



MENOS ANIMALES, MÁS CIENCIA

ESTRATEGIAS PARA REDUCIR EL NÚMERO DE ANIMALES EN LA INVESTIGACIÓN EXPERIMENTAL

Enrique Font y Matthew Kramer

Aunque no existen cifras precisas, se estima que cada año son sacrificados entre varios cientos y varios miles de millones de animales en todo el mundo en experimentos científicos, sobre todo en la investigación biomédica (Taylor, Gordon, Langley y Higgins, 2008). Solo en España se utilizaron cerca de 800.000 animales vertebrados en procedimientos experimentales en 2014. De estos –la mayoría ratones y ratas de laboratorio–, 62.000 fueron sometidos a manipulaciones que les provocaron «dolor, sufrimiento o angustia severos». Es una cantidad enorme de animales, si bien varios órdenes de magnitud por debajo de los que sacrificamos para otros usos como la alimentación o la caza.

La experimentación con animales proporciona indudables beneficios que no son alcanzables por métodos alternativos como los cultivos celulares o los modelos matemáticos, pero también plantea un serio problema ético. Por eso, cada vez son más los que cuestionan la utilidad de mucha de la experimentación que se hace actualmente con animales y que exigen mayor rigor y transparencia a los investigadores. Varios estudios recientes han alertado de que una parte importante de la experimentación con animales no produce resultados válidos. En algunos casos esto se debe a que los animales experimentales no son buenos modelos para el estudio de determinadas enfermedades humanas (Seok et al., 2013). Otro motivo, quizás más preocupante, es que muchos experimentos están mal diseñados, mal ejecutados o sus resultados no están analizados correctamente, lo que impide extraer de ellos conclusiones válidas (Kilkenny et al., 2009).

■ LAS TRES ERRES DE LA EXPERIMENTACIÓN CON ANIMALES

Un problema, desgraciadamente muy extendido, tiene que ver con el tamaño muestral, es decir, con el número de animales que intervienen en un experimento. La legislación internacional establece que la experimentación con animales, cuando está justificada, debe hacerse utilizando el número mínimo de ejemplares que permi-

ta llegar a conclusiones válidas desde el punto de vista estadístico. Esta medida responde al conocido como principio de las tres erres –refinamiento, reducción y reemplazo– que muchos consideran el principio rector que debe guiar la experimentación con animales. En concreto, la R de reducción consiste en promover el uso de métodos que minimicen el número de animales utilizados en un experimento, para permitir a los investigadores obtener la misma información con menos animales, o más información con el mismo número de animales.

Tan malo es utilizar en un experimento demasiados animales como utilizar demasiado pocos. Si el número de animales empleados es innecesariamente elevado estaremos desperdiciando inútilmente animales (y tiempo y dinero). Pero si el número de animales es insuficiente no podremos llegar a conclusiones científicamente válidas. Por ejemplo, es posible que concluyamos erróneamente que un determinado tratamiento no tiene efecto cuando en realidad sí que lo tiene, y el sacrificio de los animales implicados en el experimento habrá sido en vano. Según algunos estudios, entre el 75 % y el 90 % de los resultados de experimentos publicados en revistas de prestigio no se han podido replicar en otros laboratorios, en gran medida debido a la utilización de tamaños muestrales inadecuados (Begley y Ioannidis, 2015; Button et al., 2013).

■ DISEÑO EXPERIMENTAL Y GRUPOS DE CONTROL

En un artículo publicado recientemente en la revista *Biological Reviews* (Kramer y Font, 2017), los autores hemos propuesto varias estrategias para la reducción del número de animales de experimentación centradas en los grupos de control. Pero, ¿qué son y qué función desempeñan los grupos de control? Supongamos que un investigador recibe el encargo de comprobar si un determinado fármaco es eficaz en el tratamiento de una dolencia que provoca fiebre moderada y malestar general. El investigador administra el fármaco a un grupo de pacientes a los que se les ha diagnosticado la enferme-

«SE ESTIMA QUE CADA AÑO SON SACRIFICADOS ENTRE VARIOS CIENTOS Y VARIOS MILES DE MILLONES DE ANIMALES EN TODO EL MUNDO EN EXPERIMENTOS CIENTÍFICOS»



