

Valencia 22 de septiembre de 2003

Dra. Jacqueline Saulny de Jorges  
 Editora en Jefe de la Revista de Obstetricia y  
 Ginecología de Venezuela.  
 Su Despacho.

Luego de un cordial saludo, me permito hacerle llegar algunas observaciones sobre las respuestas que a mi anterior comunicación dieran los autores y asesores del trabajo "Velocimetría Doppler de la arteria uterina como factor de predicción de pre-eclampsia y crecimiento fetal restringido" y que fueron publicadas en la sección de cartas al editor Rev Obstet Ginecol Venez 2003;63(1):55-60.

Mucho sabré agradecerle que incluya el contenido de la presente en la misma sección en el próximo número de la revista que tan acertadamente dirige.

Sin otro particular a que hacer referencia, quedo de usted

Atentamente

Dr. Alberto Sosa Olavarría

Anexo: Lo indicado en el texto.

1. La estrategia metodológica adoptada por los autores creo que responde a un diseño **observacional, analítico** (no descriptivo como afirman), prospectivo y transversal en **cohorte única** con selección intencional de la muestra, orientado a **evaluar la capacidad diagnóstica** de una tecnología perinatal. Si esto hubiese sido explícitamente señalado al inicio del trabajo, se justificaría la afirmación del Prof. Avilán Rovira de "que la mayoría de los lectores conocen que los trabajos de determinación (sic) de la sensibilidad y especificidad de las pruebas y signos diagnósticos se basan en la metodología bayesiana".
2. Los autores reformularon los cuadros según las sugerencias, pero me resulta incomprensible que **los cuadros 3 y 4 sean idénticos**. Según mis cálculos esa tabla debería ser la que muestro en el **anexo**.
3. Los autores emplean los instrumentos adecuados al diseño: sensibilidad, especificidad, etc., omiten

los falsos positivos y negativos, y pudieran haber complementado con el índice de Youden o el de Kappa que son herramientas intrínsecas al objetivo inferencial y no el de Riesgo Relativo (su IC 95% exige Chi cuadrado corrección de Yates, en este caso), ni la prueba exacta de Fisher con las que uno de los especialistas consultados pretende reforzar las conclusiones. No comparto el criterio de aplicar todas las pruebas estadísticas inferenciales para demostrar significancia por múltiples vías y menos si estas no se ajustan al diseño pre-establecido o seleccionado por los autores.

4. Los autores emplean acertadamente en su corrección como **prueba de hipótesis el IC al 95%** de cada índice, algo que omitieron en el trabajo original. Esto me parece **excelente** pues el médico actual acostumbrado a la lectura crítica e inclinado al análisis meta-analítico, encontrará mejor fundamentación para aplicar los resultados al ejercicio clínico cotidiano.
5. Si el diseño aplicado sobre la muestra hubiese sido el enfoque de riesgo, la conclusión obvia sería la siguiente: "cuando está presente la incisura en las velocidades diastólicas de la arteria uterina (factor de riesgo), la posibilidad de HIE y CFR (efecto o daño) es de **x** y **z** veces, respectivamente, mayor que cuando la incisura está ausente". Como quiera que el diseño seleccionado es con la intención de evaluar la capacidad diagnóstica de un procedimiento la conclusión, según mi opinión ha debido ser: "**Las características de la OVF Doppler en la arteria uterina permiten en el caso de presentar incisuras en las velocidades diastólicas, predecir la existencia de HIE con una sensibilidad de 62 % y un VPP de 35,7 %, siendo la capacidad de la prueba mucho más confiable para excluir la HIE, con una especificidad de 90 % y un VPN de 96,5 %. En relación al CFR la capacidad diagnóstica de la presencia de incisura demostró una sensibilidad de 87,5 % y un VPP de 87,5 %, siendo la capacidad de la ausencia de la incisura diastólica para excluir el CFR más elevada (especificidad de 98,9 % y VPN de 98,9 %). Dados los IC95 % de elevada significancia obtenidos para estos índices, cabe esperar que sus valores predictivos se mantengan con prevalencias comprendidas entre 5 % y 14 % tanto para la HIE como para el CFR.**"

