

ARTÍCULO ESPECIAL

La revisión por pares en las revistas científicas

J. Gervas y M. Pérez Fernández

Equipo CESCA. Madrid.

Introducción

Las motivaciones de los científicos son múltiples y complejas, pues incluyen cuestiones tan dispares como la curiosidad y la posición social; entre ellas, el deseo de ser reconocido por los pares es una motivación intensa y permanente, a veces inocente, a veces interesada (por el impacto en la posición social y en la obtención de recursos). Lograr la aceptación de las propias ideas por otros investigadores es conseguir uno de los mejores premios que puede dar el mundo científico¹; es pasar del microcosmo individual, siempre amenazado por la subjetividad, al sólido macrocosmos del conocimiento científico de la comunidad científica, que da sentido a la vida, al trabajo y a las ideas del investigador.

Los científicos comparten sus ideas y resultados con los pares de manera formal e informal, desde comentarios de pasillo a charlas, conferencias y publicaciones. Este compartir permite hacer público el trabajo personal y ayuda a mejorarlo, en un proceso continuo de crítica a superar (desacuerdo creativo); además, introduce enormes dosis de autocrítica, pues el investigador sabe que sus colegas suelen ser escépticos, renuentes a cambiar las teorías establecidas, y no es fácil convencer a los que tienen otras ideas, a veces opuestas a las propias.

La publicación en una revista científica, una manera formal de

compartir ideas y de conseguir la aceptación de las propias, obliga a cumplir ciertos requisitos de calidad, particularmente la revisión por pares, en el curso de la cual se pueden detectar omisiones, errores y explicaciones alternativas. La revisión por pares se ha convertido en la norma y, de hecho, caracteriza a las revistas científicas dentro del universo de las publicaciones. A la propia revisión por pares se le añade, generalmente, un cierto grado de anonimato, bien de los revisores, bien de los autores, en un afán de conseguir, respectivamente, libertad de crítica y equanimidad. Pero, ¿tiene fundamento científico la revisión por pares de los trabajos enviados para publicación a las revistas científicas? Y, si lo tiene, ¿añade algo el anonimato de revisores o de autores? Intentaremos responder a estas dos preguntas utilizando argumentos y propuestas que contribuyan a crear el debate que intentó generar en España, hace años, el director de una revista de salud pública².

Fundamento teórico y ventajas de la revisión por pares

Las revistas científicas no tienen capacidad para publicar todos los manuscritos que reciben, pues el número de páginas disponibles (oferta) está siempre por debajo de los deseos/necesidades de publicación de los autores (demanda)³. Además, los lectores no desean/pueden enfrentarse a la lectura de la publicación indiscriminada de todos los manuscritos y esperan

que sólo se publiquen aquellos que puedan tener interés potencial. La situación es distinta según se consideren las revistas científicas que se dedican a las ciencias básicas, cuyos lectores suelen ser pares de los autores, o las revistas científicas que se dedican a las ciencias clínicas, cuyos lectores mayoritarios suelen ser médicos que trabajan con pacientes⁴; las primeras sirven para intercambiar ideas, para hacer progresar la ciencia, y los pares-lectores tienen enorme capacidad crítica para leer, sin muchos filtros, los artículos. Las segundas, las revistas científicas clínicas, añaden al intercambio de ideas un importantísimo componente educacional, de difusión de pautas clínicas aplicables directamente en la práctica diaria, y muchos de sus lectores carecen de capacidad crítica suficiente para discriminar la valía científica de los métodos empleados en los artículos publicados. Así, las revistas científicas clínicas (y las de ciencias básicas) establecen filtros que permiten ajustar la demanda a la oferta, respecto a los manuscritos que reciben, al tiempo que aseguran a los lectores calidad e interés de lo publicado. En ello hay intereses científicos y financieros, pues las revistas se sostienen gracias a los recursos que atraen —suscripciones y publicidad—, sin los que no se mantendrían en el mercado. En este juego de intereses, los autores aceptan el precio —el filtro de la evaluación y la lista de espera para la publicación de los artículos aceptados— porque la comunicación formal de sus ideas y trabajos a través de las revistas científicas

Correo electrónico: jgervasc@meditex.es

(Aten Primaria 2001; 27: 432-439)

es un poderoso estímulo para su carrera, tanto en lo personal como en lo profesional (posición social y académica, justificación de recursos obtenidos, base para obtención de nuevas ayudas, y otros).

