

Comparaciones entre cirugía convencional y laparoscópica. Ha llegado el momento de los estudios de equivalencia

Ya no constituyen una novedad para ningún cirujano las ventajas que, en general, ofrece la cirugía laparoscópica. Por ejemplo, sabemos que la inmensa mayoría de las intervenciones convencionales se pueden realizar de esta manera, aunque en algunos casos aún se roce la pura exhibición, que se necesita un período de aprendizaje, que se suele acortar casi siempre el período postoperatorio, que el consumo analgésico es menor, etc. Mucho se ha estudiado y publicado al respecto, casi siempre enfocando el correspondiente estudio o el simple análisis de una experiencia desde el punto de vista de la superioridad, es decir, planteando que la cirugía laparoscópica y la convencional no son iguales para ciertas variables, como las referidas de la estancia y el dolor. Como señal de evidencia, en la bibliografía se recurre de manera habitual a las pruebas clásicas de inferencia estadística, sin duda con la intención de conseguir un resultado que sea estadísticamente significativo ($p < 0,05$). Esto se logra repetidamente al manejar variables como las ya mencionadas del dolor o de las estancias, y también con los períodos de convalecencia, a tenor de lo que de forma ya muy repetitiva se lee en los artículos de CIRUGÍA ESPAÑOLA y de cualquier otra revista especializada.

Las pruebas de inferencia estadística clásicas son una buena herramienta para demostrar superioridad, entre otras cosas, porque fueron diseñadas para ello. Parten de una hipótesis llamada nula (H_0), algo idealizada en su concepción, pues significa la igualdad absoluta y, si es posible según el valor de la p , se llega a su propio rechazo y la subsiguiente aceptación de la denominada hipótesis alternativa (H_a), que en cierta forma corrobora la existencia de superioridad —o inferioridad— entre las variables contrastadas, para finalmente establecer que este hallazgo es plausible, por lo menos en tal magnitud, en la generalidad de casos, entiéndase población. El problema surge cuando no es posible rechazar la H_0 , porque entonces no se puede llegar a conclusión alguna sobre el análisis realizado. Hay otro problema, quizá mayor en esa circunstancia y que tal vez derive de la propia concepción mecánica de las pruebas, y es el hecho de que intuitivamente tendamos a considerar que el incidente de no poder rechazar la H_0 sugiere que se ha demostrado igualdad o equivalencia entre las variables analizadas, cuando en realidad nada más lejos de ello; se trata de una interpretación absolutamente errónea. Como decían

dos conocidos bioestadísticos ingleses¹: “la ausencia de evidencia no es evidencia de la ausencia”, comentario éste que también se puede leer en revistas médicas españolas de gran impacto².

El caso es que, sabiendo lo mucho que sabemos sobre las bondades repetidamente demostradas de la cirugía laparoscópica, quedan algunos puntos críticos por resolver cuando se la compara con la cirugía convencional: precisamente suelen ser variables de gran importancia para la integridad del enfermo o para el resultado intrínseco de la propia intervención quirúrgica, en definitiva, para la acreditación de la técnica. Ejemplos hay muchos. El más llamativo está en las dudas sobre los resultados oncológicos de la cirugía laparoscópica del cáncer colorrectal, que se pueden extender a otros cánceres digestivos, pero igualmente se podría mencionar el tema de las complicaciones de cada técnica laparoscópica con respecto a la convencional, las recidivas herniarias a medio o largo plazo, los resultados tardíos de la cirugía antirreflujo, las repercusiones sobre el esfínter biliar de la dilatación del mismo en ciertas modalidades del abordaje laparoscópico de la coledocolitiasis, y un largo etcétera. Por desgracia, precisamente en estos temas cruciales es donde más cantidad de resultados estadísticamente no significativos se suelen observar en las distintas publicaciones, de modo que la evidencia de superioridad, y también la de equivalencia, sigue en el aire para la mayoría de estos puntos cruciales, por más que en muchas oportunidades los autores de estas publicaciones —y mucho nos tememos que bastantes lectores— se deslicen subrepticamente, en sus comentarios o en su lectura, hacia aquella interpretación tan intuitiva como errónea antes mencionada, si es que no caen abiertamente en este error.

