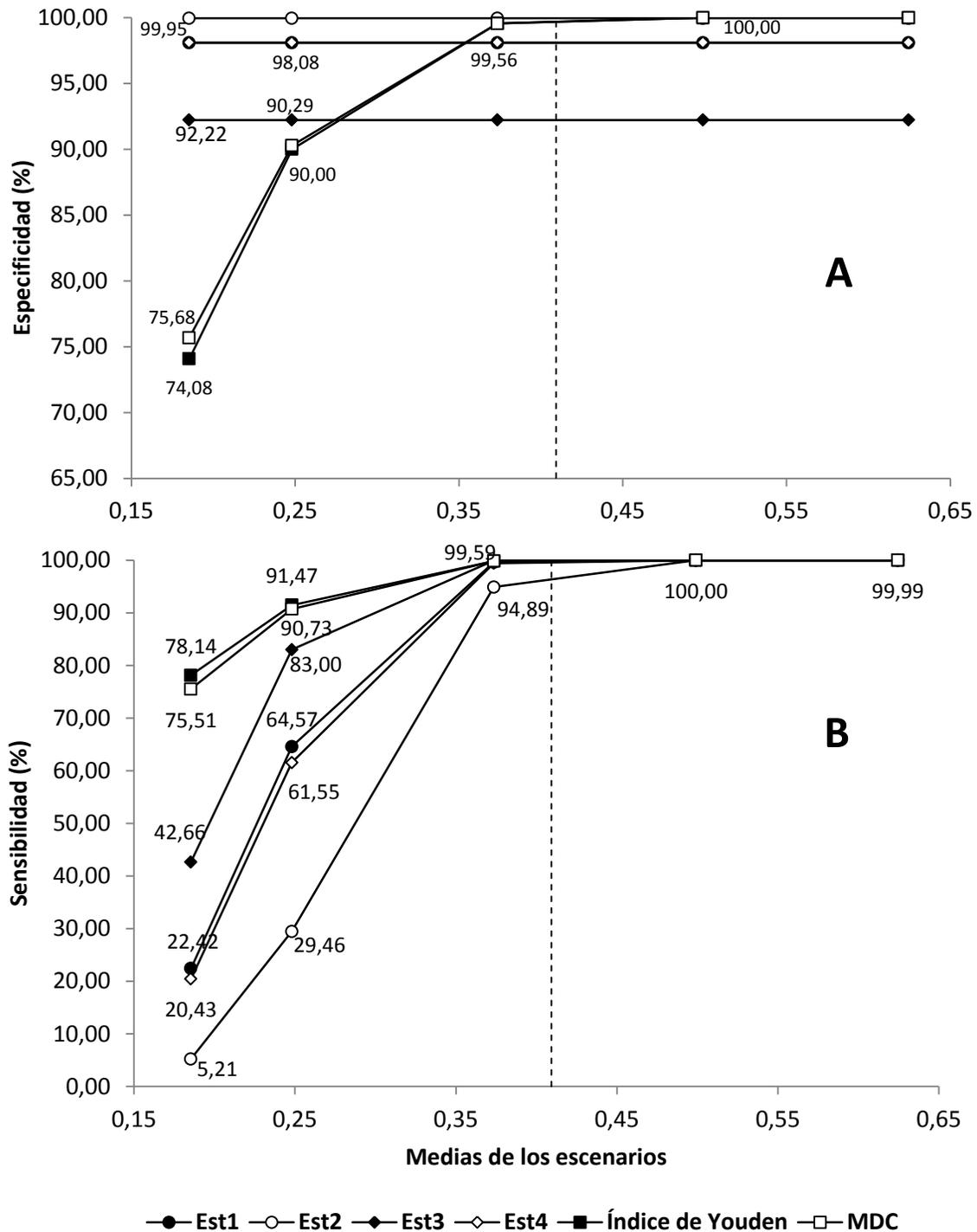


Figura 8. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 9 presenta los valores promedios de sensibilidad y especificidad obtenidos en las muestras simuladas en condiciones heterocedásticas 1. En lo que respecta a los valores promedios de especificidad, ambas técnicas (estándar y curva COR) muestran valores elevados en todos los escenarios, superando el 85%. La técnica que muestra el mayor valor promedio de especificidad es Est2, cuyo valor alcanza el 99,95%, seguida de Est4 y Est2, con cifras alrededor del 98%. En cuanto a los estimadores de la curva COR, la condición de heterocedasticidad afecta más a MDC que al índice de Youden, pues éste último muestra un mejor comportamiento en los valores de especificidad, pues inicia con 95% en el escenario 1 y llega a 99% en el escenario 5; en cambio, MDC en los dos primeros escenarios no alcanza el 90% de especificidad y obtiene valores elevados como el índice de Youden sólo el último escenario.

En relación a los valores de sensibilidad, la MDC es el estimador de umbrales de decisión que refleja un mejor comportamiento a medida que se incrementa el valor medio de las observaciones enfermas, teniendo un mínimo de 55% en el escenario 1 y llegando al 97% en el escenario 5. Seguido se encuentra el índice de Youden con un comportamiento similar a MDC. De las técnicas de la metodología estándar, Est3 es la que aporta mejores valores de sensibilidad, cuyos promedios alcanzan valores como los del índice de Youden. Por otro lado, la metodología estándar que mostró los valores de sensibilidad más bajos fue Est2, llegando a alcanzar cifras por encima del 60% sólo a partir del escenario 4.

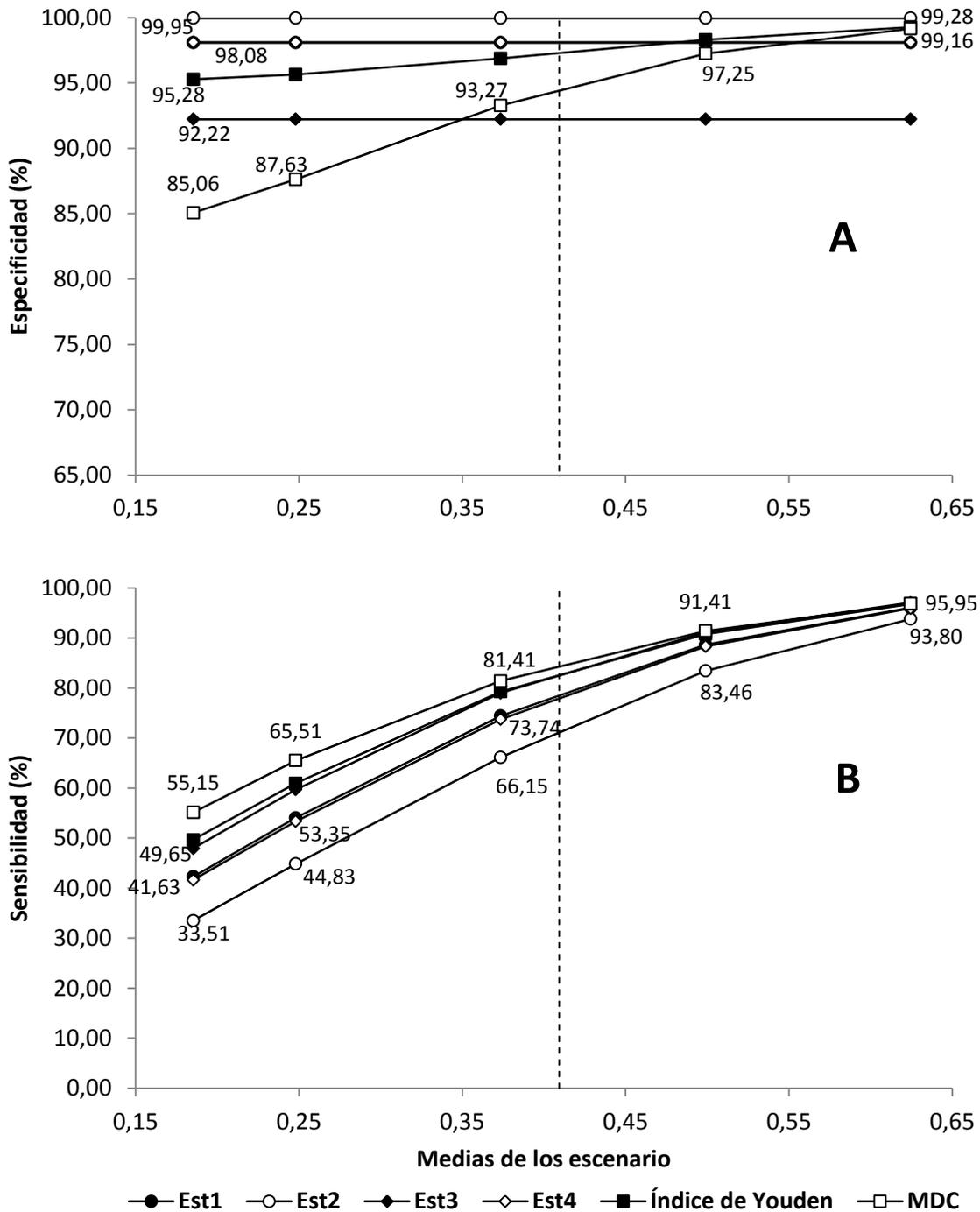


Figura 9. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en la condición heterocedásticas 1, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 10 presenta los valores promedios de sensibilidad y especificidad en la condición heterocedásticas 2. Se observa que los valores promedio de especificidad son mayores empleando Est2 (99,95%) y Est1 y Est4 (98%). Aplicar la prueba Est3 provoca una disminución en la especificidad, cuyo promedio alcanza un 92%. Por su parte, de los estimadores de la curva COR el índice de Youden, al igual que en la condición heterocedásticas 1, presentó un mejor comportamiento que MDC, iniciando en el primer escenario con una especificidad promedio del 90% e incrementándose progresivamente hasta alcanzar el 99% en el escenario 4. MDC presentó valores promedios más bajos que el índice de Youden, con una diferencia máxima del 10% en el escenario 1; distancia ésta que se reduce a medida que se incrementa progresivamente el valor medio de las observaciones enfermas.

Por su parte, los mayores valores de sensibilidad son obtenidos al emplear MDC como estimador de los umbrales de decisión, arrojando valores por encima del 75% a partir del escenario 2. Seguido de éste, se encuentran el índice de Youden y Est3, alcanzando ambos los valores reportados por MDC a partir del escenario 3. Además, la metodología que arrojó los valores promedios más bajos de sensibilidad fue la técnica clásica Est2, cuyos niveles no alcanzaron por lo menos el 50% en el segundo escenario.

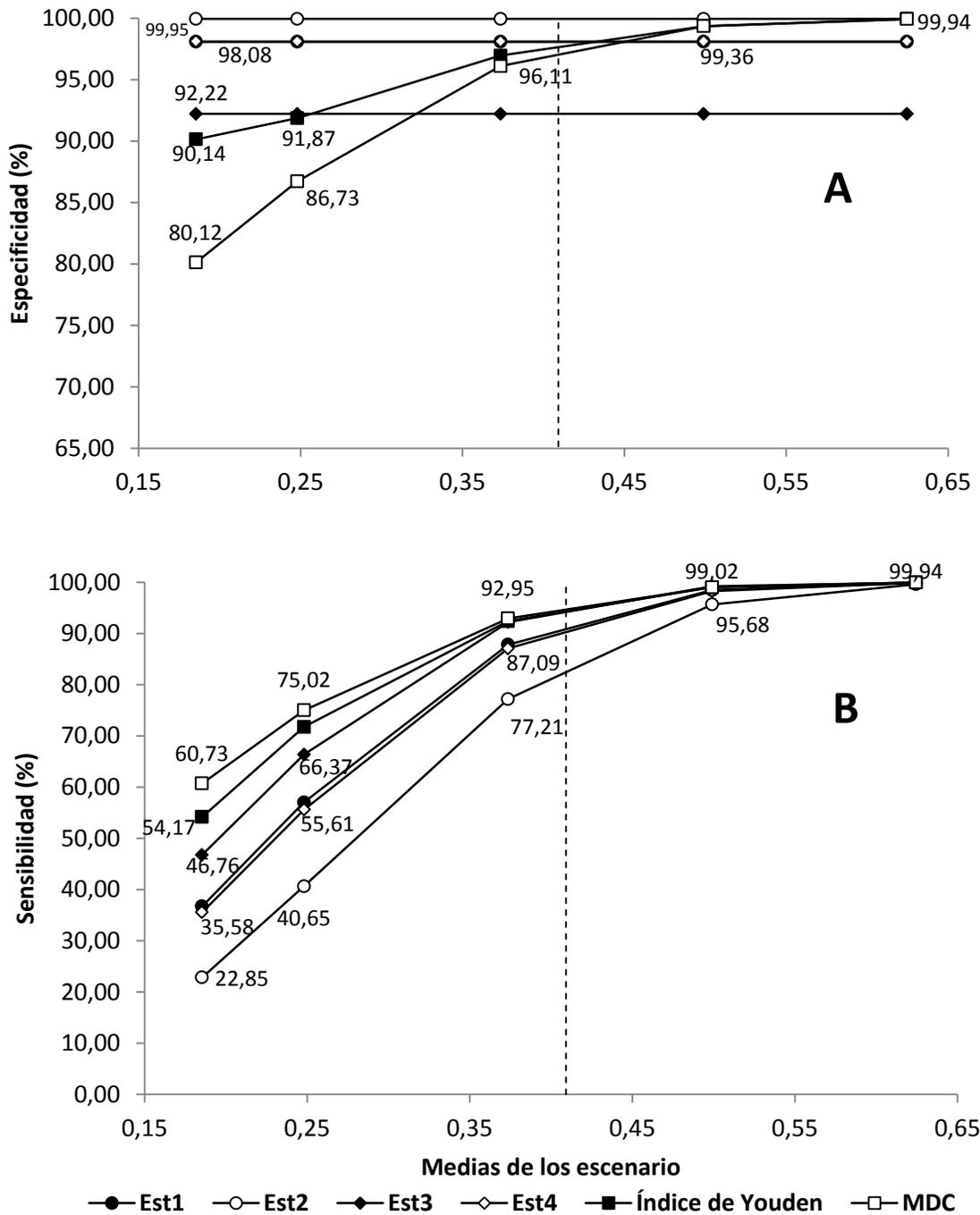


Figura 10. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en la condición heterocedásticas 2, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infección por Leishmania sp.

La figura 11 muestra la comparación de los estimadores del valor de sensibilidad y especificidad de la metodología clásica y curva COR en las muestras simuladas correspondiente al diagnóstico de *Leishmania sp.*, bajo condiciones de homocedasticidad.

Se observa que los valores promedio de especificidad son mayores empleando las técnicas Est2 y Est4. De las metodologías que emplean los datos sanos equilibrados, Est3 genera una disminución de los niveles promedio de especificidad de 6 puntos y ubicándose en 92%. En lo que corresponde a la curva COR, el índice de Youden y MDC mostraron valores de especificidad similares, teniendo un mínimo alrededor del 66% y cuyo valor aumenta progresivamente conforme se incrementa el valor promedio del grupo de pacientes enfermos, alcanzando valores elevados como los de las metodologías clásicas en el escenario 3.

Por su parte, los valores de sensibilidad muestran que en todos los escenarios, los mismos son superiores al aplicar los índices de Youden y MDC, los cuales oscilan entre el 67 y 99%. Para el caso de la metodología estándar, el índice que mostró un mejor comportamiento fue Est3, con un 24,60% de sensibilidad en el escenario 1, la cual genera los valores elevados de la curva COR a partir del escenario 4. De forma contraria, el que muestra un menor valor medio de sensibilidad es Est2, observándose un porcentaje muy bajo (2%) en el escenario 1 y llegando a superar el nivel del 50% a partir del escenario 4.