Los filtros de las revistas son múltiples, desde las normas de publicación al lenguaje en que admiten manuscritos, la primera valoración del comité editorial (expresión de una política editorial concreta) y la revisión por pares (la evaluación propiamente dicha). Los filtros, especialmente la revisión por pares, permiten: a) seleccionar el material a publicar, eligiendo lo más innovador, realizado con el máximo rigor científico y con el mayor impacto potencial sobre los lectores (y, en último término, sobre la salud de los pacientes); b) mejorar la presentación del trabajo —su claridad científica y lingüística—, eliminar errores, evitar omisiones, introducir lógica interna, considerar explicaciones alternativas, mejorar las referencias a trabajos previamente publicados, hacerlo atractivo y destacar su utilidad; c) disminuir la posibilidad de publicar artículos fraudulentos, evitar la doble publicación y valorar los aspectos éticos del trabajo realizado y de su publicación, y d) generar un ambiente de calidad y neutralidad científica, de credibilidad de lo publicado y de la revista en general, que se logra a través del consenso en torno a la consistencia interna, la posibilidad de repetir el estudio y la aceptación de críticas del trabajo a publicar⁵⁻¹². En este sentido, si se tiene en cuenta que la decisión final de publicar un artículo recae en el director de la revista, la evaluación por pares le ayuda a tomar una decisión juiciosa, más allá de la propia valoración del comité editorial.

El filtro, la evaluación por pares, atempera el deseo de publicar de los autores y les fuerza a realizar cambios que no habían introducido previamente por descuido, ig-

norancia o, incluso, mala fe. Los cambios, aunque se discutan, se suelen aceptar por su racionalidad y/o porque compensa la modificación si se obtiene la aceptación para la publicación del manuscrito modificado; pocos autores introducirían cambios si no se les forzase¹³. La revisión tiene un papel educativo en el que el evaluador comparte ideas propias con el autor —sobre todo cuando discrepa—, en este sentido, es un poderoso instrumento de avance de la ciencia^{12,14}. El evaluador concuerda, corrige y/o discrepa del autor, le ayuda a mejorar la presentación del trabajo, le defiende de críticas *ex post*, solicita aclaraciones y facilita el trabajo al lector. En cierta forma, la revisión es una defensa del lector, no sólo un ejercicio educativo para el autor y el evaluador. La revisión es, pues, al tiempo, una defensa del autor (contra sus propios errores) y del lector (contra los trabajos sin calidad) y una ayuda en la toma de decisión del director.

En la selección de trabajos para su subvención, en las convocatorias de ayudas para trabajos de investigación, se utiliza también la revisión por pares. La oferta no es páginas aquí, sino dinero en forma de ayudas o subvenciones; la demanda no es manuscritos para publicar, sino proyectos para financiar; los lectores a defender forman la población que puede beneficiarse de los resultados, y el director y el comité editorial son, respectivamente, el presidente y el comité asesor de la agencia financiadora. Por lo demás, el fundamento teórico y las ventajas son las mismas, y ambos procesos requieren una evaluación apropiada^{6,15}.

Fundamento teórico y ventajas del anónimo en la revisión por pares

En el proceso de revisión por pares suele aplicarse el anónimo del revisor (el autor no sabe quién le

ha evaluado, pero el revisor sabe quién firma el artículo; es el anónimo simple), aunque algunas revistas van más allá y también hacen a los manuscritos anónimos (ni el autor sabe quién le corrige, ni el revisor sabe a quién evalúa; es el anónimo doble). El anónimo simple es lo más habitual; se utiliza, por ejemplo, en *British Medical Journal*, *New England Journal of Medicine*, *JAMA*, *Annals of Internal Medicine*, *Obstetrics and Gynecology* y otras¹⁶; el anónimo doble se utiliza, según sabemos de ciencia propia, en las revistas *Atención Primaria*, *Gaceta Sanitaria*, *Medicina Clínica*, *Revista Española de Epidemiología* y en el *European Journal of General Practice*; también en *Annals of Emergency Medicine*, *Epidemiology* y *Journal of the American Geriatrics Society*¹⁶. El fundamento teórico del anónimo simple, y en general del anónimo del evaluador, es el aumento de libertad para la crítica. El evaluador anónimo goza de una libertad de crítica que le ayuda a: a) expresar directamente su opinión, con cortesía pero sin rodeos; b) evitar litigios y discusiones personales; c) solventar conflictos de autoridad (especialmente cuando un inferior juzga a un superior en categoría/posición académica); d) elaborar la crítica de forma que deje libertad al director para tomar la decisión final (el director goza, muchas veces, de menos autoridad científica que el evaluador, pero debe decidir y aceptar la responsabilidad de la decisión con absoluta independencia); e) facilitar la participación de miembros destacados de la comunidad científica, que rechazarían el penoso trabajo de la revisión si se hiciese en forma nominal; f) facilitar la crítica *ex ante*, y en privado, que con franqueza y tacto puede librar al autor, sin sonrojo, de la más dura y pública crítica *ex post*, y g) evitar la aceptación de trabajos de calidad dudosa, al obviar la crítica científica constructiva y el «no», a veces, para lograr

una clientela de autores aceptados agradecidos^{2,5,6,8,14}.