6. No comparto el párrafo expresado por uno de los profesores consultados de que "...por considerar que los lectores de estos temas están familiarizados con estos cálculos, no parece necesario incluirlos en el informe". Todo lo contrario, hoy en día cuando la medicina basada en la evidencia (MBE) parece ser el paradigma predominante, mientras más explícita y prolija sea la información acerca de los criterios de selección de la muestra y del procesamiento de los datos recaudados, más fácil será para el acucioso lector, establecer la solidez entre lo analizado y lo concluido por el investigador y por otra parte le facilitará al experto aplicar el meta-análisis.
7. La diferencia entre los cálculos que aporté (calculadora del Dr. Hutchon) y los del Prof. Avilán son producto del empleo de fórmulas diferentes, es harto conocido que los valores predictivos y la exactitud son influenciados por la prevalencia, y de allí la importancia de calcular estos índices tanto con la **prevalencia aparente** (encontrada en la muestra) como con la **prevalencia real** (previa o pre-test) de la población sobre la cual se pretenden aplicar los resultados; de allí la importancia del enfoque bayesiano. Resulta obvio que la probabilidad nunca puede ser mayor a la unidad, pero la relación de verosimilitud puede mostrar valores inferiores a esta y al excluirla, inclina el mejor resultado de la prueba hacia los valores predictivos negativos, es decir, que favorece el diagnóstico de casos negativos (sanos).
8. Coincido con lo afirmado por el Prof. Avilán cuando señala que: "Se piensa también que el médico utilizará los valores de sensibilidad y especificidad (o las razones de verosimilitud) para calcular las probabilidades pos test con las probabilidades pre test (o a priori) que le estime a sus pacientes"; pero en este sentido debo agregar, sin afirmar que la verdad absoluta esté de nuestra parte, que dentro de las orientaciones que doy a mis alumnos, les sugiero que **al leer críticamente un trabajo** destinado a evaluar la capacidad diagnóstica de una prueba, **observen detenidamente la prevalencia aparente**, y si esta es igual a la de la población sobre la cual ellos intentan aplicar los resultados del estudio, orienten a sus pacientes en base a los valores reportados para **la sensibilidad y especificidad**. **Si desconocen la prevalencia a priori**, real o pre test, o estas son inferiores o superiores a la

del estudio, deben **recorrer a los valores predictivos positivos y negativos realizando la corrección respectiva según la prevalencia en la población** y teniendo en cuenta que para una prevalencia baja cabe esperar baja predicción y lo contrario para una prevalencia elevada. Finalmente si queremos **prescindir de la corrección o a comparar los resultados de dos pruebas, debemos utilizar los coeficientes de probabilidad o de verosimilitud positivos y negativos**, que como se sabe no dependen de la proporción de enfermos en la muestra. Por lo anteriormente expuesto y a manera de conclusión pudiéramos afirmar que en todo trabajo de investigación que descansa sobre el tipo de diseño de evaluación de la capacidad diagnóstica de una prueba, debería incluir todos los índices que han motivado este interesante y productivo intercambio de opiniones.

Anexo:

El cuadro 4: Índice S/D y predicción de crecimiento fetal restringido

7      1:      8 positivos probados  
1      91:     92 negativos probados  
In total 100 observations

**Seguridad** eq: 98/100; proportion: 0,98  
95 % CI: 0,953<Ac<1,007; Wilson CI: 0,923<Ac<0,997  
**Kappa** medida de acuerdo: 0,864  
(98-(8528/100))/(100-(8528/100))  
95 % CI: 0,668<Kappa<1,06  
**Sensibilidad: 0,875** (7/8)  
95 % CI: 0,646<Se<1,104; Wilson CI: 0,467<Se<0,993  
**Especificidad: 0,9891** (91/92)  
95 % CI: 0,968<Sp<1,01; Wilson CI: 0,932<Sp<0,999  
**Probabilidad positiva: 80,5** (0,75/(1-0,989))  
95 % CI: 11,262<PL<575,416  
**Probabilidad negativa: 7,913** (0,989/(1-0,875))  
95 % CI: 1,265<NL<49,504  
**J de Youden: 0,864** (0,989+0,875-1)  
95 % CI: 0,634<J<1,094  
**Prevalencia** eq: 8/100; proporción: 0,08  
95 % CI: 0,002<pr<0,158; Wilson CI: 0,038<pr<0,156  
**Seguridad predictiva positiva: 0,875** (7/8)  
95 % CI: 0,646<pp<1,104; Wilson CI: 0,467<pp<0,993  
**Seguridad predictiva negativa: 0,989** (91/92)  
95 % CI: 0,917<np<1,061; Wilson CI: 0,585<np<0,993

Caracas: de noviembre de 2003.