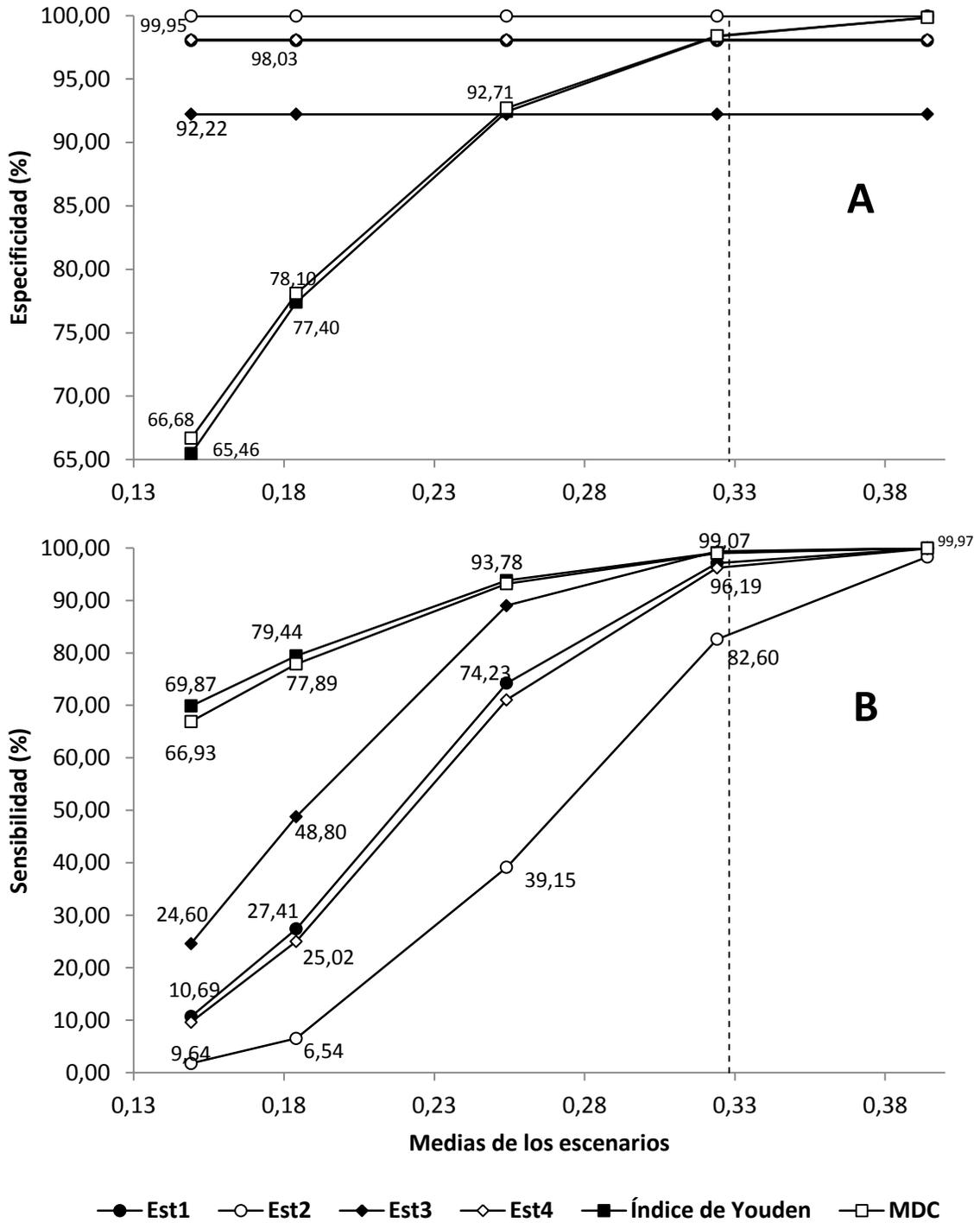


Figura 11. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

En lo que corresponde a la primera de las condiciones heterocedásticas, la figura 12 presenta los valores promedios de sensibilidad y especificidad obtenidos. Se observa que los valores promedios de especificidad de las técnicas clásicas fueron elevados, superando el 90%. De éstas, la que muestra el mayor valor promedio de especificidad es Est2, cuyo valor alcanza el 99,95%, seguida de Est4 y Est1 con valores alrededor del 98%. De la metodología clásica, la que menor posee nivel de especificidad es Est3 (92%). En cuanto a los estimadores de la curva COR, el índice de Youden se ve poco afectado por la heterocedasticidad, teniendo un buen comportamiento en los valores de especificidad, pues inicia con 92% en el escenario 1 y llega a 94% en el escenario 5. MDC, si bien presentan un aumento progresivo a medida que se incrementan los escenarios, sus valores promedios en todos los casos son menores a los del índice de Youden.

En cuanto a la sensibilidad, MDC es el estimador que refleja un mejor comportamiento a medida que se incrementa el valor medio de las observaciones enfermas; teniendo un mínimo de 53% en el escenario 1 y llegando al 92% en el escenario 5. Seguido se encuentra el índice de Youden y Est3 con valores promedios cercanos a los de MDC a partir del escenario 4. Por otro lado, la metodología estándar que mostró los valores de sensibilidad más bajos fue Est2, con cifras que no superan en ninguno de los casos el 80%.

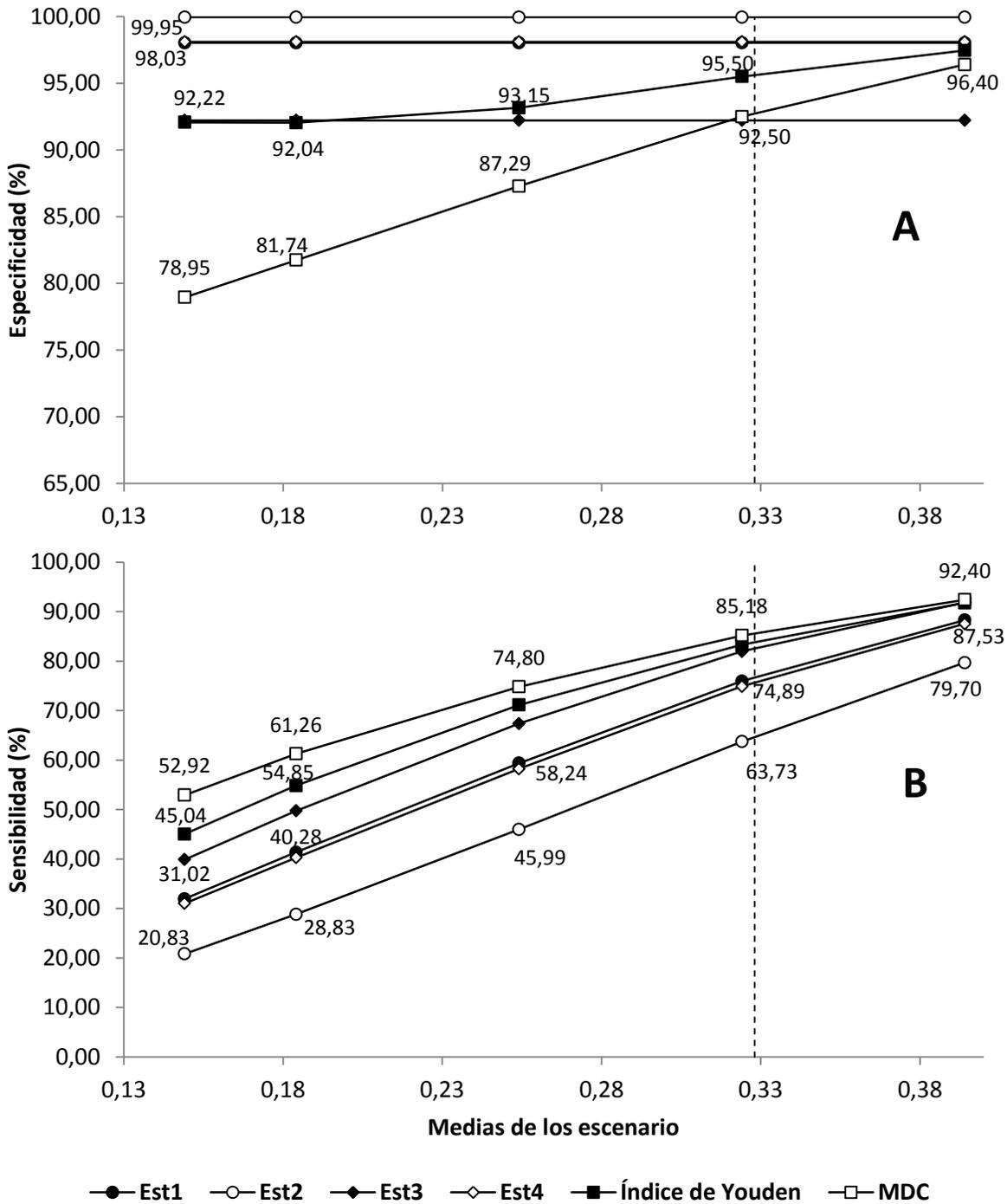


Figura 12. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

En lo que respecta a la segunda de las condiciones de heterocedasticidad, la figura 13 presenta los valores promedios de sensibilidad y especificidad. Se observa que los valores promedio de especificidad son mayores empleando Est2 (99,95%). Resultados elevados también se obtuvieron con Est4 (98%). El menor de los valores promedio de especificidad de las metodologías clásicas fue obtenido con Est3 (91%). Por su parte, de los estimadores de la curva COR, el índice de Youden presentó un mejor comportamiento que MDC pero sólo en los dos primeros escenarios, iniciando en el primer escenario con una especificidad promedio del 78% e incrementándose progresivamente hasta alcanzar, en ambos estimadores, el 99% en el escenario 4.

Por su parte, los mayores valores de sensibilidad promedios son obtenidos al emplear MDC e índice de Youden como estimadores de los umbrales de decisión, arrojando valores por encima del 50% en el primer escenario y logrando cifras del 90% a partir del escenario 3. De los estimadores de la metodología clásica, el de mejor comportamiento fue Est3, obteniendo buenas cifras como las arrojadas por los estimadores de la curva COR en el escenario 4. Del resto, todos presentan bajos promedios de sensibilidad, que se incrementan por encima del 80% a partir del escenario 4.

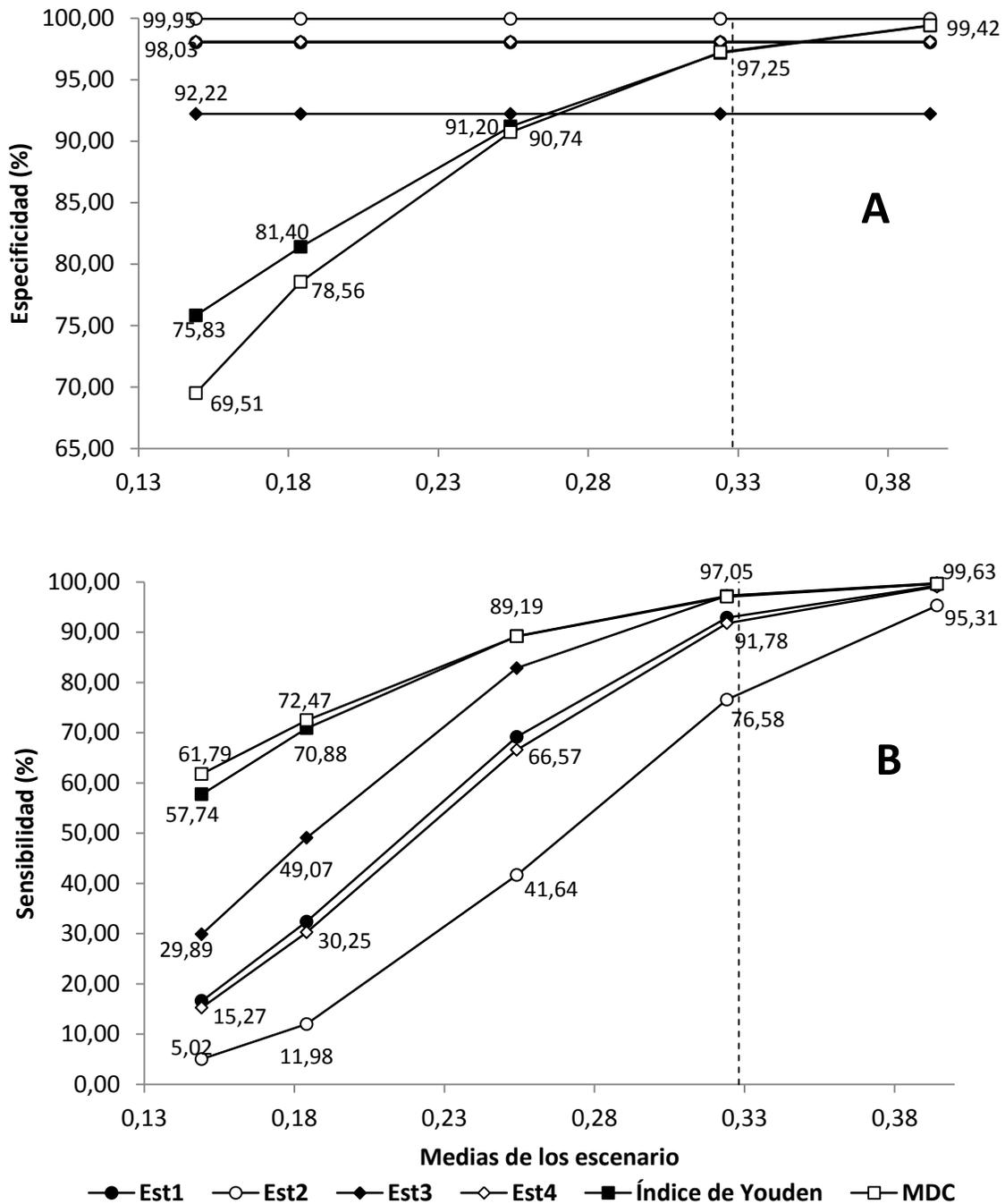


Figura 13. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infección por Toxoplasma sp.

La figura 14 muestra la comparación de los estimadores del valor de sensibilidad y especificidad de la metodología clásica y curva COR en las muestras simuladas correspondiente al diagnóstico de *Toxoplasma sp.*, bajo condiciones de homocedasticidad.

Los valores promedio de especificidad son mayores al emplear Est2 y Est4, cuyos porcentajes disminuyen si se emplea Est3 como estimador (92%). En lo que corresponde a la curva COR, el índice de Youden y MDC mostraron valores de especificidad dinámicos, teniendo un mínimo alrededor del 70% y cuyo valor aumenta progresivamente conforme se incrementa el valor promedio del grupo de pacientes enfermos, llegando a arrojar un 100% de sensibilidad en el escenario 5.

Por su parte, los valores de sensibilidad muestran que en todos los escenarios, los mismos son superiores al aplicar los índices de Youden y MDC, arrojando valores no menos del 70%. Para el caso de las metodologías estándar, el índice que mostró un mejor comportamiento fue Est3, con un 33% de sensibilidad en el escenario 1 y los cuales alcanzan las cifras de la curva COR a partir del escenario 3. De forma contraria, el que muestra un menor valor medio de sensibilidad es Est2, observándose un porcentaje bastante bajo (3%) en el escenario 1 y obteniendo valores del 99% en el escenario 4.

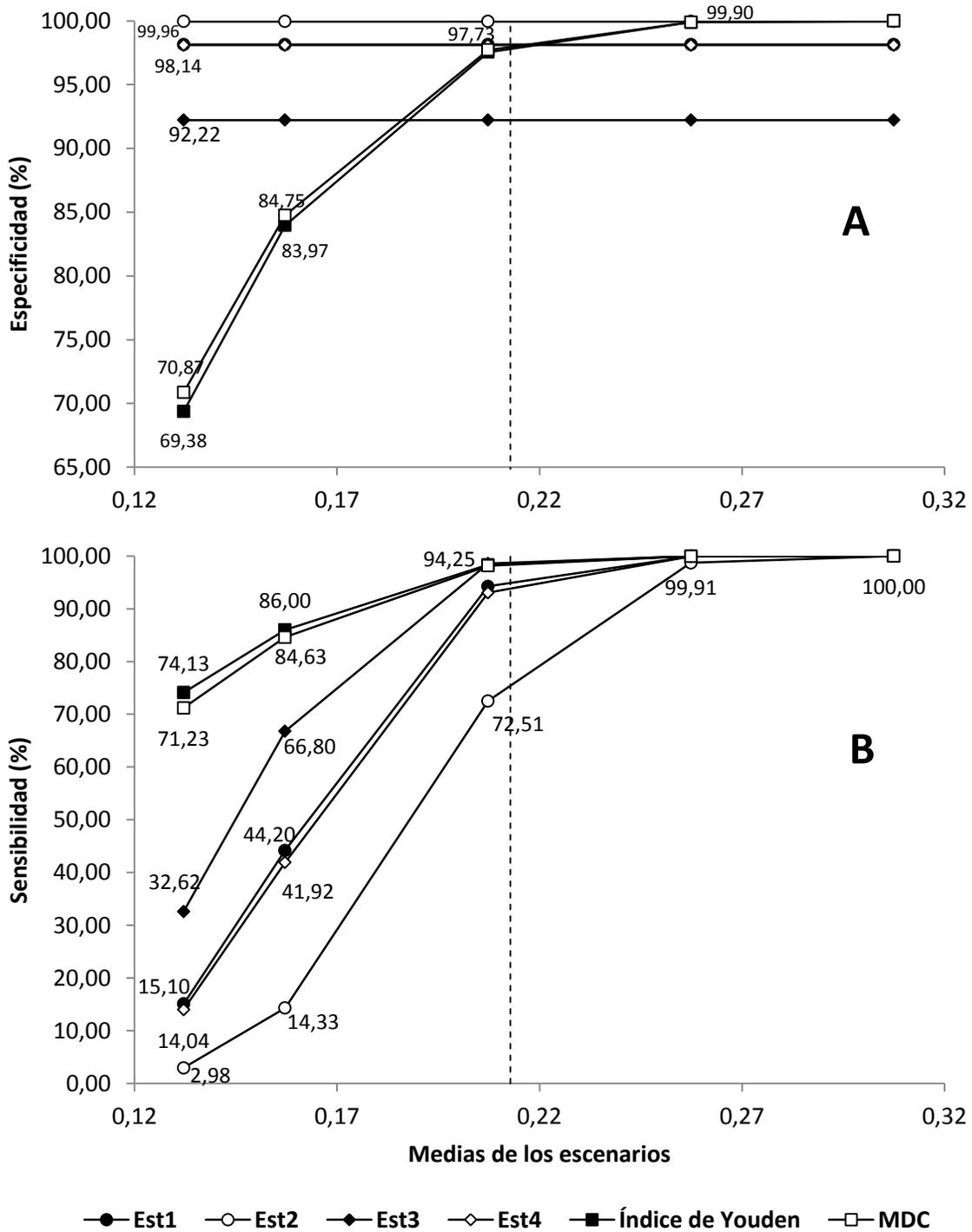


Figura 14. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

En la primera de las condiciones heterocedásticas, la figura 15 presenta los valores promedios de sensibilidad y especificidad.