Concha Molina



Concha Molina



dad y observa que la fiebre y el malestar en estos pacientes remiten después de tres días. ¿Puede el investigador concluir solo con esta información que el fármaco tiene algún efecto beneficioso? Evidentemente no. Para poder llegar a una conclusión acerca de la eficacia del fármaco sería necesario comprobar cuánto duran la fiebre y el malestar en pacientes que no han recibido ningún tratamiento. El grupo de pacientes que recibe el fármaco se conoce como grupo de tratamiento o grupo experimental, mientras que el grupo de pacientes sin tratar es el grupo de control.

Para que el experimento estuviese bien hecho sería además necesario que los individuos de los dos grupos fuesen lo más parecidos posible (sexo, edad, historial clínico, severidad de los síntomas...), de manera que la única diferencia entre ellos relevante para el curso de la enfermedad fuese la que provoca el fármaco. Una forma relativamente sencilla de conseguirlo consiste en asignar a los pacientes a uno u otro grupo de forma aleatoria. Si después de tomar estas precauciones observamos que la fiebre y el malestar en los pacientes no tratados duran, por ejemplo, cinco días (en comparación con los tres días que duran en los pacientes que reciben el fármaco),

entonces tendremos motivos fundados para sospechar que el nuevo fármaco podría ser un tratamiento eficaz para la dolencia que estamos estudiando. Antes de recomendar su uso aún deberíamos hacer muchas comprobaciones, por ejemplo, para determinar que el fármaco no tiene efectos secundarios indeseables, pero al menos nuestro experimento sería en principio correcto desde el punto de vista del diseño experimental.

Pero incluso un ejemplo aparentemente tan simple encierra cierta complejidad. La inclusión de un grupo de control se podría considerar inaceptable desde el punto de vista ético, ya que supondría privar a pacientes enfermos de un tratamiento que podría proporcionarles alivio. En la práctica, sería preferible comparar el grupo de tratamiento que recibe el fármaco experimental con un grupo de pacientes que reciban la terapia estándar tradicional para ese tipo de dolencia. Por tanto, en la decisión acerca de los controles que debemos utilizar en un experimento, intervienen consideraciones no solo de índole estadística, sino también éticas e incluso económicas.

Los grupos de control son un elemento muy importante del diseño experimental, pero lo cierto es que la mayoría de los investigadores sabemos muy poco acer-



Paula Navarro

La legislación internacional establece que la experimentación con animales, cuando está justificada, debe hacerse utilizando el número mínimo de ejemplares que permita llegar a conclusiones válidas desde el punto de vista estadístico. Todas las fotografías de este artículo han sido tomadas en el animalario del Servicio Central de Apoyo a la Investigación Experimental de la Universitat de València.

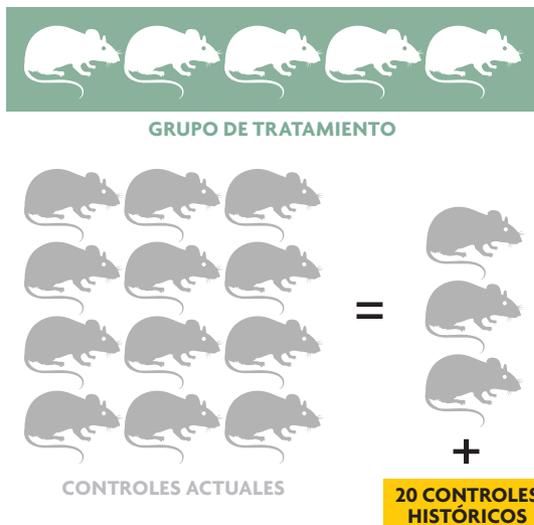
ca de los controles y de cómo utilizarlos correctamente. No es sorprendente, ya que es un tema al que la mayoría de los libros de texto dedican muy poca atención. De una veintena de textos de estadística y diseño experimental dirigidos a biólogos y psicólogos publicados desde el año 2000, más de la mitad ni siquiera tienen la entrada «control» en el índice. De los que sí incluyen información sobre controles, la mayoría se limita a mencionarlos sin explicar ni los tipos de controles ni la función que desempeñan en relación con los objetivos del experimento. Un texto dirigido a biólogos describe los controles de la siguiente manera: «Si uno de los tratamientos corresponde a algún tipo de nivel “natural” de un factor se lo puede denominar control». Con explicaciones tan crípticas como esta no debe sorprendernos que los investigadores tengan una idea muy superficial, y a menudo errónea, de lo que supone utilizar grupos de control.

