

MEDICINA BASADA EN EVIDENCIAS

**The diagnosis of acute appendicitis in a paediatric population:
To CT or not to CT.**

Journal of Pediatric Surgery 2003; 38: 367-371.

Dr. Gaston Duffau T.

Departamento de Pediatría y Cirugía Infantil, Facultad de Medicina, Campus Norte, Universidad de Chile.

El propósito declarado del estudio es determinar si la tomografía predice correctamente el diagnóstico de apendicitis aguda en niños, comparada con el diagnóstico preoperatorio hecho sólo por historia, examen físico y otros estudios de laboratorio.

Tratándose de un estudio de una prueba diagnóstica de apendicitis (tomografía), o mejor dicho, aparentemente de la comparación de dos pruebas diagnósticas, es muy importante establecer cual es el patrón ideal de comparación o "gold standard" por medio del cual se definirá en cada caso la presencia o ausencia de apendicitis aguda. Se debe recordar que este criterio, patrón de diagnóstico, no debe incluir a la prueba diagnóstica que se está valorando.

En el artículo no se establece específicamente cual será el "gold standard", pero parece darse por entendido en el texto que en definitiva sería, y fue, el estudio anatomopatológico puesto que sólo ingresaron a la investigación pacientes apendicectomizados.

En un lapso de 5 años, los autores revisaron a todos los pacientes que fueron sometidos a intervención quirúrgica por supuesta apendicitis aguda. Excluyeron los casos que tuvieron una apendicectomía incidental, como parte de otro procedimiento y aquellos con exploración diagnóstica y apendicectomía por dolor abdominal crónico. Reunieron 283 casos.

En la investigación de una prueba diagnóstica es indispensable que la patología en estudio presente un espectro apropiado, tanto cualitativamente como cuantitativamente. Ello quiere decir que el grupo participante debe incluir las diversas formas de presentación del problema, con casos altamente sospechosos, mediana y escasamente, así como con otras patologías con las cuales se puede confundir la que estamos investigando. Por otra parte, estas formas probables de presentarse el problema mostrarán una proporción similar a la esperable en la

población donde se aplicará la prueba diagnóstica posteriormente.

El estudio será prospectivo y en forma independiente y ciega se aplicará el criterio patrón diagnóstico (gold standard) y la prueba en estudio. Esto se refiere al hecho que si a un grupo de casos que entran al estudio se les practica tomografía (que aún no se sabe si es útil o no lo es para los fines del diagnóstico de apendicitis aguda) su informe por el experto no se debe ser influido por el conocimiento de otros resultados clínicos o de laboratorio que se disponga del caso.

De los 283 casos ingresados al estudio (operados), 268 fueron diagnosticados por anatomía patológica como portadores de apendicitis aguda y en 15 pacientes el órgano estaba normal.

En 96 de los casos se practicó tomografía preoperatoria y se compararon los dos grupos: aquellos con y sin tomografía (ciertamente esperando que no mostraran diferencias en características consideradas basales).

La comparación es, en principio, no válida puesto que por alguna razón a unos se les realizó tomografía y a otros no. (Más apropiado habría sido haber efectuado a todos el examen, o al menos en un grupo elegido al azar.) Por otra parte, los dos conjuntos diferían significativamente en la distribución por género, factor de mucha importancia en la patología estudiada.

Se menciona el uso de la prueba exacta de Fisher (en la comparación de proporciones), pero no se indica como se compararon los promedios de edad y de recuento de leucocitos en sangre. En la tabla respectiva (N°1), no se muestra ningún estadígrafo de dispersión en estos dos casos, de modo que el lector no puede formarse una idea de la adecuación de la comparación, ni comprobar el resultado expresado como valor de "p".

Cinco casos de 96 con tomografía, resultaron erróneamente informados como negativos para apendicitis.

Al considerar género y edad, los autores encuentran que en 81 niñas de 11 años o mayores la precisión diagnóstica, en general, es significativamente menor que la encontrada en todos los demás pacientes, es decir niñas menores y varones de todas las edades.

Tal comparación no es apropiada, porque la diferencia señalada no se informa para cada grupo, de modo que el lector pudiera creer que la precisión diagnóstica es superior en niñas de 10 años o menores así como en varones mayores y en varones menores, aspecto no señalado como estudiado.