Los seguidores acérrimos del paradigma bayesiano achacan ciertos males como éste a la propia naturaleza, algo retorcida y manifiestamente insuficiente para determinados usos, del tipo de razonamiento probabilístico que contienen las pruebas de inferencia clásica que utilizamos de manera habitual. Posiblemente sea ésta una visión exagerada, propia y común de quien es fanático seguidor de una idea o de una tendencia. Pero es bastante verosímil que el problema sea más profundo y vaya más allá de la propia naturaleza de las pruebas estadísticas. Dicho de otro modo, ¿no será cierto que, en reali-

dad, la cirugía laparoscópica no es ni mejor ni peor que la convencional para muchas de esas variables críticas? ¿No será menos cierto que lo que en realidad nos gustaría demostrar sea precisamente esto, y no que la cirugía laparoscópica sea mejor que la convencional, como en verdad lo es, para otro tipo de variables como el dolor o la estancia postoperatoria? Si aceptamos que esto pueda ser así, y pasada una primera fase histórica de los estudios sobre estas cuestiones, se podría argüir que, ciertamente, no se está haciendo, hoy por hoy, un enfoque totalmente correcto de las investigaciones, idea que se puede extender a las pruebas clásicas de inferencia porque, seguramente, con ellas no llegaremos a ninguna conclusión clara ni aún, si se me apura, ni siquiera con grandes ensayos clínicos cuyo diseño esté basado en demostrar superioridad –desigualdad– en lugar de equivalencia.

A los cirujanos se nos acusa de ser poco científicos a la hora de evaluar y aceptar las novedades que aparecen en nuestra especialidad, y la cirugía laparoscópica tiene entendidos detractores que sólo podrán ser convencidos de su hipotético error con argumentos científicos y evidencias bien fundadas en un método riguroso. Recientemente, hemos oído en persona el lamento público en este sentido de conocidos laparoscopistas, como el Dr. Azagra, por no haber transmitido sus convicciones más desde la medicina basada en la evidencia que desde pronunciamientos de otro tipo. Así pues, el proceso de investigación clínica que encuentre el lugar que haya de tener la cirugía laparoscópica necesita salir de su estancamiento actual y caminar abiertamente a la resolución de esas dudas fundamentales que se han mencionado. Presentaciones en forma de vídeos, descripciones de series personales, estudios que usan pruebas estadísticas sólo útiles para demostrar desigualdad y que sólo nos llevan a la indefinición, de esto ya hemos visto y leído mucho. Es posible que todos estos esfuerzos, por otra parte absolutamente legítimos, procedan de amplios grupos de estudio o de humildes centros comarcales, rendirían científicamente mucho más si se enfocaran desde el punto de vista de demostrar la equivalencia entre aquellas variables críticas si es que, como parece, va a ser muy difícil o imposible demostrar superioridad en estos aspectos.

Los ensayos clínicos de superioridad fueron concebidos para confirmar diferencias no espectaculares, ya que cuando lo son en algún campo, suelen percibirse con prontitud e, incluso, hacen que no sea ético plantearse este tipo de estudio. A nadie se le ocurriría en la actualidad diseñar un amplio ensayo clínico para comparar la colecistectomía electiva laparoscópica frente a la convencional, aunque quede aún pendiente conocer más detalles sobre el problema de las lesiones yatrogénicas de la vía biliar. Por su parte, los ensayos clínicos de equivalencia, que también provienen del entorno de la farmacología, son mucho menos conocidos y, aunque tienen ciertos condicionantes éticos³, se han multiplicado en la última década. Persiguen demostrar equivalencias de acción entre fármacos nuevos y antiguos cuya diferencia no radica, pues, en su efecto intrínseco sobre la dolencia, sino en su comodidad de administración, facilidad de cumplimiento de la prescripción, un

menor coste y reducción de algún efecto adverso, en ocasiones muy colateral. Pero antes de plantearlos es necesaria la condición previa de que no se haya podido demostrar entre ellos, mediante estudios de superioridad, una diferencia en su acción principal que fuera clínicamente importante, ni en el sentido del beneficio ni del daño, y que se haya demostrado alguna ventaja independiente o accesoria de lo que es su finalidad primordial. El estado actual de la cirugía laparoscópica es un fiel reflejo de esta situación: está claro que, en general, no es más dañina en sí misma que la cirugía convencional, y tiene ciertas ventajas colaterales bien acreditadas que la harían enormemente atractiva si se demostrara que, por lo menos, es igual que la cirugía convencional para lo que es el resultado principal del acto quirúrgico: resolver la dolencia contra la que se dirige sin producir más complicaciones. Tampoco es necesario que sea mejor, aunque en algunos campos lo pueda ser. La mayoría de las veces sería suficiente con ser igual en lo fundamental, y tener otras ventajas añadidas, para que pudiese ser declarada como técnica de elección, pero esa igualdad debe ser confirmada científicamente; no basta con no poder confirmar la desigualdad.