Muchos piensan, por ejemplo, que todos los experimentos deben tener un grupo de control, cuando en realidad no es así. Imaginemos por ejemplo que queremos comparar la eficacia de dos fertilizantes, A y B. Si lo único que nos interesa es averiguar si existen diferencias entre los dos productos, entonces no hay necesidad de incluir un grupo de control. En el mejor de los casos, incluir un grupo de control compuesto por plantas a las que no proporcionamos ningún fertilizante únicamente nos diría que los productos A y B mejoran el crecimiento de las plantas, algo que ya deberíamos saber dado que se trata de fertilizantes. Por el contrario, otros experimentos pueden requerir no uno, sino varios grupos de control.

Tampoco es cierto que los grupos de control tengan que ser necesariamente del mismo tamaño que los grupos de tratamiento, una idea errónea que fomentan muchos programas informáticos para el cálculo del tamaño muestral. Un diseño experimental en el que todos los grupos son del mismo tamaño se conoce como un «diseño balanceado». En general, los diseños balanceados proporcionan mejores resultados que los no balanceados, pero hay circunstancias que pueden hacer preferible utilizar grupos de control más grandes o más pequeños que los grupos de tratamiento (Bate y Karp, 2014; Ruxton y Colegrave, 2011). Además, los grupos de control tienden a ser intrínsecamente menos variables que los grupos de tratamiento, por lo que un diseño balanceado con frecuencia dará lugar a un problema de heterogeneidad de la varianza.

■ CÓMO TOMAR PRESTADA INFORMACIÓN DE EXPERIMENTOS PREVIOS: LOS CONTROLES HISTÓRICOS

Una posibilidad para reducir el número de animales de un experimento consiste precisamente en reducir el tamaño del grupo o grupos de control. La información histórica es siempre importante a la hora de diseñar un experimento: los resultados de experimentos anteriores son la base sobre la que concebimos y diseñamos nuevos experimentos. Nuestra propuesta consiste en utilizar información extraída de esos experimentos anteriores, en concreto de los grupos de control. En el caso de experimentos que se repiten de manera rutinaria en un laboratorio, siempre con el mismo tipo de animales, en condiciones estándar y utilizando los mismos protocolos experimentales, una forma muy eficaz de reducir el número de animales sin sacrificar la capacidad de llegar a conclusiones válidas consiste en «tomar prestada» información de los controles utilizados en experimentos previos. No se trata de eliminar el grupo de control, sino de reducir su tamaño incorporando información de controles de experimentos anteriores, es decir, de controles históricos.



Utilizando datos simulados de características similares a los que se manejan en biomedicina encontramos que, si los controles históricos son suficientemente similares a los actuales, la capacidad para llegar a conclusiones válidas desde el punto de vista estadístico es la misma utilizando doce controles actuales que utilizando solo tres controles actuales y tomando prestada información (media y varianza) de veinte controles históricos. Esto supone un ahorro de nueve animales.

Por ejemplo, en un laboratorio de toxicología es posible que se repita varias veces el mismo procedimiento experimental, variando únicamente la sustancia tóxica empleada, que es distinta en cada experimento concreto. ¿Es necesario incluir un control del mismo tamaño que el grupo de tratamiento cada vez que se repite el experimento? La respuesta es un rotundo no. Si ya disponemos de controles históricos, podemos incorporarlos en el diseño experimental, lo que nos permitirá utilizar un grupo de control actual con menos animales.

Algunos investigadores pueden tener reparos porque piensan que reducir el número de animales que utilizan en un experimento limitará sus posibilidades de conseguir un resultado estadísticamente significativo. Pero en nuestro trabajo demostramos, utilizando simulaciones y datos reales de la bibliografía, que esos reparos no están justificados. En general, nuestros resultados demuestran que la incorporación de información prestada de los controles históricos aumenta la capacidad para llegar a conclusiones válidas (i. e., incrementa la potencia de las pruebas estadísticas) al tiempo que permite reducir el número de animales en el grupo de control actual.

