Aplicación de un protocolo de actuación en el tratamiento de los traumatismos abdominales cerrados: los resultados no son lo que parecen

Sr. Director:

He leído con atención el interesante artículo de la Dra. Lladó et al[1] correspondiente al número de agosto de 2002 de esta revista, en el que se concluye que trabajar con un protocolo de manejo de pacientes con este problema deriva en una mucho menor utilización de la cirugía sin que se aumente, entre otras cosas, la mortalidad final. En mi opinión, este punto de la mortalidad es crítico, puesto que cualquier otro beneficio del tratamiento no operatorio (TNO) no estaría justificado si fuera a costa de aumentar la mortalidad final. Curiosamente, los autores obtienen una mortalidad del 19% con TNO, frente a un 11% en su serie histórica en la que operaban mucho más al no aplicar su protocolo actual. Es un 8% de diferencia, cualitativamente en contra del protocolo, a la que no consideran como importante, es más, afirman literalmente que "la mortalidad es similar".

Ante todo, una diferencia del 8% en mortalidad, aunque no sea una cifra abultada, no es en absoluto despreciable actualmente en ningún aspecto de la patología quirúrgica. Sin embargo, sospecho que quizás los autores no la valoran porque en la inferencia estadística obtienen una p = 0,21, que no llega a ser significativa para el corte habitual del 5%. Si es porque consideran que un

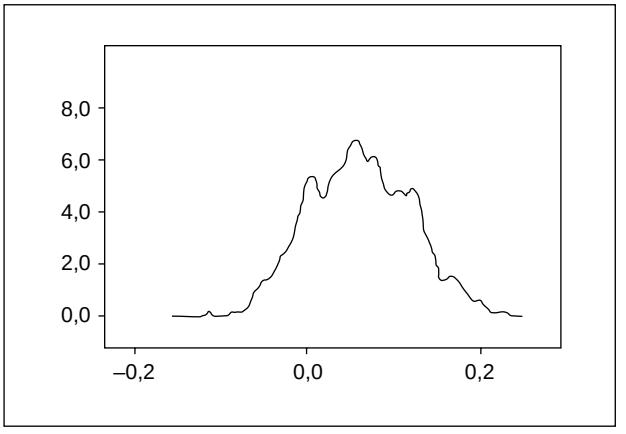

Fig. 1. Resultados de la simulación con WinBUGS.

TABLA 1. **Cálculo de la probabilidad posterior para distintas diferencias de mortalidad**

Diferencias grupos protocolo-histórico	Probabilidad posterior igual o mayor
–0,1	0,99
0,00	0,85
0,02	0,75
0,04	0,64
0,06	0,51
0,08	0,38
0,10	0,27
0,12	0,18
0,14	0,11
0,16	0,06
0,18	0,03
0,20	0,01

8% es despreciable, simplemente creo que nadie puede estar de acuerdo con ellos. Si es por el valor de la p, están cometiendo el error de interpretación más frecuente y de mayor calado que existe en las pruebas de hipótesis de la estadística clásica, y por el que éstas han recibido históricamente el grueso de las más duras críticas.

Efectivamente, si no se puede descartar la hipótesis nula, no se ha demostrado que sea cierta: si ese 8% no es estadísticamente significativo no quiere decir que sea cero. La hipótesis nula, se descarte o no, tiene sus propias probabilidades de ser cierta, igual que la hipótesis alternativa, y desde luego la cuantificación de tales probabilidades no se logra con las pruebas de hipótesis clásicas. De ahí viene la mayoría de equívocos, que como en este caso, pueden cambiar de pies a cabeza las conclusiones de un artículo. Téngase en cuenta que las pruebas de hipótesis clásicas son esclavas del tamaño de la muestra, de modo que esa misma diferencia del 8% con una muestra cuatro veces mayor hubiera resultado ser estadísticamente significativa, y las conclusiones distintas radicalmente. Por eso se ha de ser muy cauto al interpretar una $p > 0,05$.