En todas las revistas el revisor suele perder colectivamente el anónimo cuando se publica la lista de revisores que han evaluado los artículos a lo largo del año. Esta lista es la garantía última (el crédito) de que la revista trata de garantizar la calidad de los artículos publicados; por ello, suele incluir a la flor y nata de la especialidad/campo que cubra la revista, los pesos pesados que cuentan con experiencia científica y son aceptados por la comunidad respectiva^{2,5,7,10-12,14,17,18}. Si se ha seleccionado bien, la lista debe incluir autores de valía de muy distinto origen, para poder rechazar la crítica fácil de que se constituye con miembros conspicuos de la «torre de marfil»¹². Actúan con informes anónimos, pero suelen estar orgullosos de pertenecer al grupo de evaluadores que se listan, y de ver incluido su nombre junto a los pares que estiman (lo que es una forma más de aceptación por la comunidad científica; la otra, el que muchos evaluadores son también autores que publican^{5-7,12}).

El fundamento teórico del anónimo doble es añadir a la libertad de crítica, que proporciona el anónimo simple, la ecuanimidad, la imparcialidad serena de juicio; no se trata de ser objetivos, sino de lograr que la subjetividad no tenga sesgos². La evaluación por pares con anónimo doble ayuda a: a) evitar el sesgo de juzgar benévolamente los manuscritos de instituciones/autores reconocidos y acerbamente los de instituciones/autores desconocidos/noveles; b) aconsejar y corregir manteniendo la adecuada «distancia terapéutica»; c) solventar los posibles conflictos de interés, y d) poner en pie de igualdad, respecto al anónimo, al evaluador y al autor^{2,7,10,14}.

El autor que se sabe anónimo recibe el informe del revisor con sensación de alivio, pues sabe que

sólo se juzga el trabajo que se describe en el manuscrito. Si hay acrimonia se deberá al carácter del evaluador, no a un ajuste de cuentas. El anónimo protege, especialmente, al autor novel y al que presenta ideas y trabajos arriesgados e innovadores. Gracias al anónimo pueden desarrollarse trabajos e ideas que no coincidan con líneas previas, sin temor a ser criticado por ello antes de su publicación.

Como respecto a la revisión por pares, todo lo referente al anónimo simple se aplica a las convocatorias de ayudas económicas y a la evaluación de los proyectos que se presentan para obtener fondos. El anónimo doble no tiene cabida en este campo, pues uno de los criterios clave para la concesión de la ayuda es el trabajo previo y la solidez del equipo investigador que presenta el proyecto.

Experimentación y crítica a la revisión por pares

Como en otros campos^{19,20}, la experimentación contradice a la observación respecto a la revisión por pares. Pese a una historia de más de 300 años de revisión por pares en las publicaciones y concesión de ayudas^{3,14}, hace apenas una década que se ha introducido la metodología científica en la valoración de la eficacia y validez de la revisión por pares²¹. Las voces de alarma, no obstante, se empezaron a oír hace más de 30 años⁵⁻⁷; no es raro, pues, que sean muchas las revistas que están ensayando métodos que cambien, más o menos radicalmente, la revisión por pares^{11,13,14,22-25}, aunque otras se muestran satisfechas con el sistema habitual^{2,10,12,26}. Por supuesto, la inquietud también se extiende a los organismos que conceden ayudas para investigación^{15,27}. A la propia virtud no probada (o escasamente demostrada) de la revisión por pares se añade el reto de las nuevas tecnologías, especialmente el papel de

la revisión en Internet^{11,13,22,24}, pero esa cuestión se escapa a este análisis, aunque implique cambios¹¹ que pueden llevar a la desaparición de las revistas tal como las conocemos^{4,11,28,29}.

¿Qué sabemos de la revisión por pares a través de la experimentación? Los fundamentos teóricos son más sólidos que los experimentales² y están más demostrados sus efectos perjudiciales que sus beneficios experimentales²¹. Entre los estudios realizados podemos destacar:

1. Los que se refieren a la concordancia entre evaluadores, cuestión estudiada en primer lugar por sociólogos y psicólogos, en revistas como *Journal of Personality and Social Psychology*, *Sociological Quarterly* y *Personality and Social Psychology*^{6,7}, cuyo estudio clásico, en ciencias médicas, se hizo en el *New England Journal of Medicine*, en 1974⁵. Ingelfinger, director de la revista en aquel entonces, sometió 496 manuscritos consecutivos a parejas de evaluadores, con el propósito de comparar su recomendación final; hubo acuerdo perfecto en un 41,8% de los casos, ligeramente por encima del azar; el acuerdo global fue del 30%. Este «buen» resultado se debió, sobre todo, al amplio acuerdo acerca de los manuscritos de pésima calidad, que constituyeron un cuarto del total, pues los malos artículos se reconocieron con facilidad. Incidentalmente, el desacuerdo entre evaluadores justifica la remisión consecutiva del mismo manuscrito, tras ser rechazado, hasta encontrar una revista que lo acepte, a veces de más categoría que las previas; esta práctica es habitual en el mundo científico^{7,10,30,31}; por propia experiencia conozco (JJ) el caso de un artículo que evalué para Medicina Clínica, recomendando justificadamente su rechazo, y que posteriormente fue publicado en *Gaceta Sanitaria*, casi sin cambios (no se incluye la cita para respetar la confidencialidad del proceso de revisión).