Nuestra propuesta no es en absoluto original. Los controles históricos se utilizan desde hace mucho tiempo en la investigación clínica (Viele et al., 2014). Sin embargo, no todos los especialistas están convencidos

«UNA POSIBILIDAD PARA REDUCIR EL NÚMERO DE ANIMALES DE UN EXPERIMENTO CONSISTE EN REDUCIR EL TAMAÑO DEL GRUPO O GRUPOS DE CONTROL.»

de su utilidad. Algunos rechazan explícitamente el uso de controles históricos argumentando que en general será muy difícil garantizar que, si se observan diferencias entre los grupos de tratamiento y de control, estas no sean debidas a que las circunstancias en el momento en el que se recogieron los datos del control histórico eran distintas de las actuales (Bailey, 2008). Por ejemplo, un medicamento para el tratamiento del asma puede parecer muy eficaz sencillamente debido a que en el momento en el que se realizó el experimento los niveles de contaminación eran muy inferiores a los que prevalecían en periodos anteriores cuando se recogieron los datos históricos.

Estas objeciones justifican, al menos en parte, la mala reputación que para muchos tienen los controles históricos. Pero, en los ensayos clínicos, los controles históricos tradicionalmente consisten en datos de pacientes que los investigadores extraen de trabajos publicados o de los archivos de los hospitales. En ocasiones ni siquiera se trata de datos de pacientes individuales, sino de datos agregados procedentes de un conjunto de pacientes. La información así recogida es muy heterogénea, y a menudo no tiene en cuenta diferencias importantes en el diagnóstico o en el tratamiento de dichos controles históricos que podrían llevar a conclusiones erróneas. Nuestra propuesta es más conservadora, ya que recomendamos tomar prestada información de controles históricos utilizados en el mismo laboratorio, por los mismos investigadores, con los mismos animales y utilizando protocolos experimentales idénticos o muy parecidos, un tipo de situación que pensamos que es habitual en muchos centros de investigación.

situación que pensamos que es habitual en muchos centros de investigación.

■ ¿CUÁNDO PODEMOS UTILIZAR CONTROLES HISTÓRICOS?

La clave para utilizar correctamente los controles históricos consiste en procurar que los controles históricos y los actuales sean lo más parecidos posible. En nuestro trabajo proporcionamos recomendaciones para tomar prestada información de controles históricos en siete escenarios distintos que corresponden a situaciones comunes en la investigación biomédica. Los escenarios contemplan un amplio abanico de posibilidades, desde situaciones en las que podemos asumir que los controles históricos y los actuales son prácticamente intercambiables (tienen la misma media y varianza), hasta otras en las que las diferencias entre los controles históricos y los actuales son mayores, lo que limita la cantidad de información his-



Algunos investigadores pueden poner reparos porque piensan que reducir el número de animales que utilizan en un experimento limitará sus posibilidades de conseguir un resultado estadísticamente significativo. Pero en este trabajo se demuestra, utilizando simulaciones y datos reales de la bibliografía, que esos reparos no están justificados.

tórica que podemos tomar prestada. Nuestros resultados demuestran que la inclusión de datos históricos siempre mejora la capacidad del investigador para llegar a conclusiones válidas desde el punto de vista estadístico (es decir, incrementa la potencia de las pruebas estadísticas).

Con respecto al ahorro de animales, este depende del tipo de escenario considerado. Como podíamos sospechar, las situaciones más restrictivas (es decir, menos parecido entre controles históricos y actuales) permiten un menor ahorro de animales en el grupo de control actual. Pero, si podemos asumir que los controles históricos son muy similares a los actuales, el ahorro en términos de animales puede ser considerable. Por ejemplo, en algunas de nuestras simulaciones encontramos que tres controles actuales más veinte históricos proporcionan la misma potencia estadística que doce controles actuales, es decir, la incorporación de datos históricos permite un ahorro de animales del 75 % sin ninguna merma en la capacidad para llegar a conclusiones válidas.