Los promedios de especificidad de las técnicas clásicas fueron elevados, superando el 90%, teniendo Est2 el mayor valor de todas (99,96%). En cuanto a los estimadores de la curva COR, el índice de Youden es el que muestra un mejor comportamiento de sus promedios de especificidad en los cinco escenarios, iniciando con 91% y alcanzando un máximo de 99%. MDC se ve afecta por la condición de heterocedasticidad y toma valores por encima del 90% sólo a partir del escenario 3.

En relación a los valores de sensibilidad, MDC es el estimador de umbrales de decisión que refleja un mayor promedio en todos los escenarios, aumentando progresivamente a partir del 50%. Seguido se encuentra el índice de Youden y Est3 con valores menores al 50% para el escenario 1 y cifras superiores al 80% a partir del escenario 3. Por otro lado, la metodología estándar que mostró los valores de sensibilidad más bajos fue Est2, cifras que no superan en ningún caso las arrojadas por las demás técnicas.

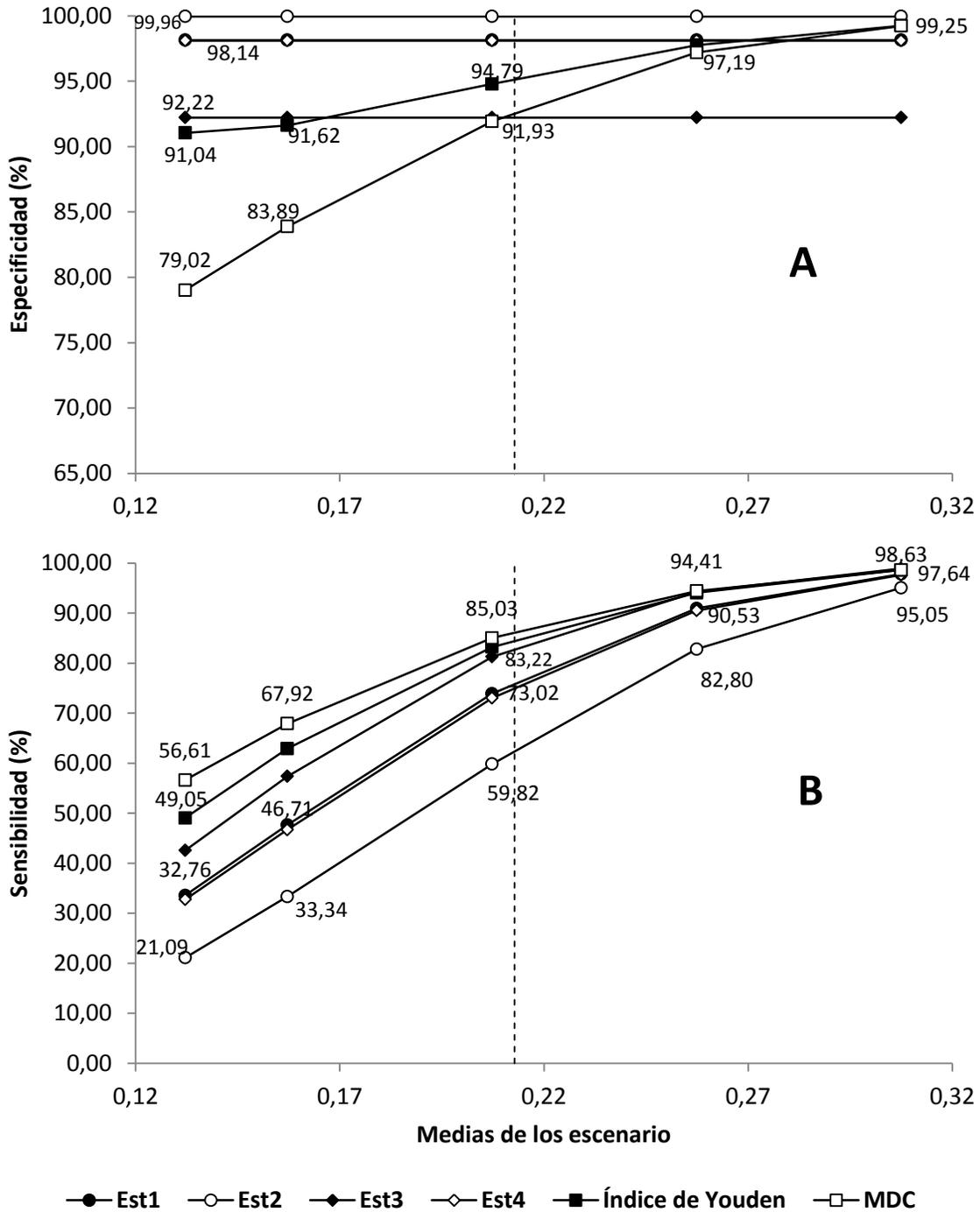


Figura 15. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

En lo que respecta a la segunda de las condiciones de heterocedasticidad, la figura 16 presenta los valores promedios de sensibilidad y especificidad obtenidos en las muestras simuladas.

Se observa que los valores promedio de especificidad son mayores empleando Est2 (99,95%). Dentro de las clásicas, la que mostró el menor de los valores promedios fue Est3, aunque con una buena media del 91%. Por su parte, de los estimadores de la curva COR, el índice de Youden presentó un mejor comportamiento que MDC, sólo en los dos primeros escenarios, iniciando en el primero con una especificidad media del 86% e incrementándose progresivamente hasta alcanzar el 99,85% en el escenario 5.

Por su parte, los mayores valores de sensibilidad promedios son obtenidos al emplear MDC e índice de Youden como estimadores de los umbrales de decisión, arrojando valores por encima del 50% en el primer escenario y logrando cifras del 90% a partir del escenario 3. De los estimadores de la metodología clásica, el de mejor comportamiento fue Est3, teniendo cifras elevadas como la de los estimadores de la curva COR a partir del escenario 3. Del resto, todos presentan bajos promedios de sensibilidad, que se incrementan por encima del 50% a partir del tercer escenario.

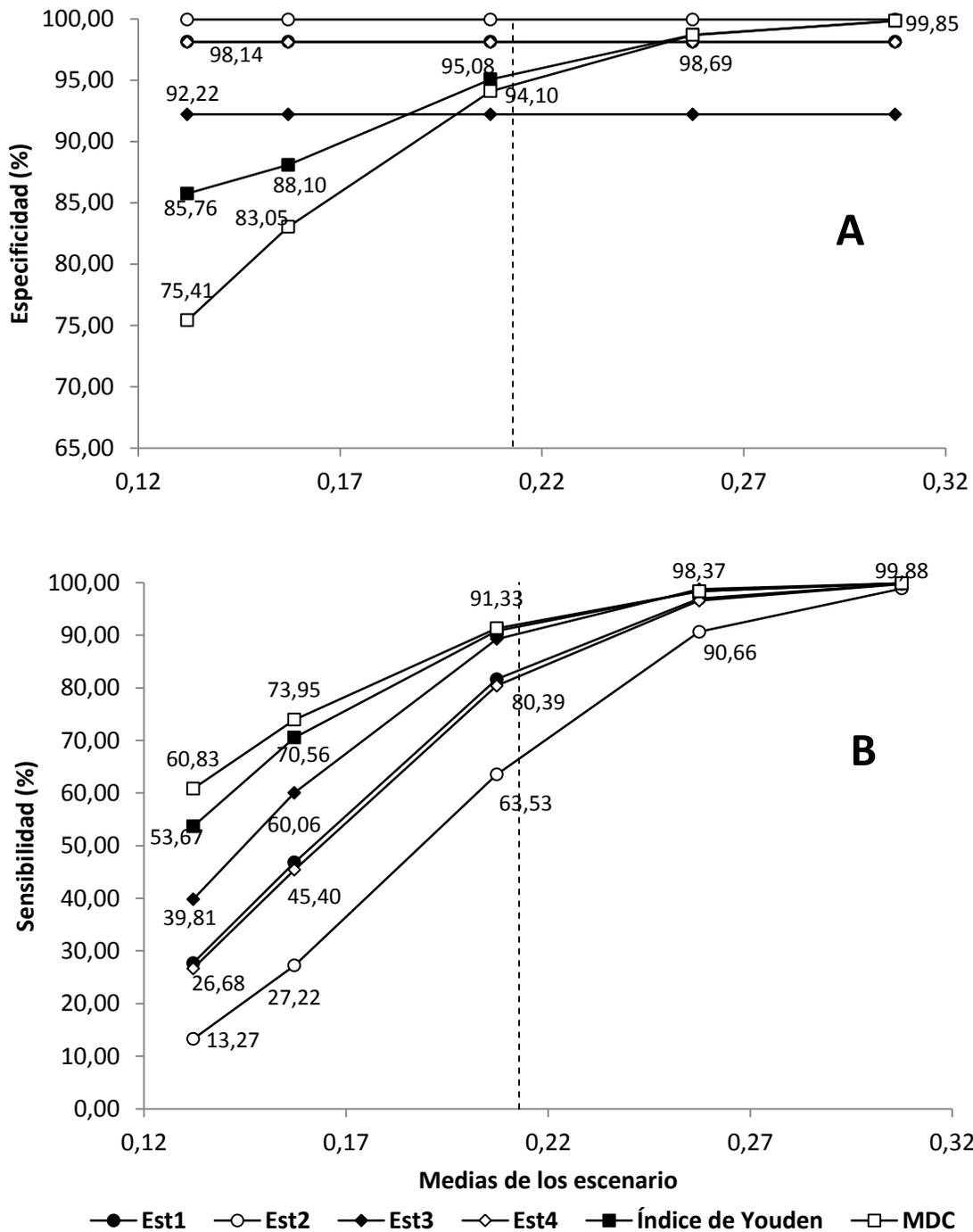


Figura 16. Estimadores del valor de Especificidad (A) y Sensibilidad (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Comparación de la precisión discriminatoria de las metodologías en estudio en las muestras simuladas

- **%Especificidad=100% y %Sensibilidad=100%**

Infección por Trypanosoma cruzi

La figura 17 muestra la estimación de la proporción de casos en donde se obtuvieron niveles de especificidad y sensibilidad igual al 100% en condiciones de homocedasticidad. En el caso de la especificidad, técnica clásica Est2 mostró en casi la totalidad de las muestras simuladas una especificidad perfecta de 100%. Seguido y con gran distancia se encuentran Est4 y Est1, cuyos porcentajes oscilan alrededor del 50%. Est3 mostró una especificidad perfecta tan sólo en el 5% de los casos. Los estimadores de la curva COR presentaron un comportamiento similar entre sí, pasando de una escasa frecuencia de especificidad perfecta en los dos primeros escenarios a un elevado porcentaje a partir del escenario 3 (87%) con un aumento progresivo hasta el 99,90%.

En cuanto a la proporción de sensibilidad igual al 100%, todas las metodologías (estándar y COR) mostraron una baja frecuencia en los dos primeros escenarios; no obstante, el porcentaje mostrado por los indicadores Youden y MDC es superior al resto. Con excepción de la metodología clásica Est2, todas las metodologías experimentaron un incremento notable en el acierto del 100% de sensibilidad a partir del escenario 3, teniendo los mayores niveles de frecuencia en el índice de Youden y Est3 (95%). El punto de convergencia en donde todas las técnicas coinciden, además con un elevado porcentaje de sensibilidad igual al 100%, es del escenario 4 en adelante.

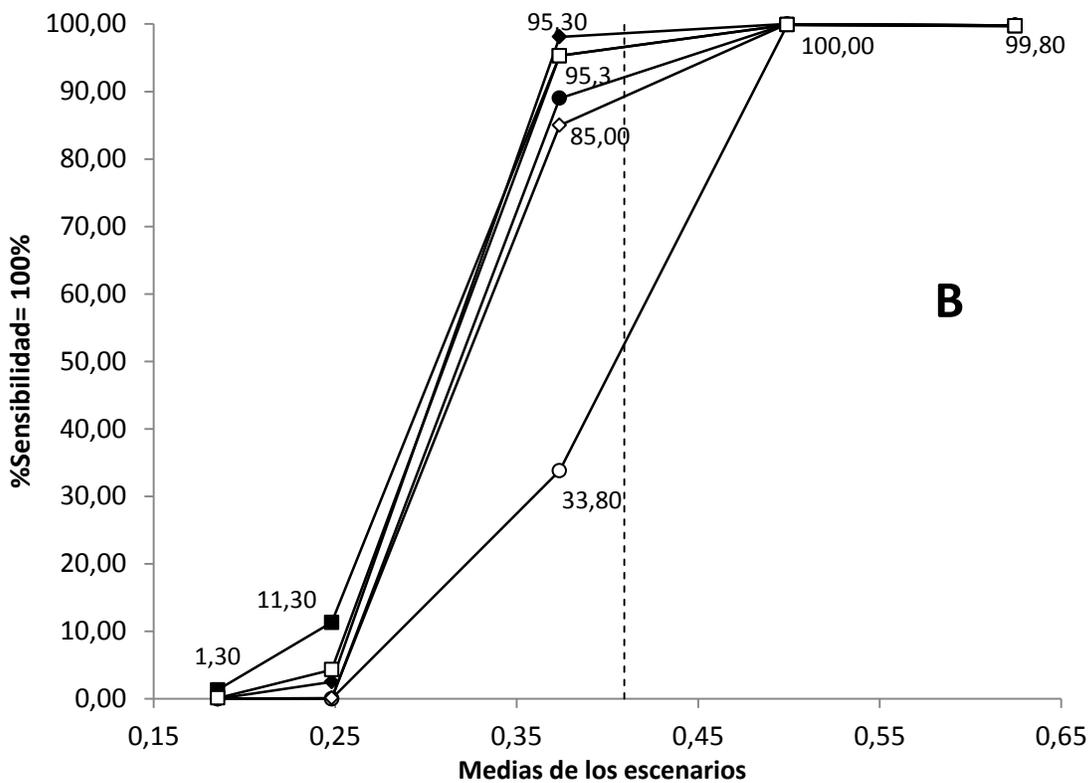
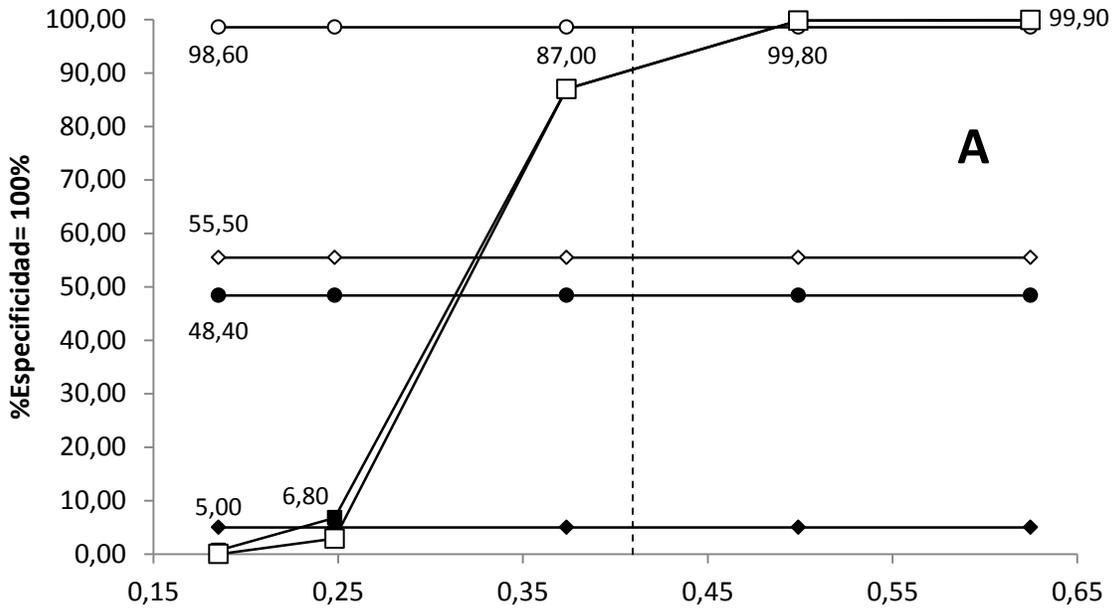


Figura 17. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 18 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en la primera de las condiciones heterocedásticas. En relación a la especificidad igual al 100%, la técnica que tuvo una mayor frecuencia de acierto fue Est2, con un 98,60% de los casos. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, el índice de Youden se ve menos afectado por la heterocedasticidad que MDC, y mostrando un mejor comportamiento, pues el porcentaje de muestras en donde se alcanzó una especificidad perfecta en el primer escenario fue del 33%; 27 veces más que lo observado con MDC. Ambos estimadores tuvieron un incremento progresivo, no obstante el índice de Youden siempre fue mayor, llegando a alcanzar el 80% al escenario 5.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas obtuvieron valores bajos hasta el escenario 4 y llegando a tener un porcentaje no mayor al 50% en el escenario 5. En éste último, la metodología que arrojó el valor más alto en cuanto a sensibilidad perfecta fue la técnica clásica Est3 (39,70%), seguida de los estimadores de la curva COR con un 35%. De igual manera, la metodología que proporcionó el valor más bajo fue la clásica Est2 (13%).