2. Los que han estudiado directamente el proceso de evaluación, replicándolo experimentalmente (y rozando los límites de lo ético). El estudio clásico es el de Peters y Ceci, publicado en 1982⁷, en el que se seleccionaron 12 artículos excelentes, publicados en otras tantas revistas de primera línea por autores famosos, de instituciones prestigiosas y con gran factor de impacto. Los artículos se reescribieron, cambiando sólo los autores y las instituciones en que se habían realizado (se pusieron nombres inventados), y se enviaron, a los 2 años de su publicación, a las mismas revistas que los habían publicado, sin advertir del engaño, y sabiendo que la revisión en ellas era de anónimo simple. En 3 revistas se advirtió la previa publicación, en 9 publicaciones no se advirtió el engaño y los artículos pasaron de nuevo la revisión por pares. Sólo uno fue admitido de nuevo; los 8 restantes fueron rechazados por «problemas metodológicos», no por ser cuestión ya conocida/publicada.

3. Ha habido revistas que carecieron de revisión por pares durante años, de los cuarenta a los setenta del pasado siglo, y sin embargo lograron un crédito que perdura; entre ellas *The American Journal of Medicine* y *The Lancet*^{3,5}. En el primer caso el director y fundador, Gutman, filtraba los manuscritos y decidía por sí mismo sobre un 95% de lo publicado; en el segundo era el comité editorial el que decidía, y lo hacía sin evaluadores externos en un 90% de los casos.

4. Los que han analizado el resultado de la evaluación a través de la calidad del manuscrito, valorando el texto antes y después de su modificación según las sugerencias de los evaluadores. Tanto en una publicación holandesa⁹, como en una australiana¹³, aun con las dificultades de la indefinición de «calidad» respecto al texto del manuscrito, se estimó que la revisión por pares había mejorado

la presentación, legibilidad y consistencia interna de los textos. En una revista estadounidense, la mejoría fue cierta pero sin importancia³², y los lectores echaron en falta, más que «calidad», relevancia clínica³³.

5. Los que han medido la capacidad de los revisores para advertir errores en el manuscrito, como el del *British Medical Journal*, publicado en 1998³⁴. Con el permiso del autor de un artículo admitido para su publicación, se introdujeron 8 errores voluntarios en el diseño, análisis e interpretación, en el manuscrito original. El texto modificado se envió para su evaluación a 420 revisores habituales de la revista. Se recibieron informes de 221 evaluadores; la media de errores detectados fue de 2; un 10% identificó 4 o más errores, y el 16% no identificó ninguno.

6. Los que revisan los graves errores históricos que ha provocado la revisión por pares, en el experimento natural que es su propia existencia. La galería va desde la publicación sistemática de artículos en el *New England Journal of Medicine* con errores estadísticos⁵ hasta la publicación de artículos con datos inventados y falsos, en todo o en parte, tanto por autores de prestigio como por autores desconocidos (que con este motivo se hicieron famosos)^{35,36}. Constituyen parte de «la población del fraude», quizá sólo una pequeña fracción (la que se ha detectado), Summerlin, Long, Alsabti, Soiman, Strauss, Darsee, Slutsky, Berger, Baltimore, Gallo, Pearce, Anderton, Nixon y otros³⁵⁻³⁷, entre los que no hay autores españoles. Error histórico ha sido, también, el rechazo de manuscritos con ideas nuevas que han tenido después enorme impacto³⁸, como el de Jenner sobre la vacunación contra la varicela, en 1876, rechazado por *Philosophical Transactions of the Royal Society* y, cien años después, otro error en estudios sobre inmunología que llevaron a obtener, poste-

riormente, el Premio Nobel^{14,38}, y que dio pie a que el autor empezara el discurso de recepción del premio con diapositivas de las cartas de las prestigiosas revistas que rechazaron su publicación²³. Muy ilustrativo es el caso del ensayo clínico canadiense (de la Universidad McGill, de Montreal) sobre la episiotomía, al que se ha dedicado un artículo con el relato de las dificultades de financiación y de publicación, en el que se incluyen los comentarios de evaluadores y directores de las revistas, y el tortuoso camino que se siguió para lograr su difusión³⁰.

7. Los que analizan la calidad de los informes de los evaluadores. Aunque se desconoce el proceso formal —si existe— que lleva a la elección de los revisores de las distintas revistas, suelen elegirse autores de prestigio, bien conocidos y aceptados^{2,5,7,10-12,14,17,18}. Si eso implica la selección de evaluadores maduros, de edad, se ha demostrado que son los revisores jóvenes que han publicado recientemente los que tienen mejor juicio y más capacidad de detectar errores en los manuscritos y de hacer mejores informes^{34,39}.