Aparte de condicionantes éticos, los ensayos de equivalencia necesitan de un planteamiento metodológico si cabe más riguroso que los de superioridad, y de un enfoque estadístico también diferente^{4,5}. Aquí, las escasas pruebas de inferencia que existen, diseñadas específicamente para este fin, al contrario que en los enfoques de superioridad, formalmente parten de la H_0 para intentar rechazarla y aceptar la H_1 . No vale, pues, aplicar pruebas de inferencia estadística como las conocemos y manejamos habitualmente, al menos de forma directa como lo haríamos en los enfoques de superioridad. No obstante, para los problemas que con más frecuencia nos plantearía la investigación de equivalencia dentro de los aspectos fundamentales en las comparaciones entre cirugía laparoscópica y convencional, hay una forma de aprovechar la información de las pruebas clásicas usando el intervalo de confianza⁵ de la diferencia observada entre las variables que se contrasten, junto con la condición de que se establezca, previamente y de forma clara y meridiana, el valor del margen de equivalencia entre esas variables. Se precisa, pues, establecer unos límites que acoten un margen de valores simétricos por arriba y por debajo de la diferencia igual a cero. Mucho más exigente de establecer, pero más concluyente en cuanto a la interpretación práctica, es el llamado límite de no inferioridad, que acota un margen en una sola dirección a partir de la diferencia cero. Un tipo u otro de márgenes serán los que se deben confrontar con el intervalo de confianza de dicha diferencia, de modo que habrá equivalencia sólo si aquel intervalo de confianza queda englobado en el margen preestablecido. Para fijar estos límites no nos podemos ayudar de la estadística, sino del conocimiento del problema, de la lectura crítica de la bibliografía y del razonamiento y la discusión. Incluso puede que se necesitara, en ciertos temas, una conferencia de consenso entre expertos para otorgar mayor consistencia y credibilidad al ensayo⁶. Como desafío intelectual resulta, pues, tanto o más interesante que buscar supe-

rioridad. Otro problema, a veces no precisamente menor, de este tipo de enfoque es que se suelen requerir muestras más amplias que las que precisaría un estudio de superioridad, generalmente entre un 10 y un 30% mayores, como escudo ante los errores de tipo I y especialmente de tipo II, si bien esto no debería ser un obstáculo insalvable cuando un diseño de equivalencia es lo único pertinente.

En resumen, y si a corto plazo no se producen por la vía de la superioridad hallazgos tan contundentes como poco previsibles, los cirujanos que gustan de investigar científicamente los resultados de la cirugía laparoscópica harían bien en dotarse de los conocimientos necesarios para diseñar ensayos de equivalencia y llevarlos a cabo, si es que queremos superar la fase descriptiva o incluso exhibicionista de dicha cirugía, huir de la vaguedad y adentrarnos definitivamente en el mundo de las evidencias científicas. El diseño, monitorización e interpretación de resultados de esta clase de ensayo suele ser algo más complejo que cuando se busca superioridad⁶, pero quizás sea la única solución para aclarar mu-

chas de las dudas que aún planean sobre la cirugía laparoscópica.

V. Javier Escrig Sos

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

Bibliografía

1. Altman DG, Bland JM. Absence of evidence is not evidence of absence. *BMJ* 1995;311:485.
2. Argimon JM. La ausencia de significación estadística en un ensayo clínico no significa equivalencia terapéutica. *Med Clin (Barc)* 2002; 118:701-3.
3. Djulbegovic B, Clarke M. Scientific and ethical issues in equivalence trials. *JAMA* 2001;285:1206-8.
4. Green WL, Concato J, Feinstein AR. Claims of equivalence in medical research: are they supported by the evidence? *Ann Intern Med* 2000;132:715-22.
5. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous method. *BMJ* 1996;313:36-9.
6. Ebbutt AF, Frith L. Practical issues in equivalence trials. *Statist Med* 1998;17:1691-701.