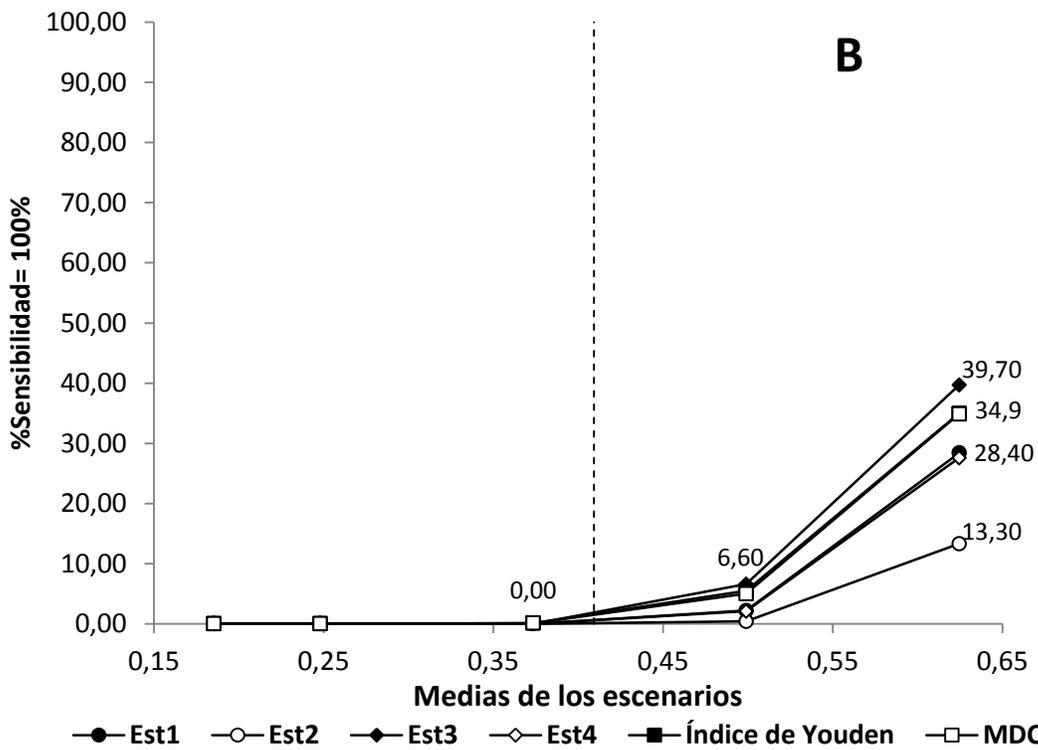
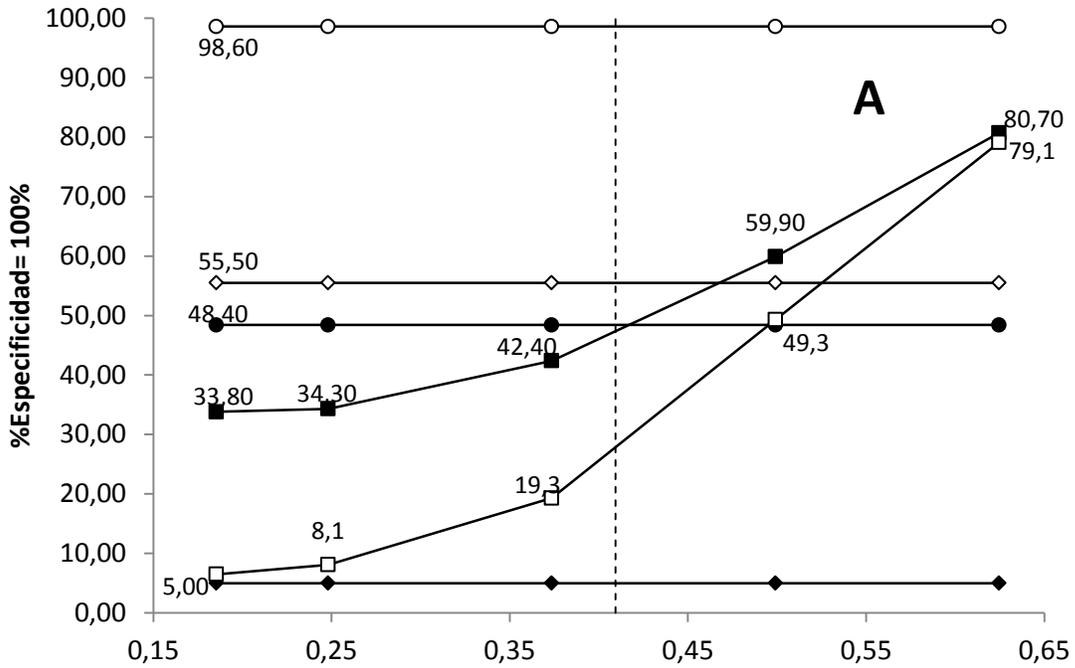


Figura 18. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 19 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en la segunda de las condiciones heterocedásticas. En relación a la especificidad igual al 100%, la técnica que tuvo una mayor frecuencia de acierto fue la clásica Est2, con un 98,60% de los casos. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, el índice de Youden fue el que mostró un mejor comportamiento; sin embargo, las frecuencias de acierto son bajas en los tres primeros escenarios (máximo un 42%), cifras que se elevan del escenario 4 en adelante. Lo mismo ocurrió con la MDC, pero como éste índice se vio afectado por la heterocedasticidad, sus valores fueron más más bajos con un mínimo en el primer escenario del 1%.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas, con excepción de Est2, obtuvieron valores superiores al 10% a partir del escenario 3, siendo muy bajos a esta nivel y llegando a tener un porcentaje elevado en el escenario 5. Los estimadores de la curva COR y la metodología Est3, fueron la que mostraron un mejor comportamiento, teniendo los valores más elevados en ésta última.

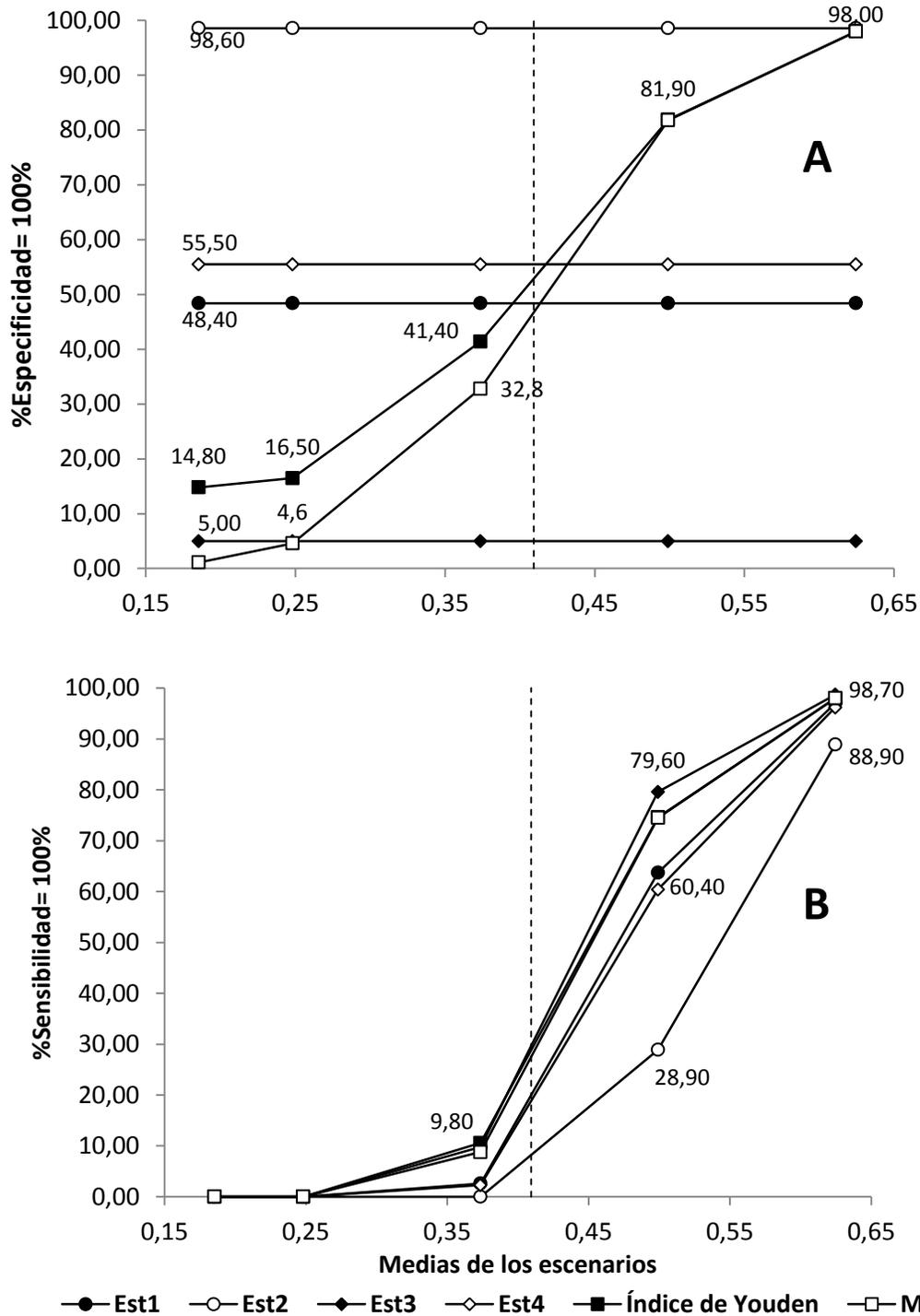


Figura 19. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infección por Leishmania sp.

La figura 20 muestra la estimación de la proporción de casos en donde se obtuvieron niveles de especificidad y sensibilidad igual al 100% en condiciones homocedásticas. En el caso de la especificidad, la metodología clásica Est2 mostró en casi la totalidad de las muestras simuladas una especificidad perfecta de 100%. Seguido y con gran distancia se encuentran Est4 y Est1, cuyos porcentajes oscilan alrededor del 50%. Est3 mostró una especificidad perfecta tan sólo en el 5% de los casos. Los estimadores de la curva COR presentaron un comportamiento similar, con una especificidad perfecta que va del 13% en el escenario 3 al 95% en el escenario 5.

En cuanto a la proporción de sensibilidad igual al 100%, ambas metodologías (estándar y COR) mostraron una baja frecuencia en los dos primeros escenarios, reportándose tan sólo un 2% para los estimadores de la curva COR. Éstos últimos son los que mostraron un mayor porcentaje de casos con sensibilidad perfecta, siendo superados solamente por la técnica clásica Est3 a partir del escenario 4. En el escenario 5, todas las técnicas, con excepción Est2, aportan una sensibilidad perfecta en la casi totalidad de las muestras simuladas.

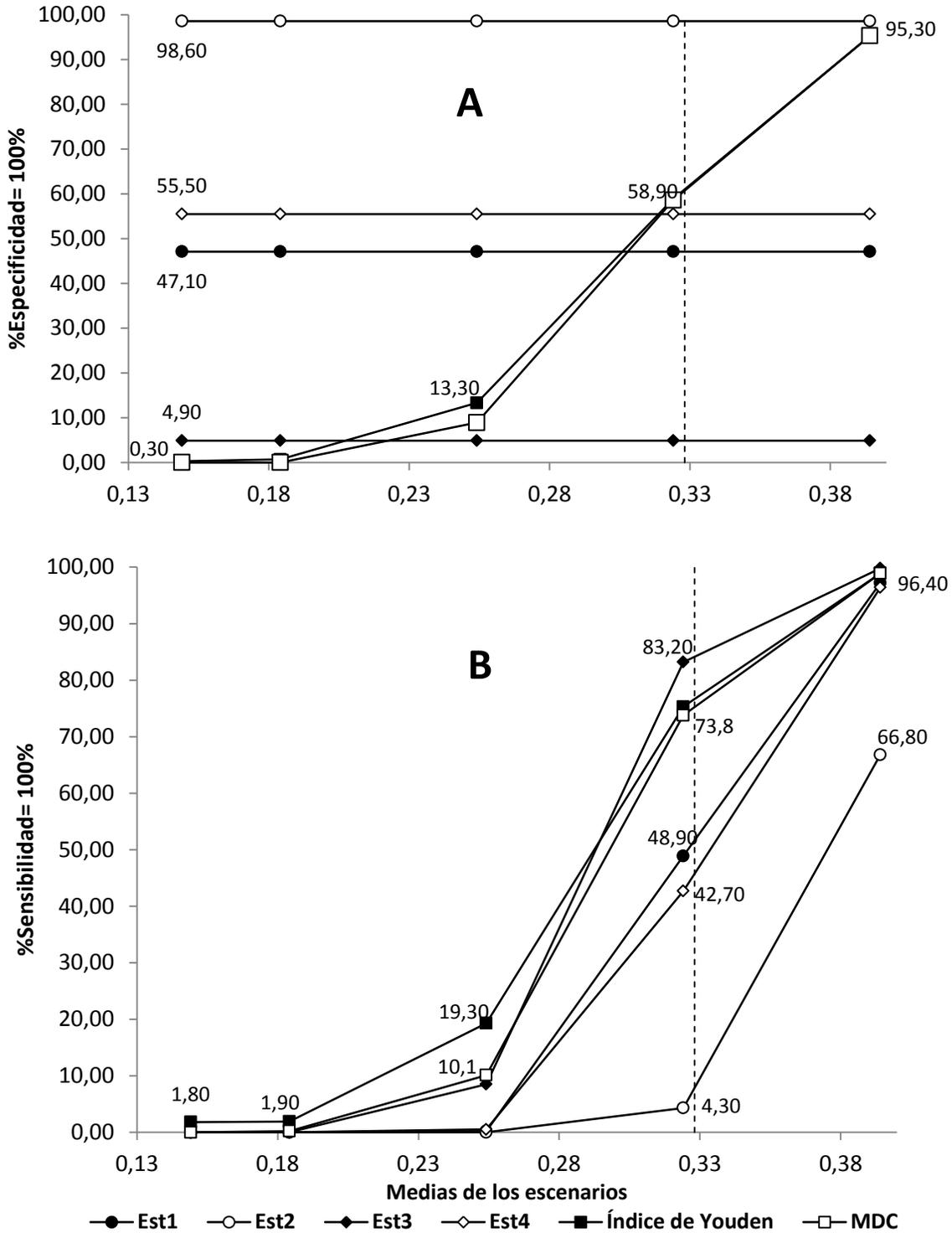


Figura 20. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 21 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en la condición heterocedásticas 1. En relación a la especificidad igual al 100%, la técnica que tuvo una mayor frecuencia de acierto fue la clásica Est2, con un 98,60%. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, si bien el índice de Youden fue el menos afectado por la heterocedasticidad y mostró un mejor comportamiento, en ninguno de los escenarios las frecuencias de los casos con especificidad perfecta alcanzó el 50%. En este mismo orden de ideas, MDC mostró casos de especificidad perfecta por debajo del 10% en los dos primeros escenarios.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas obtuvieron resultados a partir del escenario 4 y, aunque los mejores resultados fueron obtenidos con los estimadores de la curva COR y Est3, la frecuencia de casos en donde se obtuvo una sensibilidad perfecta no superó el 10% en ninguno de los escenarios. De hecho, el resto de las metodologías clásicas reportaron valores de sensibilidad perfecta siempre por debajo del 5%.