Dra.

Jacqueline Saulny de Jorges

Editora en Jefe de la

Revista de Obstetricia y Ginecología de Venezuela  
Presente.

Apreciada doctora Jorges:

Agradezco el envío de la carta del Dr. Alberto Sosa Olavarría, de fecha 22 de septiembre pasado, en respuesta a las cartas publicadas en las páginas 56-59, del volumen 63, número 1, de marzo de este año, pues en ella se me menciona directa o indirectamente en más de una ocasión.

En primer lugar, debo expresarle que me parece muy importante que la Revista que tan acertadamente dirige estimule este intercambio de opiniones entre sus lectores, pues permite aclarar conceptos, enriquecer nuestros conocimientos y en definitiva, al final, los más beneficiados serán los pacientes.

Es necesario aclarar, que cuando escribí mi carta-respuesta-1 (páginas 58,59), desconocía el contenido de la carta-respuesta de los autores (Quintero J, Brito J y Angulo D. Páginas 56,57), del cual me enteré al leer la Revista. Obsérvese que la fecha de la carta de los autores es el 14-04-03 y la mía es del 04-03-03.

Para evitar dejar algún punto fuera, comentaré punto por punto la carta del Dr. Sosa Olavarría, cerrando así, por mi parte, el intercambio de opiniones sobre este interesante trabajo publicado por ustedes.

1. Según el Dr. Sosa Olavarría la metodología del trabajo que nos ocupa (1) corresponde a un diseño “observacional, analítico”, pero para los autores es “descriptivo”. Si bien al final de la sección “Material y método”, los autores mencionan que analizaron las frecuencias y los porcentajes, para el *análisis cuantitativo* utilizaron medias aritméticas y desviaciones estándar y estimaron sensibilidad, especificidad y valores de predictivos realizaron comparaciones mediante pruebas de significancia estadística. ¿Quién tiene la razón? Sería una discusión interminable que consumiría el espacio necesario para los próximos puntos.

No comparto la idea de intentar clasificar los trabajos en las publicaciones. Si la descripción de la sección, que en general se denomina,

“Sujetos, materiales y métodos” es lo suficientemente explícita, la clasificación es innecesaria. El lector dispondrá de los elementos requeridos para su eventual clasificación.

En el segundo aspecto de este punto, el Dr. Sosa Olavarría menciona el seguimiento de una *cohorte única para evaluar la capacidad diagnóstica de una tecnología perinatal*. No compartimos lo de cohorte única, pues las embarazadas se distinguieron en 3 grupos de acuerdo a la incisura y en dos, según el índice SID. En cuanto al objetivo del estudio no fue únicamente evaluar la capacidad diagnóstica de la tecnología, sino “determinar su utilidad...para la *predicción* de preeclampsia y crecimiento fetal restringido” (1).

2. En realidad, creemos que por error, se repitieron las cifras del Cuadro 3 en el Cuadro 4, pero los resultados de los cálculos bajo este último Cuadro, corresponden a las cifras correctas de las cuatro casillas: 7, 1, 1, 91 (utilizando el mismo orden de los autores), que en general coinciden con los calculados por el Dr. Sosa Olavarría en el anexo a su carta.

3. La “J” de Youden, ideada por su autor en 1950 (2), es un índice combinado de la sensibilidad y la especificidad de una prueba diagnóstica, para intentar comparar pruebas diagnósticas. Ya está en desuso, pues las pruebas se comparan mejor evaluando las diferencias entre las sensibilidades y especificidades por separado. Para los datos del Cuadro 4, citados en el punto anterior, el Dr. Sosa Olavarría calculó una “J” de 0,864, para una sensibilidad de 0,875 y una especificidad de  $0,989 = (0,875 + 0,989) - 1$ . Si invertimos los datos: 91, 1, 1, 7, para una sensibilidad de 0,989 y una especificidad de 0,875, la “J” sería la misma, pero las pruebas diagnósticas son diferentes. Lo mismo ocurre con el índice *kappa* de Cohen.

