

Estudios de exactitud diagnóstica

¿Son los resultados del estudio válidos? (Validez interna)

¿El test diagnóstico fue evaluado en un espectro representativo de pacientes (como aquellos en los cuales se les aplicará en la práctica)?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Es ideal si el test diagnóstico es aplicado al espectro completo de pacientes – es decir, leves, severos, etapa temprana y tardía del desorden de interés. También es mejor si los pacientes se seleccionan al azar o según admisión consecutiva de manera de minimizar el sesgo de selección.	La sección de Métodos debería decir cómo fueron reclutados los pacientes y si fueron seleccionados al azar o a medida que fueron admitidos. También debería decir de donde vienen los pacientes y si es probable que sean representativos de los pacientes en quienes se usará el test.
Este artículo: Si <input type="checkbox"/> No <input type="checkbox"/> No está claro <input type="checkbox"/> Comentarios:	
¿El estándar de referencia fue aplicado independiente del resultado de la prueba diagnóstica?	
¿Qué es lo mejor?	¿Dónde encuentro esta información?
Idealmente tanto la prueba diagnóstica y el estándar de referencia deberían ser aplicados en todos los pacientes del estudio. En algunas situaciones donde el estándar de referencia es invasivo o caro podría haber reparos respecto a someter a la prueba a sujetos con la prueba diagnóstica negativa (y por lo tanto una baja probabilidad de enfermedad). Una alternativa al estándar de referencia es seguir a las personas por un tiempo apropiado (dependiendo de la enfermedad en cuestión) para ver si son verdaderos negativos.	La sección de Métodos debería indicar si el estándar de referencia fue aplicado a todos los pacientes o si un estándar de referencia alternativo (es decir, seguimiento) fue aplicado a aquellos que fueron negativos en la prueba diagnóstica
Este artículo: Si <input type="checkbox"/> No <input type="checkbox"/> No está claro <input type="checkbox"/> Comentarios:	
¿Hubo una comparación independiente, ciega entre la prueba diagnóstica y un (“gold”) estándar de referencia apropiado de diagnóstico?	
¿Qué es mejor?	¿Dónde encuentro esta información?
Hay dos temas importantes aquí. Primero que el estándar de referencia debería ser apropiado – lo más cercano a la verdad posible. Algunas veces puede que no haya un solo test de referencia apropiado y se puede usar una combinación de pruebas para determinar la presencia de enfermedad. Segundo, el estándar de referencia y la prueba diagnóstica que están siendo evaluados deberían aplicarse de manera independiente y ciega. Aquellos que interpretan los resultados de un test no deberían saber de los resultados del otro test.	En la sección de Métodos debería haber una descripción del estándar de referencia usado y si no está seguro respecto a que el estándar de referencia sea el adecuado, puede que sea necesario que realice una búsqueda más profunda en el área. La sección de Métodos también debería describir quien realizó los dos test y si cada uno fue conducido de manera independiente y ciega a los resultados del otro.
Este artículo: Si <input type="checkbox"/> No <input type="checkbox"/> No está claro <input type="checkbox"/> Comentarios:	

¿Cuáles fueron los resultados?

¿Se presentan las características del test?

Hay dos tipos de respuestas que comúnmente se reportan en estudios de test diagnósticos. Uno apunta a la precisión del test y se refleja en la sensibilidad y especificidad. El otro apunta a como se desempeña el test en la población en la que se prueba y se refleja en los valores predictivos (también llamados probabilidad post-test). Para explorar el significado de estos términos, considere un estudio donde 1000 adultos mayores con sospecha de demencia se someten a una prueba diagnóstica nueva y un estándar de referencia. La prevalencia de demencia en este grupo es de 25%. 240 personas tuvieron resultado positivo en ambas pruebas (la nueva y la referencia) y 600 personas tuvieron resultado negativo en ambas. El primer paso es dibujar una tabla de 2x2 como se muestra más abajo. Se dijo que la prevalencia de demencia era 25% por lo tanto podemos llenar en la última fila de "totales" – 25% de 1000 es 250 – con 250 personas que tendrán demencia y 750 que no tendrán demencia. También sabemos el número de personas con resultados positivos y negativos en ambos test, por lo que podemos llenar dos casillas más en la tabla.

		Estándar de Referencia		
		+	-	
Prueba diagnóstica	+	240		
	-		600	
		250	750	1000

Por sustracción, podemos fácilmente completar la tabla:

		Estándar de Referencia		
		+	-	
Prueba diagnóstica	+	240	150	390
	-	10	600	610
		250	750	1000

Ahora podemos calcular las distintas medidas.

¿Cuál es la medida?	¿Qué significa?
<p>Sensibilidad (Sn) = proporción de personas con la condición y tienen un resultado positivo en el test.</p> <p>En nuestro ejemplo, la $S_n = 240/250 = 0.96$</p>	<p>La sensibilidad nos dice que tan bien la prueba identifica a las personas con la condición. Un test con alta sensibilidad no dejará pasar muchas personas</p> <p>10 personas (4%) con demencia fueron clasificadas falsamente como que no la tenían. Esto significa que el test es bastante bueno en identificar personas con la condición.</p>
<p>Especificidad (Sp) = proporción de personas sin la condición que tienen resultado negativo en el test.</p> <p>En nuestro ejemplo, la $S_p = 600/750 = 0.80$</p>	<p>La especificidad nos dice que tan bien el test identifica a las personas sin la condición. Un test altamente específico no identificará muchas personas falsamente como enfermos</p> <p>150 personas (20%) sin demencia fueron falsamente identificadas como que la tenían. Esto significa que el test es sólo moderadamente bueno en identificar personas sin la condición.</p>
<p>Valor predictivo positivo (VPP) = proporción de personas con un test positivo que tienen la condición</p> <p>En nuestro ejemplo, el $VPP = 240/390 = 0.62$</p>	<p>Esta medida no dice que tan bien se desempeña el test en la población. Depende de la precisión del test (principalmente especificidad) y la prevalencia de la condición.</p> <p>De las 390 personas que tuvieron un resultado positivo, 62% realmente tenía demencia.</p>
<p>Valor predictivo negativo (VPN) = proporción de personas con un test negativo que no tienen la condición.</p> <p>En nuestro ejemplo, el $VPN = 600/610 = 0.98$</p>	<p>Esta medida nos dice que tan bien se desempeña el test en la población. Depende de la precisión del test y la prevalencia de la condición.</p> <p>De las 610 con resultado negativo, 98% no tiene demencia.</p>

Aplicabilidad

¿Los métodos para realizar el test fueron descritos con suficiente detalle para permitir que sean replicados?	
¿Qué es mejor?	¿Donde encuentro esta información?
El artículo debe tener una descripción suficiente del test para permitir su repetición y la interpretación de los resultados	La sección de Métodos debe describir el test en detalle
Este artículo: Si <input type="checkbox"/> No <input type="checkbox"/> No está claro <input type="checkbox"/>	
Comentarios:	