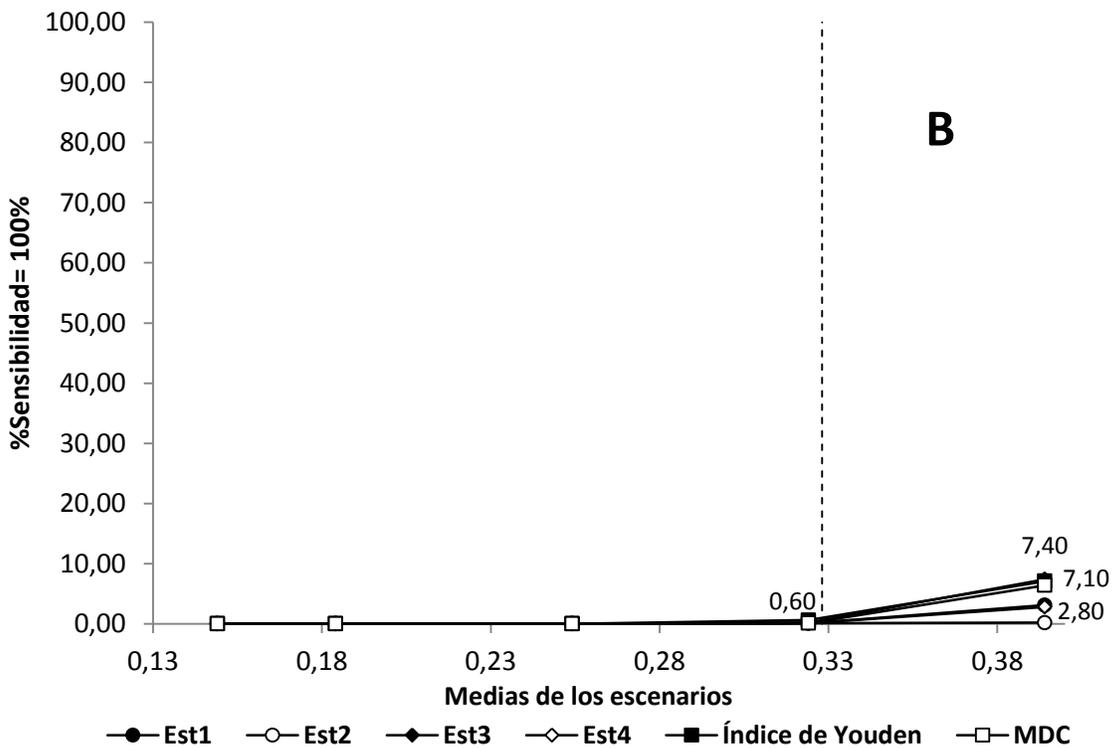
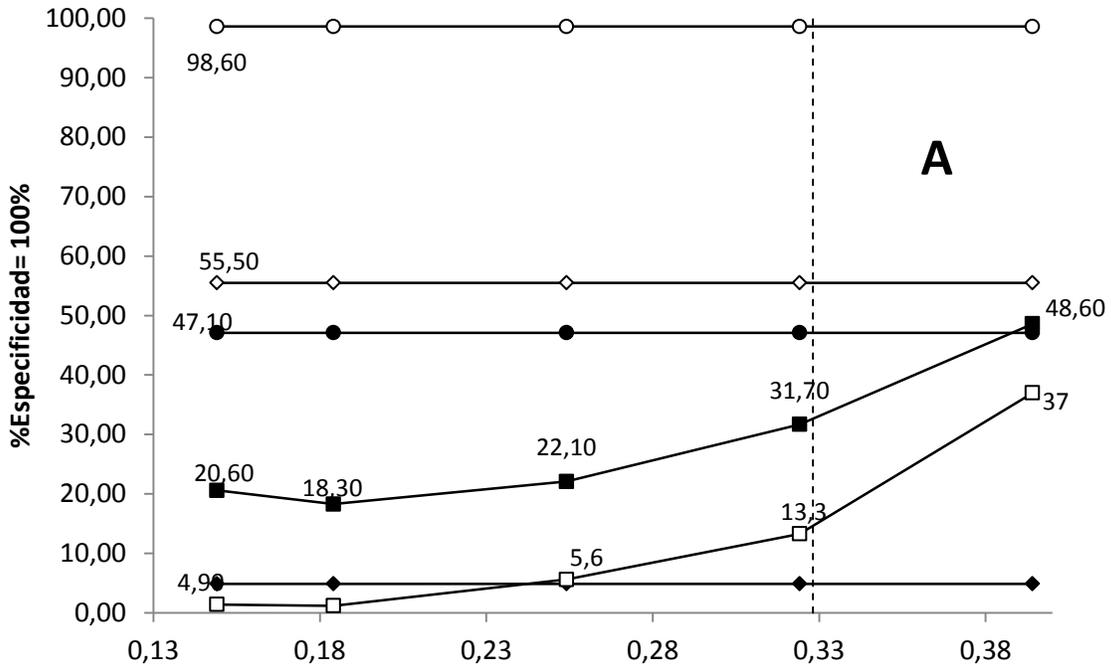


Figura 21. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 22 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en condiciones heterocedásticas 2. En relación a la especificidad igual al 100%, la técnica que tuvo una mayor frecuencia de acierto fue la clásica Est2, con un 98,60% de los casos. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, el índice de Youden fue el que mostró un mejor comportamiento; sin embargo, las frecuencias de acierto son bajas, llegando a incrementarse por encima del 50% sólo a partir del escenario 4.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas, con excepción de la metodología Est2, obtuvieron valores a partir del escenario 3, siendo muy bajos a esta nivel y llegando a tener un porcentaje elevado en el escenario 5. La metodología Est3 fue la que mostró un mejor comportamiento, con un 94% de sensibilidad perfecta en el escenario 4, seguida de los estimadores de la curva COR (89%). El resto de las metodologías clásicas, mostraron valores inferiores al 80%, teniendo el mínimo con la Est2 (32%).

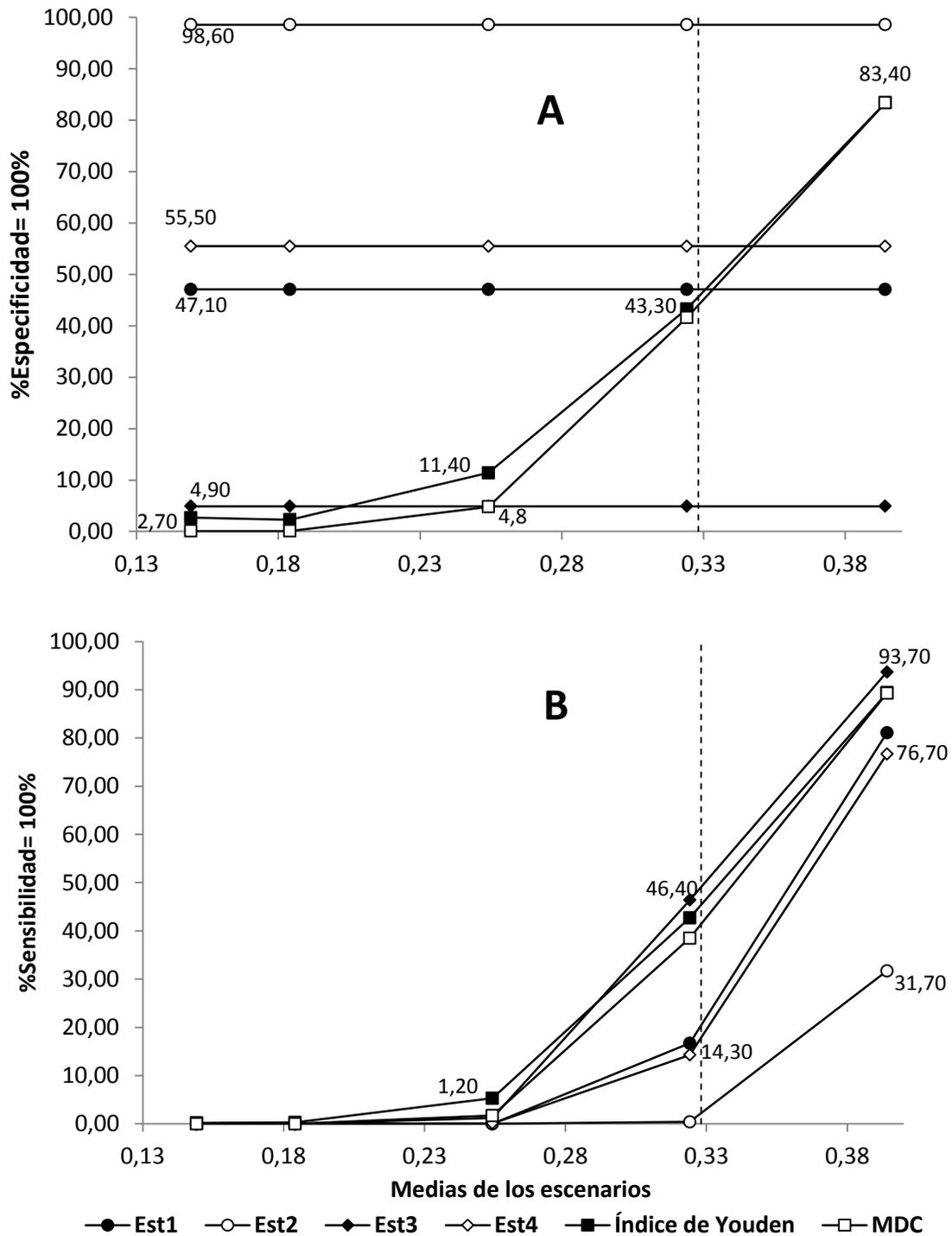


Figura 22. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infección por Toxoplasma sp.

La figura 23 muestra la estimación de la proporción de casos en donde se obtuvieron niveles de especificidad y sensibilidad igual al 100% en condiciones homocedásticas. En el caso de la especificidad, la metodología Est2 mostró en casi la totalidad de las muestras simuladas una especificidad perfecta de 100%. Seguido se encuentran Est4 y Est1 con porcentajes cercanos al 50%, mientras que Est3 mostró una especificidad perfecta tan sólo en el 5% de los casos. Los estimadores de la curva COR presentaron un comportamiento similar, con una especificidad perfecta muy baja en los dos primeros escenarios pero que se incrementa notablemente a partir del escenario 3, llegando a obtener la mayor frecuencia de casos perfectos de todas las técnicas en el escenario 5 (99,9%).

En cuanto a la proporción de sensibilidad igual al 100%, ambas metodologías (estándar y COR) mostraron baja frecuencia en los dos primeros escenarios, con incrementos en los escenarios sucesivos. Sólo la técnica Est2 tuvo una frecuencia baja que se extendió hasta el escenario 3; teniendo además los más bajos niveles de sensibilidad perfecta. Los estimadores de la curva COR junto con la metodología Est3 fueron los que aportaron más casos de sensibilidad igual al 100%, cuyos valores convergen a partir del escenario 4, con un 99%.

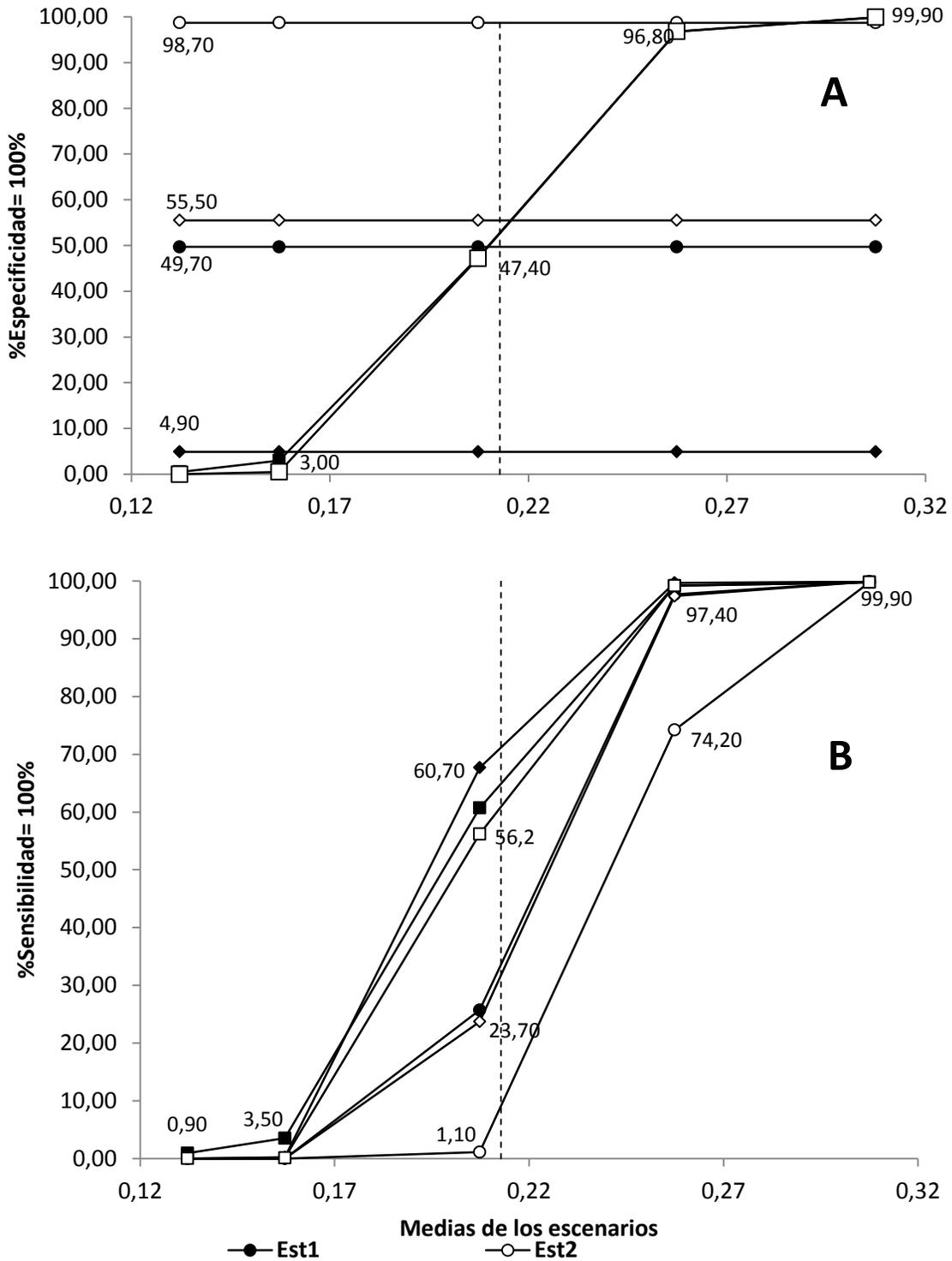


Figura 23. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 24 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en la condición heterocedásticas 1. En relación a la especificidad, la técnica que tuvo una mayor frecuencia fue la metodología Est2, con un 98,60% de los casos. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, si bien el índice de Youden fue el menos afectado por la heterocedasticidad y mostró un mejor comportamiento, ambos empezaron a tener frecuencias de especificidad perfecta alrededor del 50% en el escenario 4.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas obtuvieron resultados a partir del escenario 4, siendo la metodología Est3 la que arrojó el mayor porcentaje (72% en el escenario 5). Seguido se encuentran los estimadores de la curva COR, con un 66% de sensibilidad perfecta en el último escenario. La frecuencia de casos para la metodología Est2 no superó el 25% en ninguno de los escenarios, siendo la técnica con el menor porcentaje de sensibilidad igual al 100%.

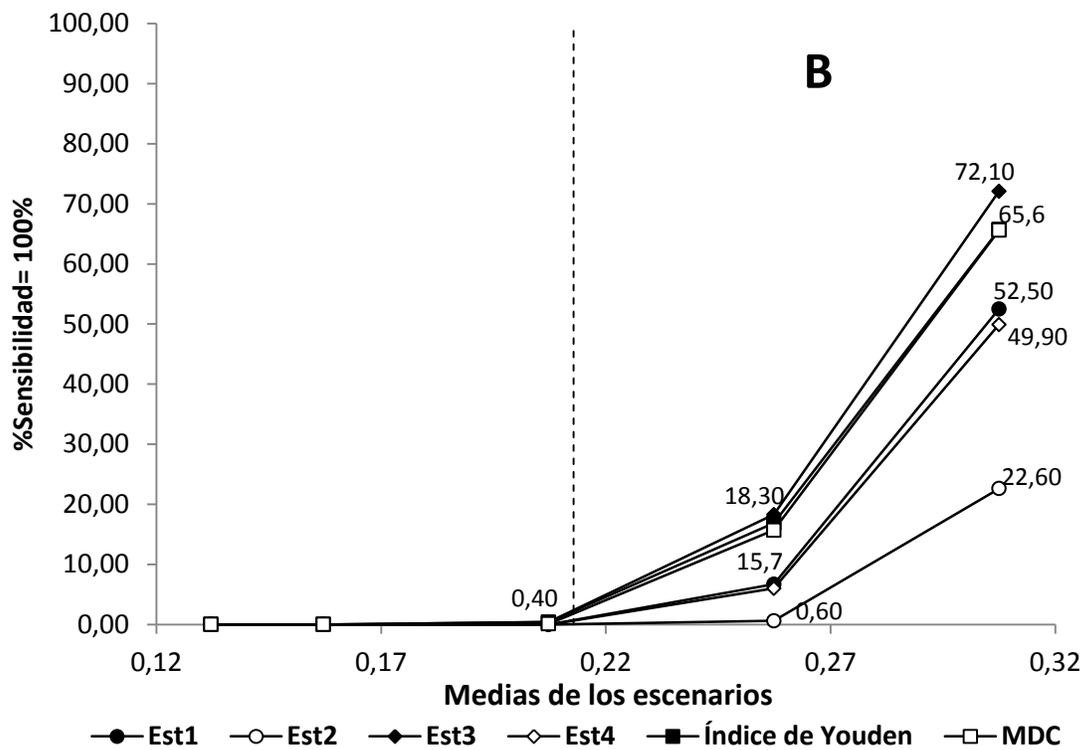
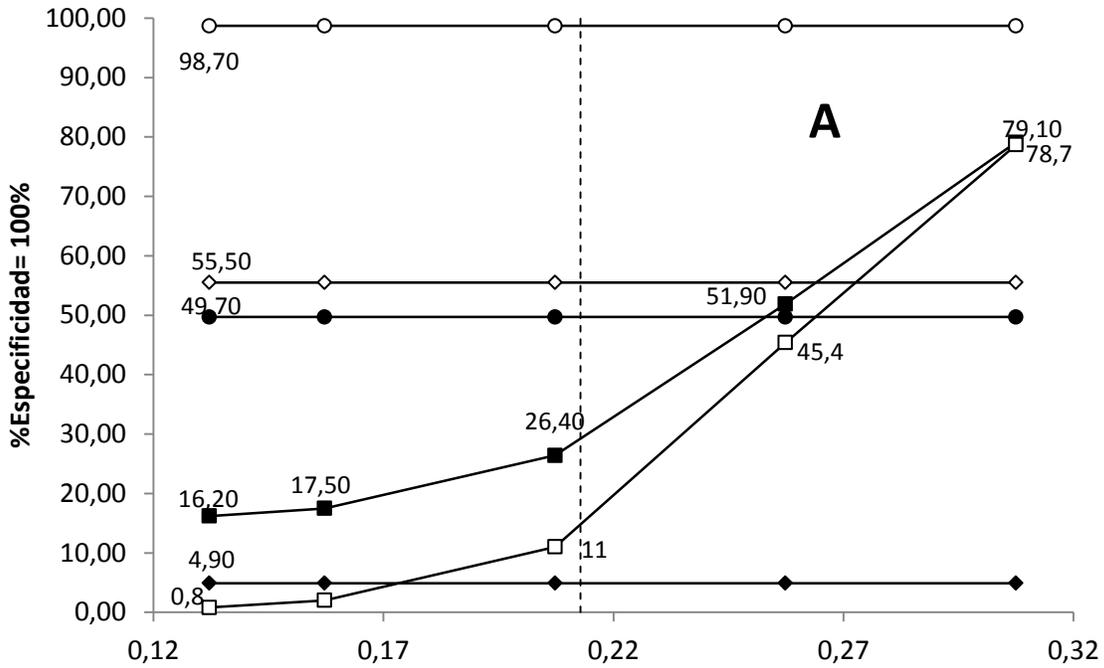


Figura 24. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 25 muestra la frecuencia en que cada una de las técnicas mostró sensibilidad o especificidad igual al 100% en la condición heterocedásticas 2. En relación a la especificidad igual al 100%, la técnica que tuvo una mayor frecuencia de acierto fue la prueba clásica Est2 con un 98,60% de los casos. El resto de las técnicas clásicas tuvo un porcentaje de acierto del 55% o menos, siendo el valor más bajo el observado con Est3 (5%). De los estimadores de la curva COR, el índice de Youden fue el que mostró un mejor comportamiento; sin embargo, las frecuencias de acierto superan el 50% sólo a partir del escenario 4.