Los autores del trabajo, en su carta a la Revista (página 57), afirman en el penúltimo párrafo: la persistencia de la incisura diastólica bilateral constituye un grupo de alto riesgo para el desarrollo de complicaciones como HIE y CFR”. Para decidir estadísticamente debe calcularse el riesgo relativo, entre los dos grupos, al finalizar el seguimiento de las cohortes. Para verificar su significancia estadística se calcula la probabilidad asociada a la diferencia entre las tasas con la prueba exacta de Fisher. Calculamos los límites

de confianza del riesgo relativo, no para verificar su significancia estadística, sino para conocer la amplitud de su variación. No hay pues redundancia de pruebas estadísticas. Para este cálculo no utilizamos el Ji cuadrado, que resultaría inválido en este caso, dada la magnitud de las frecuencias esperadas, pues existen otros métodos que no lo requieren, disponibles en el programa Epi-info versión N° 5. Nunca utilizamos corrección de Yates, pues con la facilidad de la ejecución de la prueba exacta de Fisher, hasta con una calculadora corriente, en la actualidad cada vez se usa menos.

En mi opinión, este es el análisis que correspondería al diseño. En efecto, los autores pretenden sostener sus conclusiones con “un valor predictivo positivo de 37,5 % con relación a HIE; y un valor predictivo positivo del 50 % en cuanto a CFR”, además de los valores correspondientes de las sensibilidades respectivas. Ambos valores de predicción están por encima del umbral para excluir un diagnóstico y muy por debajo del dintel para aceptar un diagnóstico, según Pauker y Kassirer (3).

4. Los autores no utilizan el intervalo de confianza como prueba (sic) de hipótesis, sino para mostrar la variabilidad de las estimaciones de sensibilidad, especificidad o valores predictivos, calculados con las cifras obtenidas. Para utilizar los intervalos de confianza en docimasia de hipótesis, se requiere estimarlos con las diferencias entre las sensibilidades (o especificidades o valores predictivos según sea la comparación).
5. Lo consideramos respondido en 3,
6. El Dr. Sosa Olavarría tiene razón. Pero los cálculos a los cuales nos referimos no son los del tamaño de la muestra ni los “del procesamiento de los datos recaudados”. Nos referimos a las estimaciones de valores postests con variación de valores pretests, pues asumimos que los lectores de este tipo de trabajos los dominan. Además, el médico utiliza los resultados de estos trabajos para estimar los valores postests en sus propios pacientes, basándose en los valores pretests que estime en esos mismos pacientes (4).

7. Se utilizó exactamente la misma fórmula, es decir, la de las razones de verosimilitud (*likelihood ratios*). La diferencia en los resultados se debe a que con esta fórmula sólo pueden usarse razones. En lugar de la probabilidad pretest, se usa su razón: P/1-P. Así, para calcular la probabilidad postest al variar la prevalencia de 0,08 a 0,144, se debió usar la razón 0,144/0,856.
8. Sería muy largo discutir este punto. Además, en los comentarios precedentes ya lo consideramos incluido. Sólo nos referiremos a la posibilidad “Si desconocen la prevalencia *a priori*, real (sic) o pretest...” Insistimos nuevamente: el médico para poder utilizar la metodología bayesiana, en el cálculo de los valores postests de sus pacientes, necesita de 3 probabilidades: la pretest y la sensibilidad y la especificidad de la prueba diagnóstica que indique. Si no sabe estimar la probabilidad pretest de la enfermedad de su paciente, de muy poco valen las otras dos.

Atentamente,

J M Avilán Rovira.

Director de la Gaceta Médica de Caracas.

#### REFERENCIAS

1. Quintero J, Villamediana J, Paravisini I, Brito J, Cadena L. Velocimetría Doppler de la arteria uterina como factor de predicción de preeclampsia y crecimiento fetal restringido. *Rev Obstet Ginecol Venez*. 2002; 62(3):153-159.
2. Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950;3:32-35.
3. Panzer Ri, Black ER, Griner PF (Editores). *Diagnostic strategies for comon medical problems*. Filadelfia PA: American College of Physicians; 1991.
4. Sackett DL, Brian-Haynes R, Tugwell P. *Clinical Epidemiology*. Boston: Little Brown & Co; 1985.