El fundamento experimental de la revisión por pares es débil, lo que dificulta su impacto en la publicación/financiación de proyectos y estudios cuya base experimental juzga. Esta contradicción ha sido señalada por los directores de las principales revistas científico-médicas, como Ingelfinger⁵ y Kassirer⁸ de *New England Journal of Medicine*, Lock y Smith²² de *British Medical Journal*, Rennie⁴⁰ de *JAMA*, Fletcher^{32,33} de *Annals of Internal Medicine* y Horton²⁷ de *The Lancet*. La revisión por pares parece tener una norma laxa que no se aplica a los artículos/trabajos que evalúa, y su fundamento científico es escaso —o nulo— pese a los 300 años de historia^{3,14,41}. Sin llegar a ser un proceso inútil y peligroso, puede ser arbitrario, caro, conservador, elitista, lento, opaco y rígido, y carece de base científica;

para algunos, se funda en expectativas irreales acerca de la capacidad de los evaluadores^{5,7}; para otros, no es más que la forma de conseguir la apariencia de respetabilidad científica que necesitan las revistas/tribunales, y debería eliminarse pues su grado de acierto lo convierte en una lotería^{28,38,42}. No sabemos nada acerca de cómo se eligen los revisores y qué información se les requiere para valorar los posibles conflictos de interés —aunque, curiosamente, en un estudio³⁴ se excluyó a los amigos del autor, pero no se dice cómo sabían de la amistad—; tampoco sabemos cómo trabajan para criticar los trabajos, ni por qué se suelen elegir dos para cada artículo (cuando hay desacuerdo entre ambos, la opinión de un tercero es inútil, según Ingelfinger⁵), ni cómo se usan sus informes para tomar la decisión final. En fin, ni siquiera se ha delimitado una lista de cuestiones por estudiar, aunque algo se ha avanzado en los últimos años^{3,5,8,11,13-16,21-25,31-34,39,40}. Hay que aceptar que nunca será un sistema perfecto^{6,8,15}, pero hay que ponerlo en cuestión si queremos que perviva; por ahora, el mayor enemigo es la confianza injustificada que inspira²². Sin investigación, sin innovación y sin cambios se logrará dinamitar la credibilidad de las revistas científicas y de los revisores de manuscritos y de proyectos científicos.

Si admitimos que la selección de los artículos tiene cierto paralelismo con las pruebas de cribado/diagnóstico⁸, será importante determinar la sensibilidad y especificidad de los revisores y las revistas. La preocupación de los directores son los falsos negativos (rechazar los buenos artículos)¹⁷, pero quizá la de los revisores sean los falsos positivos (aceptar trabajos de baja calidad). A los lectores les interesa el impacto de los estudios en su práctica clínica³³, lo que puede llegar a olvidarse si sólo se mide el éxito de una revista por su «factor de impacto» (impac-

to sobre otros autores —muchos de los cuales son, a su vez, evaluadores— no sobre la actividad clínica); el fundamento filosófico de la evaluación por pares descansa tanto en la valoración del rigor científico y en la mejora de la presentación de los manuscritos, como en la selección de los trabajos innovadores y de mayor impacto potencial sobre la salud de los pacientes³⁸.

Experimentación y crítica del anónimo en la revisión por pares

La palabra «anónimo» alude a desconocido u oculto, y en una acepción común, a «carta o informe sin firma en que se dice algo ofensivo». Lejos de esta última acepción, el anónimo es norma en la evaluación de los manuscritos para publicar y de los trabajos para financiar, y tiene ventajas e inconvenientes¹⁴. Respecto a las revistas, es más popular el anónimo simple que el anónimo doble, pero este último es norma que se reafirma en España^{2,10}. ¿Qué sabemos del anónimo en la revisión por pares a través de la experimentación? Pese a su popularidad hay poco fundamento experimental; entre lo publicado podemos destacar:

1. Revistas en las que se ha eliminado totalmente el anónimo sin que se resienta su calidad ni popularidad. British Medical Journal tomó la decisión en 1999²⁵, pero The Medical Journal of Australia empezó en 1996, con un estudio piloto que continúa¹³ en el que el artículo admitido se podía leer antes de su publicación en Internet junto a las críticas firmadas de los revisores; incidentalmente, se admitían sugerencias de los lectores, algunas de las cuales aceptaron los autores (con ello los lectores se convierten en evaluadores).

2. Estudios como el citado de Peters y Ceci⁷, en el que se cambió el nombre de los autores y de las

instituciones, en un sistema de anónimo simple; los autores sugirieron que se pasara a un sistema de anónimo doble que evitara el «efecto Rosenthal» (la expectativa positiva, en este caso, de que los autores conocidos y las instituciones prestigiosas enviaran trabajos de gran calidad e innovadores).