Reducir al mínimo indispensable el número de animales que participan en un experimento es, además de un imperativo legal, un objetivo deseable desde el punto de vista ético. Además, reducir el tamaño muestral de un experimento reduce también su coste económico y el tiempo y el esfuerzo necesario para llevarlo a cabo. Des-

graciadamente, de las tres erres, la de la reducción es la que menos avances ha experimentado en las últimas décadas. Las recomendaciones que hacemos en nuestro trabajo podrían contribuir a mejorar la calidad de la investigación, reduciendo a la vez el número de animales necesarios para obtener conclusiones válidas.

Básicamente, se trata de que los investigadores presten más atención a los grupos de control. Es necesario que, en lugar de considerarlos como un detalle puramente cosmético, conozcan los tipos de controles y las funciones que desempeñan en relación con el diseño experimental. En muchos experimentos los grupos de control son enteramente prescindibles. En particular, cuando se somete a los animales del grupo de control a procedimientos que provocan sufrimiento y dolor y que tienen resultados enteramente previsibles, el investigador debería plantearse la posibilidad de eliminar o reducir drásticamente, por motivos éticos, el grupo de control. Si es necesario incluir uno o más grupos de control en el experimento, entonces los investigadores deberían buscar estrategias que permitan reducir su número, incluyendo—cuando sea razonable hacerlo— tomar prestada información de controles históricos. ☺

REFERENCIAS

- Bailey, R. A. (2008). *Design of comparative experiments*. Cambridge: Cambridge University Press.
- Bate, S., & Karp, N. A. (2014). A common control group: Optimising the experiment design to maximise sensitivity. *PLOS ONE*, 9(12), e114872. doi: [10.1371/journal.pone.0114872](https://doi.org/10.1371/journal.pone.0114872)
- Begley, C. G., & Ioannidis, J. P. (2015). Reproducibility in science: Improving the standard for basic and preclinical research. *Circulation Research*, 116(1), 116–126. doi: [10.1161/CIRCRESAHA.114.303819](https://doi.org/10.1161/CIRCRESAHA.114.303819)
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience*, 14(6), 365–376. doi: [10.1038/nrn3475](https://doi.org/10.1038/nrn3475)
- Kilkenny, C., Parsons, N., Kadyszewski, E., Festing, M. F. W., Cuthill, I. C., Fry, D., ... Altman, D. G. (2009). Survey of the quality of experimental design, statistical analysis and reporting of research using animals. *PLOS ONE*, 4, e7824. doi: [10.1371/journal.pone.0007824](https://doi.org/10.1371/journal.pone.0007824)
- Kramer, M., & Font, E. (2017). Reducing sample size in experiments with animals: Historical controls and related strategies. *Biological Reviews of the Cambridge Philosophical Society*, 92(1), 431–445. doi: [10.1111/brv.12237](https://doi.org/10.1111/brv.12237)
- Ruxton, G., & Colegrave, N. (2011). *Experimental design for the life sciences*. Chippingham: Oxford University Press.
- Seok, J., Warren, H. S., Cuenca, A. G., Mindrinos, M. N., Baker, H. V., Xu, W., ... Tompkins, R. G. (2013). Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 110(9), 3507–3512. doi: [10.1073/pnas.1222878110](https://doi.org/10.1073/pnas.1222878110)
- Taylor, K., Gordon, N., Langley, G., & Higgins, W. (2008). Estimates for worldwide laboratory animal use in 2005. *Alternatives to Laboratory Animals*, 36(3), 327–342.
- Viele, K., Berry, S., Neuenschwander, B., Amzal, B., Chen, F., Enas, N., ... Thompson, L. (2014). Use of historical control data for assessing treatment effects in clinical trials. *Pharmaceutical Statistics*, 13(1), 41–54. doi: [10.1002/pst.1589](https://doi.org/10.1002/pst.1589)

Enrique Font. Laboratorio de Etología, Instituto Cavanilles de Biodiversidad y Biología Evolutiva. Universitat de València (Espanya).

Matthew Kramer. Statistics Group, Agricultural Research Service (Beltsville). Departamento de Agricultura de Estados Unidos.