En cuanto a la frecuencia obtenida de sensibilidad igual al 100%, todas las técnicas, con excepción de Est2, obtuvieron valores a partir del escenario 3, siendo muy bajos a este nivel y llegando a tener un porcentaje elevado en el escenario 5. La metodología clásica Est3 fue la que mostró un mejor comportamiento, con un 97% de sensibilidad perfecta en el escenario 4, seguida de los estimadores de la curva COR. El resto de las metodologías clásicas, obtuvieron valores que alcanzaron un máximo del 92%, con excepción de Est2, cuyo valor más alto fue del 73%.

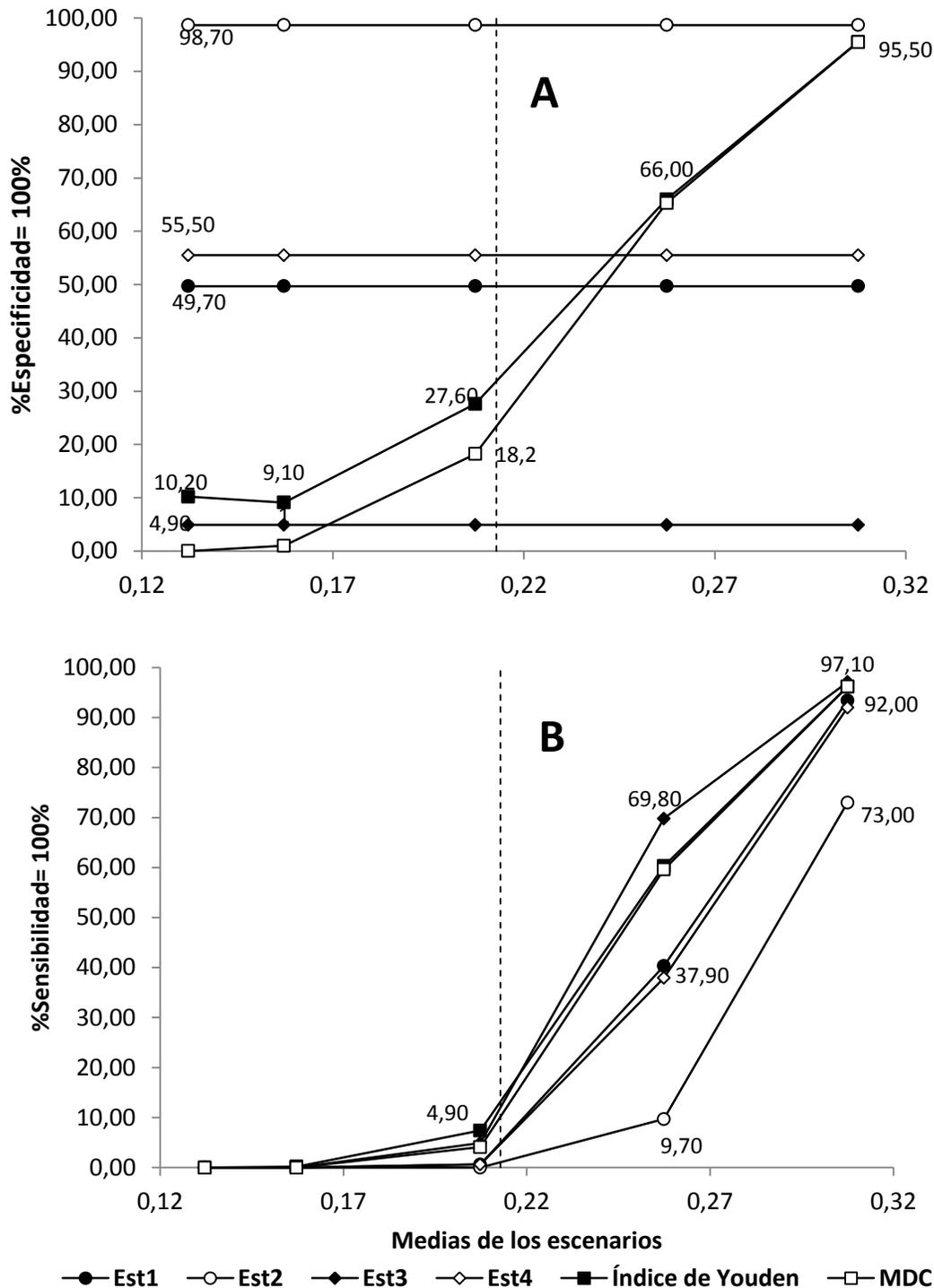


Figura 25. Estimadores de la proporción del %Especificidad= 100% (A) y %Sensibilidad= 100% (B) obtenidos con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

- **%Decisión perfecta**

Infección por Trypanosoma cruzi.

La figura 26 presenta frecuencia del número de muestras en donde se obtuvo una decisión perfecta, con una sensibilidad y especificidad igual al 100% en condiciones homocedásticas. En los dos primeros escenarios no se obtuvo con ninguna de las metodologías umbrales con decisión perfecta. A partir del escenario 3 las metodologías con el mayor porcentaje de decisión perfecta fueron las asociadas a la curva COR, en donde tanto el índice de Youden como MDC mostraron exactamente los mismos valores de aciertos, oscilando entre el 83 y el 99%. De las técnicas clásicas, tan sólo Est2 valores elevados como los de la curva COR a partir del escenario 4. El resto de las pruebas clásicas alcanzan máximo el 55% de decisión perfecta. Cabe resaltar la baja frecuencia de decisión perfecta presentada por la técnica Est2, cuyo porcentaje máximo no supera el 10%.

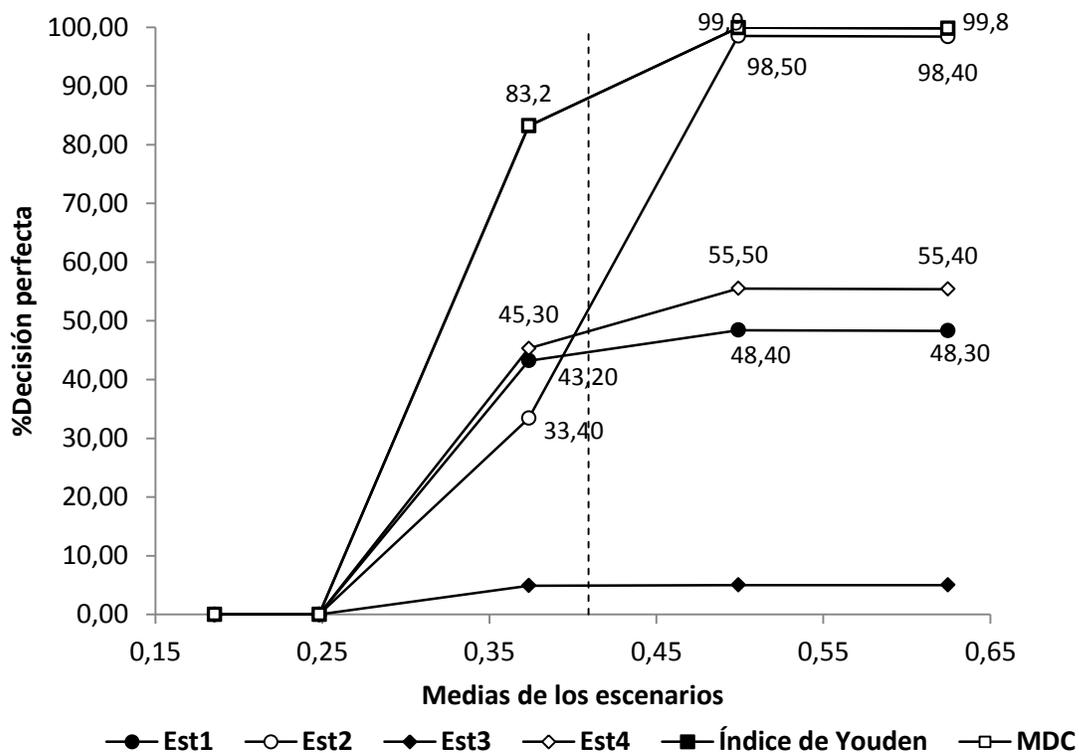


Figura 26. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La frecuencia de los casos en donde se obtuvo decisión perfecta en la condición heterocedásticas 1 es presentada en la figura 27. Es a partir del escenario 4 en donde se empiezan a observar casos de decisión perfecta siendo los mayores al emplear los estimadores de la curva COR (28%). Los mismos mostraron el doble de las frecuencias que las metodologías Est1 y Est4. Además, se observaron pocos casos de decisión perfecta con la técnica clásica Est3 (2%).

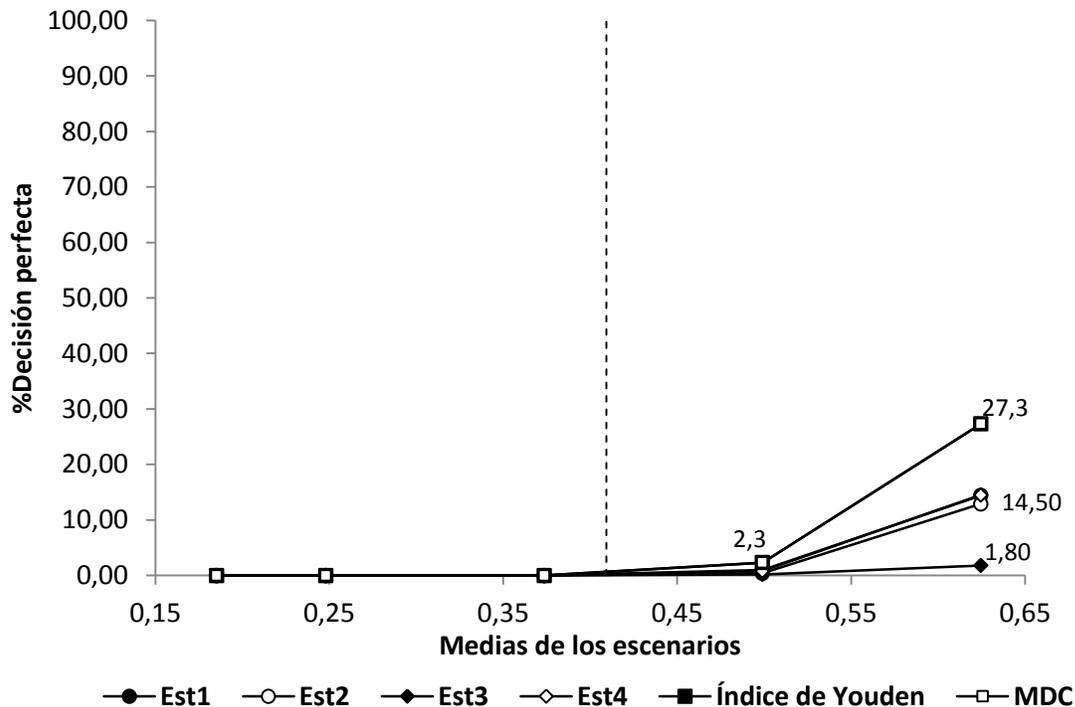


Figura 27. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La frecuencia de los casos en donde se obtuvo decisión perfecta en condiciones heterocedásticas 2, es mostrada en la figura 28. Es a partir del escenario 3 en donde se empiezan a observar casos de decisión perfecta siendo los mayores y con amplio margen los resultantes de los estimadores de la curva COR, con cifras entre los 60 y 96% y sin distinción entre el índice de Youden y MDC. De las metodologías clásicas, tan sólo Est2 alcanzó cifras superiores al 80%, mientras que el estimador de umbrales de decisión con los porcentajes más bajos de decisión perfecta fue Est3.

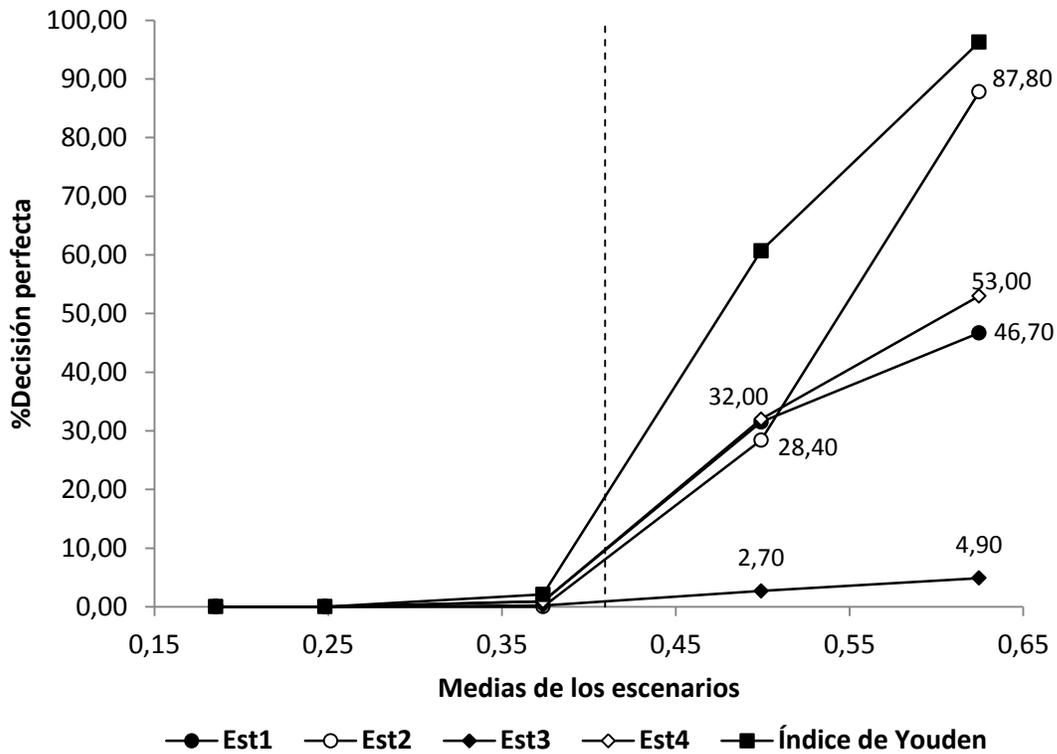


Figura 28. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Trypanosoma cruzi*. MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infeción por Leishmania sp.

La figura 29 presenta la frecuencia del número de muestras en donde se obtuvo una decisión perfecta, con una sensibilidad y especificidad igual al 100% en condiciones homocedásticas. Es a partir del escenario 3 en donde se observan casos de decisión perfecta siendo los mayores los obtenidos con los estimadores de la curva COR (95%). De las metodologías clásicas, tan sólo Est2 alcanzó cifras superiores al 60%, teniendo además como el estimador de umbrales de decisión con más bajos porcentajes de decisión perfecta a Est3 (5%).