3. Los estudios en que se ha comparado la calidad de la evaluación de manuscritos según se valoran en un sistema de anónimo simple o de anónimo doble. En algunos trabajos se demostró un efecto positivo del anónimo doble, pero siempre escaso⁴³⁻⁴⁵. En un complejo ensayo clínico, con 527 manuscritos enviados a British Medical Journal, distribuidos para su evaluación en 8 grupos (cuatro de anónimo simple y cuatro de anónimo doble, en que dos de cada grupo se subdividieron para comparar revisores que se cruzaban los informes), no hubo diferencias respecto a la calidad de los informes, ni respecto al tiempo de revisión, ni a la recomendación acerca de su publicación, en ningún grupo; tampoco hubo «efecto Hawthorne» (modificación del resultado por saberse en una situación experimental)⁴⁶. En otro complejo ensayo clínico en la misma revista (comentado respecto a la capacidad de los evaluadores de detectar 8 errores introducidos artificialmente en un manuscrito³⁴), con 420 revisores divididos en 5 grupos (control —anónimo simple, como habitualmente—, y 4 experimentales cruzados, dos en que los revisores firmaban los informes y dos con anónimo doble) no hubo diferencias entre los grupos respecto a la detección de errores, pero los evaluadores con anónimo doble rechazaron menos manuscritos. En otro ensayo clínico, con 118 manuscritos de 5 revistas estadounidenses, no hubo diferencias de calidad entre los informes de evaluadores con anónimo simple comparados con los de anónimo doble⁴⁷.

4. Los estudios en que se ha valorado el grado de anónimo (en el sistema de anónimo doble). Estos trabajos demuestran que se logra en un porcentaje variable, 50-80%^{43,44}. El éxito no parece depender de la política de la revista, pues en un estudio en que se valoraron manuscritos con anónimo doble, en 3 revistas que utilizaban habitualmente el anónimo doble y 4 que usaban el anónimo simple, el autor no fue identificado en un 60 y 58%, respectivamente; los revisores con más experiencia identificaron más frecuentemente a los autores¹⁶. En el estudio estadounidense comentado⁴⁷, con 118 manuscritos y 5 revistas, se logró el anónimo doble en un 68% de los casos, pero no hubo diferencias respecto a los manuscritos en los que se identificó al autor; los manuscritos de autores muy conocidos fueron más fácilmente identificados.

El fundamento experimental del anónimo en la revisión por pares es muy débil, tanto para el simple como para el doble. Parece más una cuestión de costumbre, de apariencia de neutralidad y de ecuanimidad, que una cuestión bien fundada. La ciencia se aleja progresivamente de los informes y actividades anónimas, que exponen a los autores a una caja negra²¹, a un tribunal kafkiano²⁵, a un soviético sin control⁴² y al oscurantismo de la Inquisición. Por ejemplo, en Estados Unidos las cartas de recomendación para los puestos de trabajo en las universidades son públicas en la actualidad; empezaron siendo secretas, pasaron después a ser públicas y anónimas, y han terminado siendo públicas y firmadas. El único argumento contra un proceso de revisión abierto en ambos sentidos —evaluaciones firmadas, manuscritos con autores identificados— es la dificultad que pueden tener los revisores que evalúen manuscritos

de autores de mayor autoridad científica o académica²⁵, pero las desventajas son múltiples, algunas no bien estudiadas todavía; por ejemplo, arbitrariedad, dogmatismo, freno a la publicación de resultados, rechazo a la innovación, robo de ideas y métodos y sesgos a favor de los autores establecidos y de las instituciones prestigiosas (cuyos trabajos se identifican con facilidad, por más que se quiera ocultar su filiación). Frente a estos últimos, el «efecto Rosenthal», la expectativa de que sus trabajos serán de calidad e innovadores, hace a los evaluadores sensibles a evitar los falsos negativos (evitar el rechazo de trabajos de buena calidad)⁷, como si fueran el director de la revista; por el contrario, frente a los trabajos que se sabe, o se suponen, de instituciones/ autores no conocidos, estos evaluadores serán sensibles a evitar los falsos positivos (aceptar trabajos de baja calidad). Si el proceso es abierto, al menos se puede intentar medir la sensibilidad y especificidad sin el falso espejo del anónimo «parcial».

Frente al anónimo hay, finalmente, razones éticas. Repugna a la mente que la transparencia que se pide y exige al proceso, al método y al resultado científico, se rompa, precisamente, cuando se valora para su publicación o su financiación (y sin que haya experimentación científica que lo justifique). Todo informe científico debe tener elegancia, calidad y neutralidad suficientes para ser exhibido en público sin que su firmante se avergüence. Por las mismas razones no se entiende que la calidad mejore al dar opinión sobre autor (aparentemente) desconocido, y así lo demuestra la experimentación. Las revisiones científicas deberían ser, pues, abiertas en los dos sentidos. Si el mundo científico es una fraternidad, y no un mundo en guerra, sobran las celadas.

