

# Cartas al director



## Sobre un análisis incorrecto de la utilidad del ganglio centinela de la mama

**Sr. Director:**

Me ha llamado mucho la atención la lectura del artículo de Armas et al<sup>1</sup> publicado recientemente en su revista, pues aparece en él un viejo y terrible sesgo que afecta al patrón de referencia, o *gold standard*, que se utiliza en la validación de una prueba diagnóstica concreta, en este caso, la linfadenectomía axilar en la validación de la biopsia del ganglio centinela (GC) en el cáncer de mama. Digo viejo, porque ya hace tiempo que es difícil verlo en publicaciones sobre pruebas diagnósticas, concretamente en una amplia revisión de Whiting et al<sup>2</sup> ya no se encuentra ninguno, y terrible porque invalida absolutamente un estudio que lo contenga.

Se trata del llamado "sesgo de incorporación", que se produce cuando la prueba que sirve de patrón de referencia incorpora o se contamina de alguna manera por el resultado de la prueba que se evalúa. El resultado de la prueba patrón debe ser totalmente independiente de la prueba en evaluación, es más, siempre que sea factible se aconseja incluso que su valoración sea ciega, es decir, que patrón y prueba evaluada sean analizados por personas distintas que no conozcan, además, el veredicto que da cada una de ellas. Tener una buena prueba patrón no es siempre fácil, pero en el caso del GC de la mama es una excepción, porque lo hay y es tan asequible como fiable, nada menos que el análisis patológico de la linfadenectomía axilar. Por eso es una lástima que los autores incurran en este sesgo cuando dicen literalmente: "Se definieron como verdaderos positivos (VP) los casos en que el GC mostró afectación por metástasis... independientemente del resultado de la linfadenectomía, ya que el GC forma parte de la cadena axilar ganglionar y cuando está afectado se considera la linfadenectomía como positiva, aunque el resto de ganglios estén libres de enfermedad".

De este modo, el sesgo de incorporación es aquí indiscutible. Es curioso, además, que los autores parece que pasen por alto que la biopsia del GC se realiza para predecir qué ocurre más allá de él en la axila, esté o no esté localizado en la misma axila, por lo cual ha de haber una línea divisoria, exactamente la que distingue la prueba patrón de la prueba en evaluación. El resultado es que se infla, sobre todo y hasta el máximo, la especificidad, y a resultas, el valor predictivo positivo, ambos del 100% en el estudio, al no dejar este diseño posibilidad alguna de que exista un falso positivo. Ellos mismos vuelven a escribir más adelante, como si tal resultado les resultara sorprendente y lo pretendieran justificar: "Por las características del estudio no existen falsos positivos...". Que me

perdonen la contundencia, pero esto es una falacia. Los falsos positivos existirán o no dependiendo de la bondad o la potencia diagnóstica de una prueba, y para que no los haya, la prueba ha de ser perfecta en ese sentido, algo bastante improbable en la realidad. Por lo tanto, es esa realidad de las cosas y no el diseño de un estudio el que determina su existencia. Aunque menos afectada, tampoco es cierta la sensibilidad que reportan, si bien al final resulta que lo más importante de la evaluación, el valor predictivo negativo, es la única cifra que no está influida por tal sesgo.

En el artículo se ofrecen 2 tablas de resultados. La tabla 1 es la que deberían haber usado para el cálculo de los índices diagnósticos, pues no está afectada por el sesgo. Pero la tabla 2 es producto de ese diseño alterado que introducen a partir del citado sesgo. Consecuentemente, la sensibilidad real no es del 95,6% sino del 88%. La especificidad no es del 100%, sino del 65%. El valor predictivo positivo no es del 100%, sino del 34%. La única cifra bien calculada es, como señalaba antes, el valor predictivo negativo, que efectivamente es del 96,4%.

En otro párrafo se puede leer otra frase poco acertada: "En 29 pacientes sólo fue positivo el GC; el resto de vaciamiento axilar fue negativo, lo que demuestra la fiabilidad de la técnica...". Precisamente estos 29 casos (el 29% del total) son los auténticos falsos positivos de la técnica, cifra relativamente abultada que determina esa especificidad y ese valor predictivo positivo reales bastante mediocres, y que no constituyen, pues, ningún estándar de fiabilidad cuando el GC es positivo. Por un lado, esos valores son bajos como para predecir rotundamente que la positividad del GC implica más ganglios afectados, con lo que habrá linfadenectomías innecesarias y, por otro, tampoco son lo suficientemente bajos como para que, si el GC es positivo, se deba obviar con garantías la linfadenectomía porque haya poca probabilidad de que existan otros ganglios positivos más adentro en la axila. Para esto último, el valor predictivo positivo debería ser menor del 5%, del mismo modo que para no realizar linfadenectomía tras la biopsia negativa del GC se pide un valor predictivo negativo mayor del 95%. Ahí radica verdaderamente uno de los retos pendientes de la biopsia del GC, una vez que ya hay bastante evidencia de que su negatividad es muy fiable: encontrar algún rasgo que también indique con igual seguridad cuándo el GC es el único ganglio afectado.

En resumen, creo que los autores han hecho unos cálculos en falso al caer en el sesgo de incorporación, pero, a la postre, los resultados fundamentales que acreditan su evaluación no están alterados, tanto el porcentaje de identificación del GC que es del 100%, como un excelente valor predictivo negativo. Aunque sea por estas 2 cifras, merecen ser felicitados.

**Javier Escrig-Sos**

Servicio de Cirugía General y Digestiva. Hospital General de Castellón. Castellón. España.

### Bibliografía

1. Armas F, Hernández MJ, Vega V, Gutiérrez I, Jiménez C, Pavcovich M, et al. Utilidad y ventajas de la inyección subareolar única del trazador en la biopsia selectiva del ganglio centinela en el cáncer de mama. *Cir Esp*. 2005;78:260-5.

2. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt P, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy. A systematic review. *Ann Intern Med.* 2004;140:189-202.