El artículo informa que, comparando quienes fueron o no sometidos a tomografía preoperatoria, no se observó diferencia significativa en la exactitud diagnóstica en el total de casos ni

considerando sólo niñas de 11 años de edad y mayores, o los restantes.

En esta última comparación se encuentra sorprendentemente que $p=1,00$, lo que es erróneo: 96,8% ($n=63$) vs. 96,4% ($n=139$) genera $p=0,79$ por Epi6 y $p=0,789$ por Medcalc.

En esta investigación se declara en la discusión: "Este es un estudio retrospectivo no aleatorizado y por lo tanto, tiene sus limitaciones. Incluye sólo aquellos pacientes que fueron apendicectomizados y no incluye un análisis de aquellos pacientes a los que se les efectuó tomografía interpretada como negativa para apendicitis y tratados sin operación. La ausencia de este subgrupo de pacientes impide las determinaciones de especificidad, valor predictivo negativo y la precisión global de la tomografía en la evaluación de apendicitis aguda".

Tabla 1. Guía de revisión para estudios sobre pruebas diagnósticas.

1. ¿Se establece el objetivo de la prueba?	Si
2. ¿Hay un estándar diagnóstico ideal?	No explicitado como tal
3. ¿Hay fundamento para la elección del estándar ideal propuesto?	No comentado
4. ¿Los resultados de la aplicación del estándar ideal son dicotómicos?	Si
5. ¿Se definió el universo y el procedimiento de muestreo?	No
6. ¿Se da una justificación del tamaño muestral?	No
7. ¿Hay evidencias que se encuentra bien representado el espectro de la enfermedad en estudio, tanto cualitativa como cuantitativamente?	No
8. ¿Se incluyeron pacientes con y sin la enfermedad en estudio pero con clínica similar?	No
9. ¿Los casos con y sin la patología de interés son reconocibles previo a la aplicación del estándar ideal?	No
10. ¿Se efectuó una comparación independiente y ciega con el estándar diagnóstico ideal?	No explicitado
11. ¿Se describió el medio donde se desarrolló el estudio y el filtro por el que pasaron los casos?	Parcialmente
12. ¿Se describió la reproducibilidad de la prueba (precisión) y de su interpretación (concordancia inter-observador)?	No
13. ¿Se definió clara y adecuadamente el término "normal"?	No
14. ¿Se presenta la respectiva tabla 2 x 2 o 2 x n con los valores correspondientes que permitan calcular los índices escogidos?	No
15. ¿Se proporcionan los intervalos de confianza de los índices, cuando corresponde?	No
16. Si la prueba es parte de un grupo o secuencia de pruebas, se determinó su contribución a la validez global?	No
17. ¿La prueba está suficientemente descrita como para reproducirla?	Si
18. ¿Se estableció la "utilidad" de la prueba?	No

Por otra parte, no se discute acerca de la prevalencia, es decir, la proporción de casos con apendicitis respecto del total estudiado. En una prueba diagnóstica, tal información es crucial, puesto que los valores predictivos son afectados de manera importante por la prevalencia.

Con limitaciones tan serias, más otras mencionadas previamente en este análisis, es muy probable que muchos hubieran decidido no hacer la revisión.

Ahora, si consideramos el total de aspectos a tomar en cuenta en el estudio de una prueba diagnóstica, veremos lo que ocurre en la situación analizada. Tabla 1.

Lectura recomendada

1. Duffau G. Estándar ideal imperfecto y variable. Rev Chil Pediatr 1996; 67: 125 -126.
2. Duffau G. Tabla de 2 x 2 en el análisis de pruebas diagnósticas por razones de verosimilitud y usos. Rev Chil Pediatr 1997; 68: 194 – 203.
3. Duffau G. Tamaño muestral en estudios sobre pruebas diagnósticas. Rev Chil Pediatr 1998; 69: 122 – 125.
4. Duffau G. Medicina basada en evidencias: Aspectos del análisis crítico de la literatura biomédica. Rev Chil Pediatr 1998; 69: 110 – 115.
5. Herrera P, Duffau G. Errores metodológicos aún en revistas de corriente principal. Rev Med Chile 2001; 129: 464 – 465.
6. Duffau G. Medicina Basada en Evidencias (MBE) y Análisis Crítico de la Literatura. Rev Psiq Clin 2003; 40: 6 – 11.