Conclusión

Las revistas científicas y las agencias que financian la investigación necesitan filtrar los trabajos de investigación, pues no todo lo que se produce merece ser publicado y/o financiado. La revisión por pares y el anónimo, simple o doble, de la evaluación han permitido el avance de la ciencia pero carecen de base experimental que justifique su amplia aceptación y su implantación rutinaria. Existe un movimiento mundial de crítica y experimentación acerca de estas cuestiones al que podría sumarse España sin dificultad, pero el primer escollo es lograr romper la unanimidad, establecer un debate que agite las tranquilas aguas de la evaluación científica. Nada impide que: a) se experimente con Internet, en el sentido propuesto por los australianos^{11,13}, por los pediatras norteamericanos²⁴ y por Varmus, el director de los National Institutes of Health de Estados Unidos⁴; b) se abra por completo el proceso de revisión, como practica *British Medical Journal*²⁵ y se valoren sus consecuencias; c) se analice con detenimiento el proceso de decisión de los evaluadores, con el esquema propuesto por Kassirer y Champion⁸ o con otro alternativo, y también que se estudie el proceso de selección de los evaluadores y los conflictos de interés que se generan, y d) se consideren alternativas tan sencillas como las propuestas por *Medical Care*, de pedir al autor que envíe el manuscrito con los informes de revisores elegidos según su buen criterio²³, o las sugerencias de Fabiato, de pedir al autor que sugiera revisores y de hacer constar en el texto publicado los nombres de los revisores del manuscrito¹⁴ u otras alternativas, como incluir los informes de los revisores junto a las notas de «retracción» de los artículos fraudulentos o que se han publicado por duplicado. En cualquier caso, es necesaria una cura de humildad,

pues la revisión por pares no evita que las publicaciones sirvan más a los autores que a los lectores⁴⁸ o la publicación duplicada^{49,50}. La ciencia no existe en el vacío, sino que depende de los científicos. Los científicos progresan cuando conocen sus límites, cuando exponen sus ideas y hallazgos al público escrutinio. Hay que buscar un equilibrio entre la innovación y la calidad que ayude al progreso científico, al trabajo de los investigadores, a la difusión y financiación de sus trabajos. La revisión por pares no tiene fundamentos claramente establecidos y no es un método infalible; vale la pena, pues, el debate científico acerca de su continuidad.