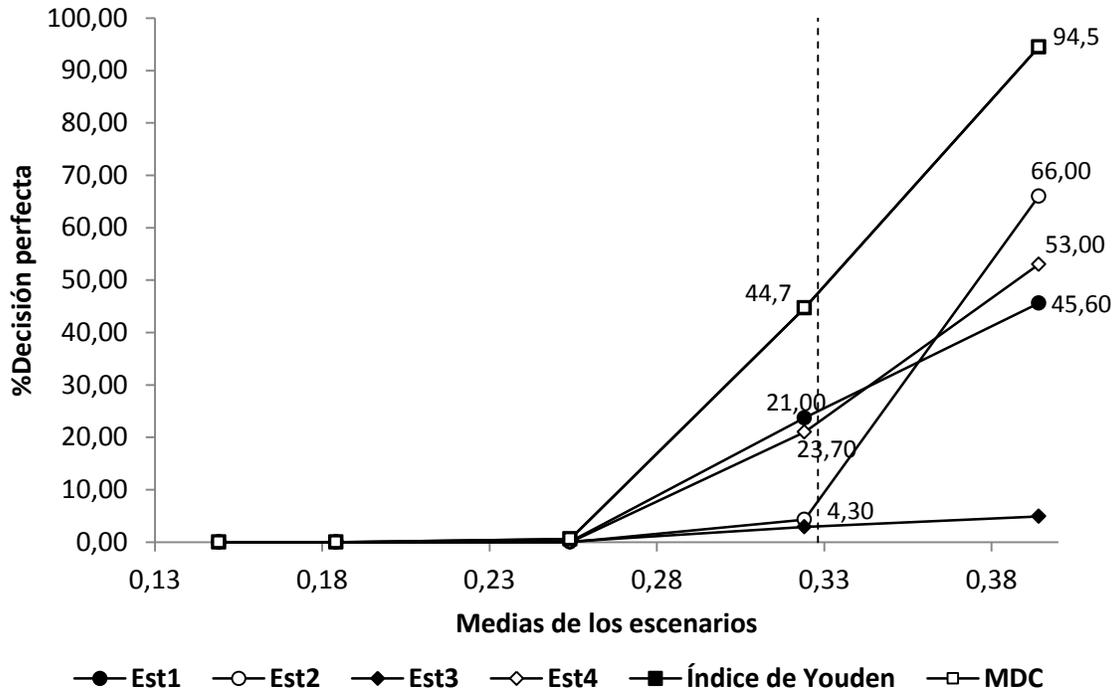


Figura 29. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La frecuencia de los casos en donde se obtuvo decisión perfecta en condiciones heterocedásticas 1 es presentada en la figura 30. Si bien, es a partir del escenario 4 en donde se empiezan a observar casos de decisión perfecta, los mismos son escasos en todas las técnicas aplicadas, pues no llegan a superar el 3%. Dentro de éstos, las cifras más elevadas son las presentadas por los estimadores de la curva COR, donde tanto el índice de Youden como MDC alcanzan un 2,8% de decisión perfecta en el escenario 5.

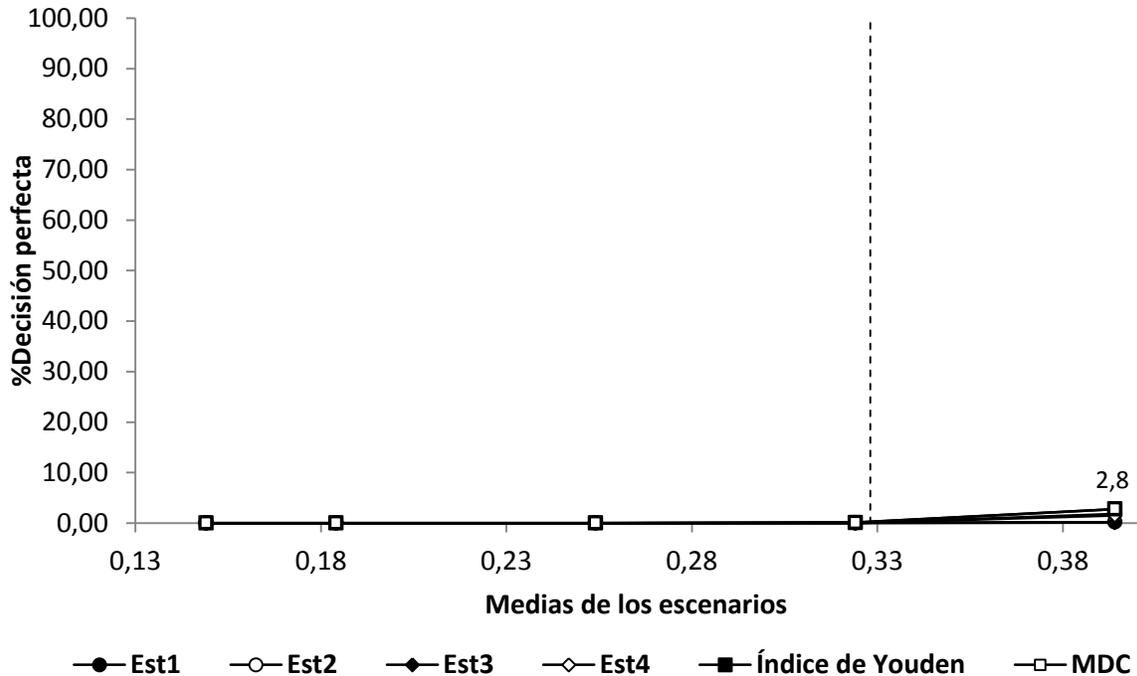


Figura 30. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La frecuencia de los casos en donde se obtuvo decisión perfecta en condiciones heterocedásticas 2 es mostrada en la figura 31. Es a partir del escenario 4 en donde se empiezan a observar casos de decisión perfecta siendo los mayores y con amplio margen los resultantes de los estimadores de la curva COR, con cifras entre los 19 y 76%, sin distinción entre el índice de Youden y MDC. De las metodologías clásicas, Est4 y Est1 mostraron el mejor porcentaje de decisión perfecta, ubicándose alrededor del 40%. La metodología Est3 fue la técnica con la menor frecuencia de decisión perfecta, ubicándose su mayor nivel en 4%.

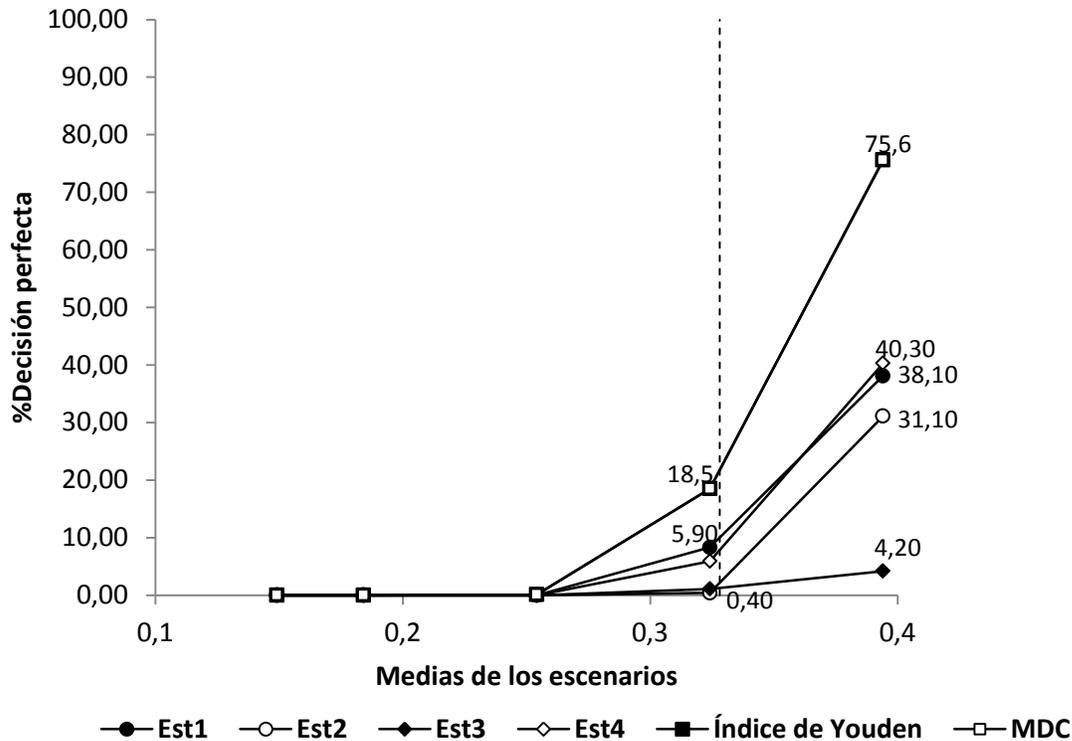


Figura 31. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Leishmania sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

Infección por Toxoplasma sp.

La figura 32 presenta la frecuencia del número de muestras en donde se obtuvo una decisión perfecta en condiciones homocedásticas. Ambos estimadores de la curva COR reportaron la mayor frecuencia de decisión perfecta en todos los escenarios. De las metodologías clásicas, tan sólo Est2 alcanzó cifras superiores al 60%, teniendo como estimador de umbrales de decisión con más bajos porcentajes de decisión perfecta a Est3 (5%).

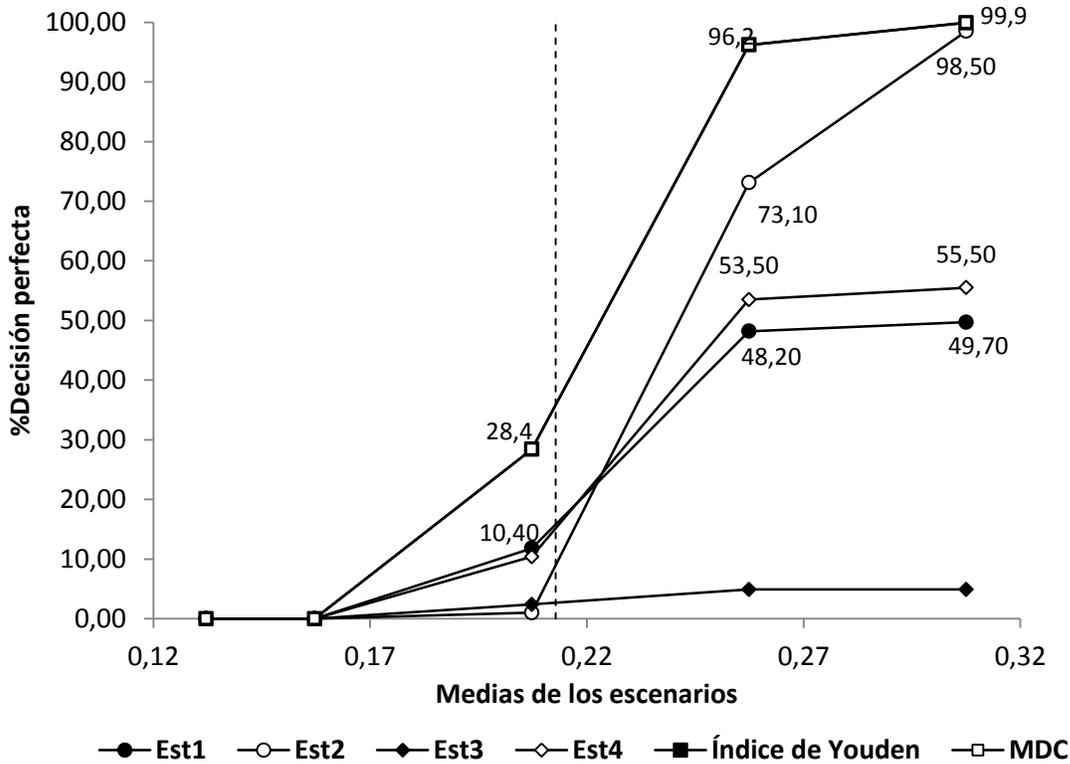


Figura 32. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones homocedásticas para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

La figura 33 muestra la proporción de casos en donde se obtuvo una decisión perfecta en la primera de las condiciones heterocedásticas. Los mayores casos de decisión perfecta fueron observados con los estimadores de la curva COR, los cuales alcanzaron un 51% de frecuencia en el escenario 5; cifra que duplica las metodologías estándar más cercanas siendo representadas por Est4 y Est1. Las cifras más bajas son arrojadas Est3.

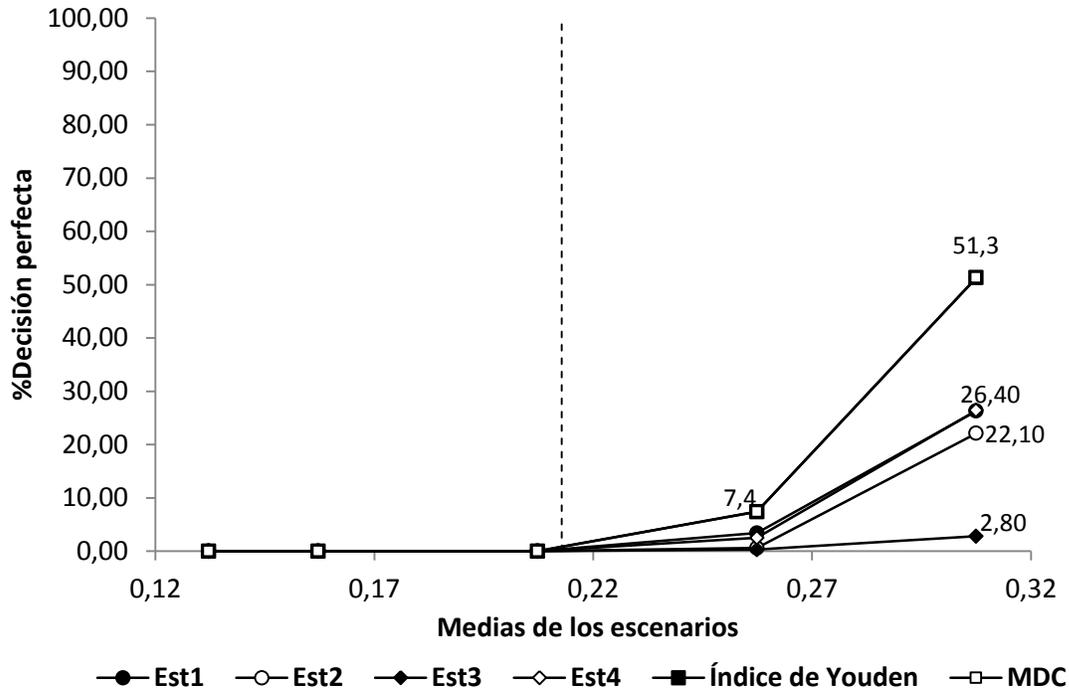


Figura 33. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 1, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

De forma similar, en la segunda de las condiciones heterocedásticas, es a partir del escenario 4 en donde se empiezan a observar casos de decisión perfecta siendo los mayores y con amplio margen los resultantes de los estimadores de la curva COR, con cifras entre 38 y 92%, sin distinción entre el índice de Youden y la MDC. De las metodologías clásicas, Est4 y Est1 mostraron el mejor porcentaje de decisión perfecta (72%) a la vez que la metodología Est3 fue la técnica con la menor frecuencia de decisión perfecta (ver figura 34).

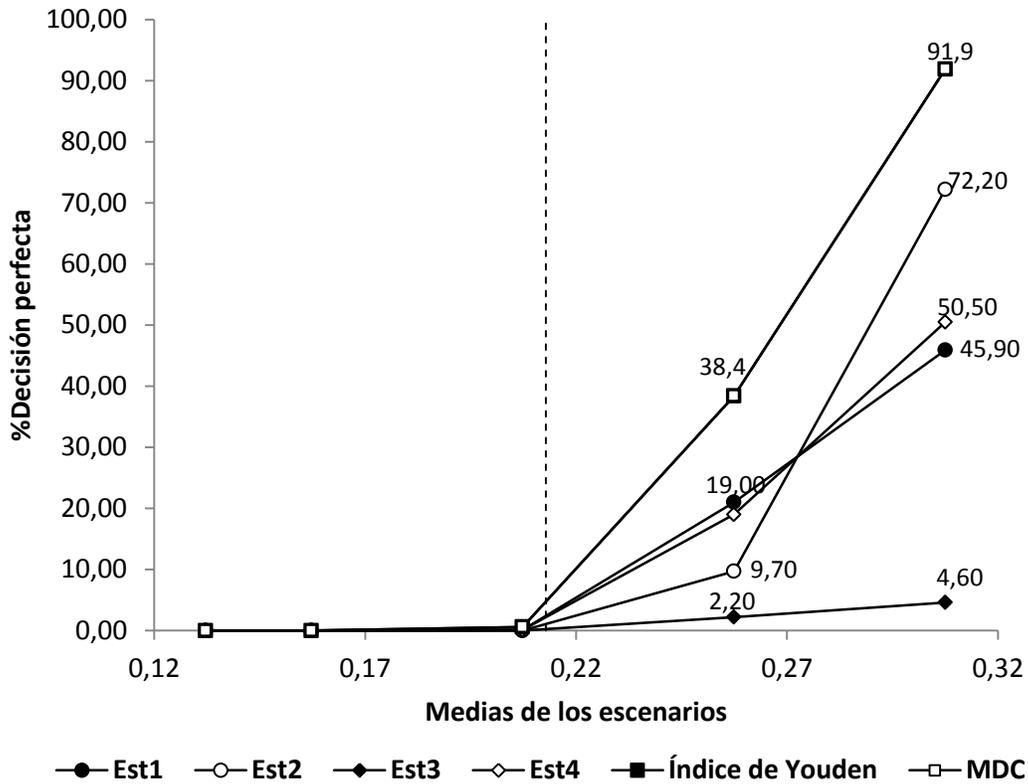


Figura 34. Estimadores de la proporción del %Decisión perfecta obtenido con los umbrales de decisión calculados a través de la metodología estándar y curvas COR en condiciones heterocedásticas 2, para el diagnóstico serológico de *Toxoplasma sp.* MDC= mínima distancia cuadrática; Est1=Metodología estándar 1; Est2=Metodología estándar 2; Est3= Metodología estándar 3; Est4= Metodología estándar 4. La línea punteada representa la media de las observaciones reales de pacientes enfermos.