Con los datos del artículo me he permitido hacer un análisis diferente: he escrutado la hipótesis alternativa según las técnicas de simulación bayesiana, tal como se describe en un reciente artículo de divulgación[2] escrito, además, en castellano. Si en el grupo TNO hubo un 19% de muertes sobre 65 casos, deduzco que murieron 12 pacientes y sobrevivieron 53. En el grupo histórico sin protocolo, con un 11% de muertes sobre 77 casos, deduzco que murieron unos 9 pacientes y sobrevivieron 66. Sin tomar ninguna idea preconcebida sobre en qué grupo es más probable la mortalidad (matemáticamente aplico una distribución uniforme a priori para ambos grupos), la distribución posterior para el grupo TNO será una beta (13,54) y para el grupo histórico otra beta (10,67). Una simulación sobre 1.000 casos de la diferencia entre ambos grupos se puede hacer con el programa MINITAB®, o con el llamado WinBUGS, que se consigue de forma gratuita en Internet. Los resultados son iguales en ambos (fig. 1). En la tabla 1 se muestra una serie de posibles diferencias de mortalidad entre el grupo TNO y el histórico, junto con las probabilidades finales de que dichas diferencias sean “iguales o mayores”, de acuerdo con la gráfica de la figura 1. Estas probabilidades son de interpretación directa, al contrario de lo que ocurre con la p.

Así pues, que haya más mortalidad en el grupo TNO es un 85% probable, frente al 15% probable en sentido inverso, siempre de acuerdo con los datos del estudio, por lo cual en el grupo protocolizado hay hasta unas 5,7 veces más probabilidades de morir. Y por poner otro ejemplo, que esa diferencia entre ambos grupos iguale o supere el 10% tiene la nada despreciable probabilidad del 27%, y por el contrario, que ese 10% de diferencia fuera cualitativamente a favor del grupo protocolizado tiene sólo un 1% (tabla 1).

En fin, los resultados no son pues lo que parecen, y son desde luego extraños, porque no se acoplan a lo que sería lógico, supuesto que trabajar con un buen protocolo tenga que dar por fuerza mejores resultados incluso en mortalidad, al menos a medio-largo plazo. Es posible que la muestra pudiera contener sesgos al tratarse de una serie de casos. En ese contexto, pruebas bivariantes como las de χ^2, Mann-Whitney, entre otras, no tienen capacidad de control de ciertos sesgos, y por lo tanto, no sirven más que como herramientas de análisis preliminar, de modo que hubiera sido conveniente acompañarlas de algún tipo de análisis multivariante. Quizás así se hubiese controlado mejor la presencia de algún efecto perturbador, como por ejemplo, el derivado de que en el grupo protocolizado (TNO) hubo una proporción bastante mayor de lesiones asociadas. Pero de nuevo, como la diferencia para eso tampoco fue estadísticamente significativa, los autores consideraron que la presencia de lesiones asociadas era equivalente. Otra vez ronda el enorme error de interpretación antes mencionado.

Sea lo que sea, y por más que en la inferencia clásica no se hayan encontrado diferencias significativas, de lo que no me queda duda es que desde los datos aportados en el estudio no se puede afirmar que la mortalidad real entre ambos grupos parezca similar, por lo tanto, las conclusiones dadas en dicho estudio no son correctas en lo que a ella se refiere, y las repercusiones podrían ser muy distintas de lo que se desprende de la lectura de éste.

V. Javier Escrig

Servicio de Cirugía General y del Aparato Digestivo.
Hospital General de Castellón. Castellón. España.

Bibliografía

1. Lladó L, Jorba R, Parés D, Borobia FG, Biondo S, Farran L, et al. Influencia de la aplicación de un protocolo de actuación en el tratamiento de los traumatismos abdominales cerrados. Cir Esp 2002; 72:79-83.
2. Silva LC, Benavides A. El enfoque bayesiano: otra manera de inferir. Gac Sanit 2001;15:341-6.