Bibliografía

1. Committee on the Conduct of Science. On being a scientist. *Proc Natl Acad Sci USA* 1989; 86: 9053-9074.
2. Rodríguez Artalejo F. Ocultar la identidad de los autores y evaluadores de artículos. *Gac Sanit* 1996; 10: 159-160.
3. Burnham JC. The evolution of editorial per review. *JAMA* 1990; 263: 1323-1329.
4. Relman AS. The NIH «E-Biomed» proposal. A potential threat to the evaluation and orderly dissemination of new clinical studies. *N Engl J Med* 1999; 340: 1828-1829.
5. Ingelfinger FJ. Peer review in biomedical publication. *Am J Med* 1974; 56: 686-692.
6. Harnard S. Peer commentary on peer review. *Behav Brain Sciences* 1982; 5: 185-186.
7. Peters DP, Ceci SJ. Peer-review practices of psychological journals. The fate of published articles, submitted again. *Behav Brain Sciences* 1982; 5: 187-255.
8. Kassirer JP, Champion EW. Peer review. Crude and understudied, but indispensable. *JAMA* 1994; 272: 96-97.
9. Pierie JPEN, Walvoort HC, Overbeke JPM. Readers' evaluation of effect of peer review and editing on quality of articles in the *Nederlands Tijdschrift voor Geneeskunde*. *Lancet* 1996; 348: 1480-1483.
10. Jiménez Villa J. La calidad de las publicaciones. *Aten Primaria* 1996; 17: 367-368.
11. Bingham C. Peer review on the Internet: a better class of conversation. *Lancet* 1998; 351 (Supl 1): 10-14.
12. Reid AJ. Canadian Family Physician's peer reviewers. *Unsung heroes*. *Can Fam Physician* 1998; 44: 13-14.
13. Bingham CM, Higgins G, Coleman R, Weyden MBVD. The Medical Journal of Australia Internet peer-review study. *Lancet* 1998; 352: 441-445.
14. Fabiato A. Anonymity of reviewers. *Cardiovasc Res* 1994; 28: 1134-1139.
15. Taylor R. NIH panel to monitor peer review in action. *Nature* 1995; 375: 438.
16. Cho Mk, Justice AC, Winker MA, Berlin JA, Waeckerle JF, Callahan ML et al. Masking author identity in peer review. What factors influence masking success? *JAMA* 1998; 280: 243-245.
17. Yankauer A. Who are the peer reviewers and how much do they review. *JAMA* 1990; 263: 1338-1340.
18. Olsen CM. Peer review of the biomedical literature. *Am J Emerg Med* 1990; 8: 356-358.
19. Rodríguez Artalejo F. Prevención de enfermedades vasculares en las mujeres. *Aten Primaria* 1998; 22 (Supl 1): 200-295.
20. Petitti DB. Hormone replacement therapy and heart disease prevention. Experimentation trumps observation. *JAMA* 1998; 280: 650-652.
21. Smith R. Peer review: reform or revolution. Time to open up the black box of peer review. *BMJ* 1997; 315: 759-760.
22. Smith R. Promoting research into peer review. An invitation to join it. *BMJ* 1994; 309: 143-144.
23. Neuhauser D. Peer review and the research commons. A problem of success. *Med Care* 1997; 35: 301-302.
24. Anderson KR, Lucey JF. Una nueva capacidad: la revisión post-publicación por expertos para *Pediatrics*. *Pediatrics* (ed. esp.) 1999; 48: 1.
25. Smith R. Opening up BMJ peer review. *BMJ* 1999; 318: 4-5.
26. Plasència A. Gaceta Sanitaria: un mensajero en la casa común de la salud pública. *Gac Sanit* 1999; 13: 4-5.
27. Horton R. Luck, lotteries and loopholes of grant review. *Lancet* 1996; 348: 1255-1256.
28. LaPorte RE, Marler E, Akazawa S, Sauer F, Gamboa C, Shenton C et al. The death of biomedical journals. *BMJ* 1995; 310: 1387-1390.
29. Squires BP. Peer review under scrutiny. Report on the third International Congress in Prague, 1997. *Can Fam Physician* 1998; 44: 15-16.
30. Klein MC. Studying epistomology: when beliefs conflict with science. *J Fam Pract* 1995; 41: 483-488.
31. Abby M, Massey MD, Galanduk S, Polk HC. Peer review is an effective screening process to evaluate medical manuscripts. *JAMA* 1994; 272: 105-107.
32. Roberts J, Fletcher RH, Fletcher SW. Effects of peer review and editing on the readability of articles published in *Annals of Internal Medicine*. *JAMA* 1994; 272: 119-121.
33. Justice AC, Berlin JA, Fletcher SW, Fletcher RH, Goodman SN. Do readers and peer reviewers agree on manuscript quality? *JAMA* 1994; 272: 117-119.
34. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports. A randomized controlled trial. *JAMA* 1998; 280: 237-240.
35. Lock S, Wells F, editores. *Fraud and misconduct*. Londres: BMJ Pub., 1996.
36. Vilarroya O, editor. *Ética de la publicación médica*. En: Varios, editores. *Medicina clínica. Manual de estilo*. Barcelona: Doyma, 1993; 117-135.
37. Smith R. Misconduct in research: editors respond. The Committee on Publications Ethics (COPE) is formed. *BMJ* 1997; 315: 201-202.
38. Horrobin DF. The philosophical basis of peer review and the suppression of innovation. *JAMA* 1990; 263: 1438-1441.
39. Polak JF. The role of the manuscript reviewer in the peer review process. *Am J Roentgenol* 1995; 165: 685-688.
40. Rennie D, Knoll E. Investigating peer review. *Ann Intern Med* 1988; 109: 181.
41. Kronick DA. Peer review in 18th-century scientific journalism. *JAMA* 1990; 262: 1321-1322.
42. Horrobin DF. Peer review of grant applications: a harbinger for mediocrity in clinical research? *Lancet* 1996; 348: 1293-1295.
43. Evans AT, McNutt RA, Fletcher SW, Fletcher RH. The characteristics of peer review who produce good quality reviews. *J Gen Intern Med* 1993; 8: 422-428.
44. McNutt RA, Evans AT, Fletcher RH,

- Fletcher SW. The effects of blinding on the quality of peer review. A randomized trial. *JAMA* 1990; 263: 1371-1376.
45. Fischer M, Friedman SB, Strauss B. The effects of blinding on acceptance of research papers by peer review. *JAMA* 1994; 272: 143-146.
46. Rooyen SV, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review. A randomized trial. *JAMA* 1998; 280: 234-237.
47. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D, PEER investigators. Does masking author identify improve peer review quality? A randomized controlled trail. *JAMA* 1998; 280: 240-242.
48. O'Donnell M. Evidence-based illiteracy: time to rescue «the literature». *Lancet* 2000; 355: 489-491.
49. Navarro V. La relevancia de la experiencia norteamericana en la reforma del sistema nacional de salud británico. El caso del GP budget holding. *Gac San* 1991; 5: 276-283.
50. Navarro V. The relevance of the US: experience to the reforms in the British National Health Service: the case of general practitioner fund holding. *Internat J Health Services* 1991; 21: 381-397.