DISCUSIÓN

La precisión discriminatoria representa una de las características más importantes de las pruebas diagnósticas, pues con la misma se garantiza una correcta clasificación de pacientes como sanos cuando el microorganismo a evaluar se encuentre verdaderamente ausente, o como enfermos cuando realmente posean la infección en cuestión (9). Esta precisión discriminatoria (expresada en términos de sensibilidad y especificidad), dependerá del umbral de decisión elegido que permita separar las poblaciones de pacientes sanos y enfermos, con el menor rango de error posible. El error viene dado, principalmente, producto la superposición que ocurre entre ambas poblaciones, teniendo pacientes sanos clasificados como enfermos y viceversa. En este estudio, al analizar los poblaciones, se obtuvo que efectivamente, que los datos de anticuerpos contra los tres microorganismos evaluados poseen una región de superposición entre pacientes positivos y negativos a la enfermedad respectiva, cumpliendo así con la premisa básica de las pruebas serológicas asociada a la distribución de las variables (8). Además, dicha región de no discriminación variaba de acuerdo a la enfermedad, siendo pequeña para el caso de *Leishmania sp.* y bastante amplia para *Toxoplasma sp.*, teniendo ésta última una menor precisión en la correcta clasificación de los pacientes sanos y enfermos.

Es por ello que, la metodología a utilizar para obtener el punto de corte debe ser aquella que optimice los valores de especificidad y sensibilidad, generando consigo resultados más confiables. En este orden de ideas, en la presente investigación se compararon los umbrales de decisión obtenidos en una prueba diagnóstica aplicando tanto curvas COR como metodología estándar. De acuerdo a los resultados, los puntos de corte de la metodología estándar tienden a presentar valores muy elevados que desvían el equilibrio de la precisión discriminatoria hacia la especificidad, siendo ésta alcanzada en un 100% en la mayoría de los casos. Al ser muy elevados, afectan notablemente los valores de sensibilidad, pues sólo aquellos pacientes con respuesta inmune más evidente serán los que resulten positivos a la enfermedad, teniendo gran cantidad de pacientes enfermos que serán clasificados erróneamente como sanos. Por su parte, los umbrales de decisión obtenidos mediante la curva COR (tanto con el índice de Youden como con la Mínima distancia cuadrática) generaron valores de sensibilidad y especificidad más equilibrados, alcanzando la mayoría de las veces cifras alrededor del 90%.

La principal diferencia a la cual se debe esta discrepancia en los umbrales de decisión, es que la metodología estándar no contempla las dos poblaciones en estudio (enfermas y sanas) para el cálculo del punto de corte, siendo éste obtenido solamente del promedio y desviación estándar de los datos de pacientes sanos, mientras que la curva COR considera para su construcción ambas poblaciones, siendo el umbral de decisión el resultado de un

equilibrio entre los valores de sensibilidad y especificidad de la prueba. Los resultados de ésta última, además de generar una mayor precisión discriminatoria, aportan una confianza elevada al analista, tomando como base la superposición de las poblaciones sanas y enfermas (8).

Estos resultados dan soporte a investigaciones previas relacionadas con la metodología estándar. Uno de ellos es el de Barajas-Rojas et al (28), quienes refieren que aplicar la metodología clásica de la media más dos desviaciones estándar conlleva automáticamente a obtener un valor de especificidad aproximadamente igual al 97,5%. No obstante, esta asunción sólo se mantiene en caso que la variable a estudiar se distribuya normalmente, donde Greiner and Biihning (29) hacen énfasis en que éste método no hace ningún tipo de consideración sobre la sensibilidad, siendo por tanto una metodología que no refleja la mayor función de un umbral de decisión: diferenciar las subpoblaciones de individuos infectados de aquellos con ausencia de la enfermedad. En efecto, en este trabajo los valores de especificidad para la metodología estándar fueron muy elevados (superiores al 95%) pero con una sensibilidad muy baja, que en varios casos no llegó al 60%.

Esto es muy importante, pues una prueba que arroje resultados bajos de sensibilidad generan resultados falsos negativos en gran proporción. Así lo especifican Sharma y Jain (30), quienes refieren que el método clásico está sujeto a generar resultados falsos negativos, sobre todo en casos donde pacientes positivos están recibiendo tratamiento médico y los títulos de anticuerpos o antígenos no son lo suficientemente elevados como al inicio de la infección; pues la prueba diagnóstica debería tener un punto de corte que le permita declarar al paciente exento de la enfermedad cuando realmente ya se encuentre sano. Los falsos negativos constituyen un serio problema en el manejo clínico, sobre todo cuando se trata de enfermedades graves, ya que se pierde la oportunidad de darle al paciente un tratamiento eficaz en etapas iniciales (31). En casos como en las infecciones evaluadas en el presente estudio, un error en el diagnóstico temprano conlleva al paciente sin tratar hasta una etapa crónica avanzada donde el abordaje clínico es más complicado, tal como ocurre con la infección por *Trypanosoma cruzi*.

Sharma y Jain, además, sugieren el uso de métodos estadísticos basados en evidencias analíticas y empíricas para calcular con mayor exactitud los valores de punto de corte, donde el análisis de regresión logística y las curvas COR han demostrado ser herramientas muy beneficiosas (30). Para esta investigación, los umbrales de decisión obtenidos mediante los estimadores de la curva COR generaron un mejor equilibrio entre los valores de sensibilidad y especificidad, minimizando con ello el error en la precisión discriminatoria. Otros estudios han reportado resultados exitosos empleando la Curva

COR, tal como Fernández y col (32), quienes evaluaron la utilidad de la procalcitonina para el diagnóstico precoz de infección bacteriana invasiva en el lactante febril, obteniendo una sensibilidad del 95,5% y una especificidad del 84,6%. De igual forma, Pérez y col (33), estudiaron, igualmente mediante las curvas COR, la sensibilidad y especificidad de indicadores antropométricos de adiposidad y distribución de grasa en niños y adolescentes venezolanos, obteniendo para el caso del índice de masa corporal, una sensibilidad del 86% y una especificidad del 92%. Además, una ventaja que ofrece la curva COR sobre la metodología clásica, es que permite en poblaciones de sanos y enfermos con medias muy cercanas entre sí y con una zona de no discriminación bastante amplia evaluar la calidad de la prueba serológica durante su estandarización, pues en estos casos los porcentajes de sensibilidad son bajos (tal como los obtenidos en las muestras simuladas), lo que permitiría, por ende, descartar dicha prueba.

Es necesario resaltar que, de los estimadores de curva COR evaluados, el índice de Youden es menos afectado por las condiciones heterocedásticas consideradas que la mínima distancia cuadrática, pues el primero mostró mejores valores de especificidad. Esto coincide con estudios que reportan inconsistencia en el umbral de decisión obtenido por ambos métodos, y, si bien, las dos técnicas le dan el mismo peso a los valores de sensibilidad y especificidad, diversos autores recomiendan más el cálculo del índice de Youden debido a que éste refleja la intención de maximizar el porcentaje de clasificación correcta de los pacientes sanos y enfermos (15).

Durante el proceso de validación de una prueba diagnóstica, se tiene como expectativa deseable que la misma clasifique correctamente a los pacientes de acuerdo a su condición de salud real, lo que se traduce en tener un 100% de sensibilidad y un 100% de especificidad (34). En base a esto, se evaluó en este trabajo la cantidad de veces en que el umbral de decisión, establecido por las metodologías en estudio, arrojaba un 100% de sensibilidad y 100% de especificidad de forma simultánea. Este porcentaje de decisión perfecta se observó con una frecuencia mayor al utilizar los estimadores de la curva COR, siendo superior en todos los casos a la metodología estándar. Estos resultados por ende, apoyan la hipótesis de que, entre curvas COR y metodología estándar, las primeras constituyen la opción a elegir para el cálculo de los puntos de corte en las pruebas serológicas.

Es necesario enfatizar que los estimadores de la curva COR (índice de Youden y la Mínima distancia cuadrática) identifican, tal como ocurrió en este trabajo, el punto de corte que determina la sensibilidad y especificidad conjuntamente más alta. Sin embargo, dicho punto de corte no necesariamente determina la sensibilidad ni la especificidad más alta

que podría alcanzar la prueba serológica (generalmente la sensibilidad más alta es determinada por un punto de corte, mientras que la especificidad más alta es determinada por otro) (35). De hecho, existen situaciones en las que se requiere disponer de un test diagnóstico altamente sensible (ejemplo, tamizaje de enfermedades) o bien altamente específico (ejemplo confirmación de enfermedades). En tales circunstancias, no es aconsejable utilizar el punto de corte identificado por el índice de Youden o Mínima distancia cuadrática; por el contrario, resulta más útil conocer los valores de sensibilidad y especificidad determinados por diferentes puntos de corte, y optar por aquel que determine la mayor sensibilidad, o la mayor especificidad, según sea el objetivo (36).

CONCLUSIONES

- En el análisis para las poblaciones, las curvas COR generaron umbrales de decisión menores que los de la metodología estándar, ubicándose además dentro de la región de superposición de datos.
- En las poblaciones, las curvas COR mostraron valores de sensibilidad y especificidad más equilibrados que las metodologías estándar; sin embargo, tales valores fueron afectados de forma inversa por el tamaño relativo de la región de superposición de datos, siendo mejores en la población de *Leishmania sp.*, y de menor calidad para *Toxoplasma sp.*, las cuales a su vez presentaron la menor y la mayor región de superposición, respectivamente.
- La metodología estándar en las muestras simuladas generó, en la mayoría de los casos, umbrales de decisión más elevados a los obtenidos por los estimadores de la curva COR, estando estos últimos afectados por la varianza de las poblaciones de pacientes enfermos.
- Los estimadores de la curva COR en las muestras simuladas generaron valores de sensibilidad y especificidad dinámicos, los cuales aumentaron de forma progresiva a medida que la media de la población de enfermos se aleja de la media de la población sana, sin embargo, la sensibilidad tendió a ser mayor para el índice de Youden en presencia de heterocedasticidad.
- Es mayor el número de casos en donde la sensibilidad y especificidad son iguales al 100% de forma simultánea (decisión perfecta) al utilizar la curva COR para estimar el umbral de decisión de la prueba diagnóstica, que utilizando la metodología estándar, por lo cual la precisión discriminatoria de las pruebas diagnósticas tendió a ser mejor con el uso de las curvas COR.

RECOMENDACIONES

- En el caso de las infecciones evaluadas en este estudio, aplicar los estimadores de la curva COR para el establecimiento de los umbrales de decisión en las pruebas serológicas en los laboratorios clínicos e investigación del país, debido a que muestra una mejor precisión discriminadora que la metodología estándar.
- Fomentar, en los centros de diagnóstico clínico e investigación universitaria, estudios que comparen la precisión de los estimadores de la curva COR y metodología estándar en el cálculo de los umbrales de decisión de pruebas serológicas que diagnostiquen otro tipo de enfermedades infecciosas a las contempladas en esta investigación y que incluyan covariables clínicas que puedan influir en la sensibilidad y especificidad de la prueba.
- Diseñar un software que incluya en su paquete estadístico el cálculo de los umbrales de decisión en pruebas diagnósticas usando ambos estimadores de la curva COR, siguiendo la rutina planteada en esta investigación.

REFERENCIAS BIBLIOGRÁFICAS

1. Ingraham J, Ingraham C. Introducción a la microbiología. Barcelona: Editorial Reverté, S.A. 1998
2. Akobeng A. Understanding diagnostic tests 3: receiver operating characteristic curves. *Acta Pædiatrica* 2007;96:644–647.
3. Rha S, Yang W, Gong S, Kim J, Kim B, Chung H. Correlation of Tissue and Blood Plasminogen Activation System in Breast Cancer. *Cancer Letters* 2000;150:137-145.
4. Ridge S, Vizard A. Determination of the Optimal Cutoff Value for a Serological Assay: an Example Using the John's Absorbed EIA. *J Clin Microbiol* 1993;31(5):1256-1261.
5. Altman D, Bland J. Diagnostic tests 3: receiver operating characteristic plots. *BMJ* 1994;309:188.
6. Obuchowski N, Lieber M, Wians F. ROC Curves in clinical chemistry: uses, misuses, and possible solutions. *Clin Chem* 2004;50:1118–1125.
7. Fan J, Upadhye S, Worster A. Understanding receiver operating characteristic (ROC) curves. *Can J Emerg Med* 2006;8(1):19-20.
8. Worster A, Innes G, Abu-Laban R. Diagnostic testing: an emergency medicine perspective. *CJEM/JCMU* 20032;4(5):348-354.
9. Shapiro D. The interpretation of diagnostic tests. *Statistical Methods in Medical Research* 1999;8:113-134.
10. Lorca M, Contreras M, Salinas P, Guerra A, Raychaudhuri S. Evaluación de una prueba rápida para el diagnóstico de la infección por *Trypanosoma cruzi* en suero. *Parasitol Latinoam* 2008;63:29-33.
11. Borbolla D, Otero C, Lobach D, Kawamoto K, Gomez A, Lopez G, et al. Evaluación de la Sensibilidad y Especificidad de un Sistema de Ayuda a la Toma de Decisiones como modelo de servicio. *RevistaSalud.com* 2009;5(18):1-9.
12. Carvajal C, Morales M. Sensibilidad y especificidad de los marcadores tumorales. *Rev. méd. Cochabamba* 2010;21(1):86-97.
13. Martínez-Camblora P. Comparación de pruebas diagnósticas desde la curva ROC. *Revista Colombiana de Estadística* 2007;30:163–176.
14. Grzybowski M, Younger J. Statistical Methodology: III. Receiver operating Characteristic (ROC) curves. *Acad Emerg Med* 1997;4:818-826.
15. Perkins N, Schisterman E. The inconsistency of 'optimal' cutpoints obtained using two criteria based on the receiver operating characteristics curve. *Am J Epidemiol* 2006;163:670–675.
16. Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. *Biom J* 2005;47:458–472.
17. Schisterman E, Perkins N, Liu A, Bondell H. Optimal Cut-point and Its Corresponding Youden Index to Discriminate Individuals Using Pooled Blood Samples. *Epidemiology* 2005;16:73–81.
18. Faraggi D, Reiser B. Estimation of the area under the ROC curve. *Stat Med* 2002;21:3093-3106.

19. Bamber D. The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology* 1975;12:387–415.
20. Peterson W, Birdsall T, Fox W. The theory of signal detectability. *Transactions of the Institute of Radio Engineers, Professional Group on Information Theory, PGIT* 1954;4:171-212.
21. Lusted L. Signal detectability and medical decision making. *Science* 197;171:1217–1219.
22. Metz C. Roc methodology in radiologic imaging. *Invest Radiol* 1986;21:720-733.
23. Robertson E, Zweig M. Use of Receiver Operating Characteristic Curves to evaluate the clinical performance of analytical systems. *Clin Chem* 1981;27:1569-1574.
24. Zweig M, Campbell G. Receiver-Operating Characteristic (ROC) Plots: A fundamental evaluation tool in clinical medicine. *Clin Chem* 1993;39:561-577.
25. Kim I, Choiy Y, Chungz H, Rhax S, Kim B. Statistical method of determining a cut off value between normal and disease groups. *Bulletin of Informatics and Cybernetics* 2004; 36:63-72.
26. Trajstman A. Diagnostic tests, sensitivity, specificity, efficiency and prevalence. *Aust Vet J* 1979; 55:501.
27. MEDIAactive. Manual de Excel 2010. España: MARCOMBO, S.A. 2010.
28. Barajas-Rojas J, Riemann H, Franti C. E. Notes about determining the cut-off value in enzyme-linked immunosorbent assay (ELISA) [letter to the editor]. *Prev Vet Med* 1993; 15:231-233.
29. Greiner M, Bohning D. Notes about determining the cut-off value in enzyme-linked immunosorbent assay (ELISA) –Reply. *Prev Vet Med* 1994; 20:307-310.
30. Sharma B, Jain R. Right choice of a method for determination of cut-off values: A statistical tool for a diagnostic test. *Asian Journal of Medical Sciences* 2014; 5:30-34.
31. Donis J. Evaluación de la validez y confiabilidad de una prueba diagnóstica. *Avances en Biomedicina* 2012; 1(2):73-81.
32. Fernández A, Luaces C, Valls C, Ortega J, García J, Mira A, Pou J. Procalcitonina para el diagnóstico precoz de infección bacteriana invasiva en el lactante febril *Anales de Pediatría* 2001; 55(4):321-328.
33. Pérez B, Landaeta-Jiménez B, Amador J, Vásquez M, Marrodán M. Sensibilidad y especificidad de indicadores antropométricos de adiposidad y distribución de grasa en niños y adolescentes Venezolanos. *INCI* 2009; 34(2):84-90.
34. Smith S. Diagnostic tests (1) – sensitivity and specificity. *Phlebology* 2012; 27:250-251.
35. Cerda J, Cifuentes L. Uso de curvas ROC en investigación clínica. Aspectos teórico-prácticos. *Rev Chil Infect* 2012; 29(2): 138-141.
36. Tarabala H. Validación de pruebas diagnósticas. *Revista FAVE* 2002; 37-42.