

# MENYS ANIMALS, MÉS CIÈNCIA

## ESTRATÈGIES PER A REDUIR EL NOMBRE D'ANIMALS EN LA RECERCA EXPERIMENTAL

Enrique Font i Matthew Kramer

Encara que no hi ha xifres precises, s'estima que cada any són sacrificats entre uns quants centenars i diversos milers de milions d'animals arreu del món en experiments científics, sobretot en la recerca biomèdica (Taylor, Gordon, Langley i Higgins, 2008). Només a Espanya es van utilitzar prop de 800.000 animals vertebrats en procediments experimentals en 2014. D'aquests –la majoria ratolins i rates de laboratori–, 62.000 van ser sotmesos a manipulacions que els van provocar «dolor, patiment o angoixa severa». És una quantitat enorme d'animals, si bé diversos ordres de magnitud per sota dels que sacrificuem per a altres usos com l'alimentació o la caça.

L'experimentació amb animals proporciona indubtables beneficis que no es poden obtenir per mètodes alternatius com els cultius cel·lulars o els models matemàtics, però també planteja un problema ètic ben seriós. Per això, cada vegada són més els que qüestionen la utilitat de molta de l'experimentació que es fa actualment amb animals i que exigeixen major rigor i transparència als investigadors. Diversos estudis recents han alertat que una part important de l'experimentació amb animals no produeix resultats vàlids.

En alguns casos perquè els animals experimentals no són bons models per a estudiar determinades malalties humanes (Seok et al., 2013). Un altre motiu, potser més preocupant, és que molts experiments estan mal dissenyats, mal executats o els seus resultats no s'han analitzat correctament i per tant no se'n poden extreure conclusions vàlides (Kilkenny et al., 2009).

### ■ LES TRES ERRES DE L'EXPERIMENTACIÓ AMB ANIMALS

Un problema, desgraciadament molt estès, té a veure amb la grandària de la mostra, és a dir, amb el nombre d'animals que intervenen en un experiment. La legislació internacional estableix que l'experimentació amb animals, quan està justificada, ha de fer-se utilitzant

el nombre mínim d'exemplars que permeta arribar a conclusions vàlides des del punt de vista estadístic. Aquesta mesura respon al conegut com a principi de les tres erres –refinament, reducció i reemplaçament– que molts consideren el principi rector que ha de guiar l'experimentació amb animals. En concret, la R de reducció consisteix a promoure l'ús de mètodes que minimitzen el nombre d'animals utilitzats en un experiment, per a permetre als investigadors obtenir la mateixa informació amb menys animals, o més informació amb el mateix nombre d'animals.

Tan perjudicial és utilitzar en un experiment massa animals com utilitzar-ne massa pocs. Si el nombre d'animals usats és innecessàriament elevat estarem malbaratant inútilment animals (i temps i diners).

Però si el nombre d'animals és insuficient no podrem arribar a conclusions científicament vàlides. Per exemple, és possible que concloquem erròniament que un determinat tractament no té efecte quan en realitat sí que en té, i el sacrifici dels animals implicats en l'experiment haurà estat en va. Segons alguns estudis, entre el 75 % i el 90 % dels resultats d'experiments publicats en revistes de prestigi no s'han pogut replicar

en altres laboratoris, en gran manera a causa de la utilització de grandàries de mostra inadequades (Begley i Ioannidis, 2015; Button et al., 2013).

### ■ DISSENY EXPERIMENTAL I GRUPS DE CONTROL

En un article publicat recentment en la revista *Biological Reviews* (Kramer i Font, 2017), els autors hem proposat diverses estratègies per a reduir el nombre d'animals d'experimentació centrades en els grups de control. Però, què són i quina funció exerceixen els grups de control? Suposem que un investigador rep l'encàrrec de comprovar si un determinat fàrmac és eficaç en el tractament d'una malaltia que provoca febre moderada i malestar general. L'investigador administra el fàrmac a un grup de pacients a qui se'ls

**«S'ESTIMA QUE CADA ANY SÓN SACRIFICATS ENTRE UNS QUANTS CENTENARS I DIVERSOS MILERS DE MILIONS D'ANIMALS ARREU DEL MÓN EN EXPERIMENTS CIENTÍFICS»**





Concha Molina



Concha Molina



ha diagnosticat la malaltia i observa que la febre i el malestar en aquests pacients remeten després de tres dies. Pot l'investigador concloure només amb aquesta informació que el fàrmac té algun efecte beneficiós? Evidentment no. Per a poder arribar a una conclusió sobre l'eficàcia del fàrmac seria necessari comprovar quant duren la febre i el malestar en pacients que no han rebut cap tractament. El grup de pacients que rep el fàrmac es coneix com a grup de tractament o grup experimental, mentre que el grup de pacients sense tractar és el grup de control.

A més, perquè l'experiment estiguera ben fet seria necessari que els individus dels dos grups foren tan semblants com fóra possible (sexe, edat, historial clínic, severitat dels símptomes...), de manera que l'única diferència entre ells rellevant per al curs de la malaltia fóra la que provoca el fàrmac. Una manera relativament senzilla d'aconseguir-ho consisteix a assignar els pacients a un o altre grup de manera aleatòria. Si després de prendre aquestes precaucions observem que la febre i el malestar en els pacients no tractats duren, per exemple, cinc dies (en comparació amb els tres dies que duren en els pacients que reben el fàrmac), llavors tindrem mo-

tius fonamentats per a sospitar que el nou fàrmac podria ser un tractament eficaç per a la malaltia que estem estudiant. Abans de recomanar-ne l'ús encara hauríem de fer moltes comprovacions, per exemple, per determinar que el fàrmac no té efectes secundaris indesitjables, però almenys el nostre experiment seria en principi correcte des del punt de vista del disseny experimental.

Però fins i tot un exemple aparentment tan simple conté una certa complexitat. La inclusió d'un grup de control es podria considerar inacceptable des del punt de vista ètic, ja que representaria privar pacients malalts d'un tractament que podria proporcionar-los alleujament. En la pràctica, seria preferible comparar el grup de tractament que rep el fàrmac experimental amb un grup de pacients que reben la teràpia estàndard tradicional per a aquell tipus de malaltia. Per tant, en la decisió sobre els controls que hem d'utilitzar en un experiment, intervenen consideracions no sols d'índole estadística, sinó també ètiques i fins i tot econòmiques.

Els grups de control són un element molt important del disseny experimental, però la veritat és que la majoria dels investigadors sabem ben poc sobre els controls i com utilitzar-los correctament. No és sorprenent, ja



Paula Navarro

La legislació internacional estableix que l'experimentació amb animals, quan està justificada, ha de fer-se utilitzant el nombre mínim d'exemplars que permeta arribar a conclusions vàlides des del punt de vista estadístic. Totes les fotografies de l'article han estat preses a l'animalari del Servei Central de Suport a la Investigació Experimental de la Universitat de València.

que la majoria dels llibres de text dediquen molt poca atenció a aquest tema. D'una vintena de textos d'estadística i disseny experimental adreçats a biòlegs i psicòlegs publicats des de l'any 2000, més de la meitat ni tan sols tenen l'entrada «control» en l'índex. Dels que sí que inclouen informació sobre controls, la majoria es limita a mencionar-los sense explicar ni els tipus de controls ni la funció que exerceixen en relació amb els objectius de l'experiment. Un text adreçat a biòlegs descriu els controls de la manera següent: «Si un dels tractaments correspon a algun tipus de nivell "natural" d'un factor se'l pot denominar control». Amb explicacions tan críptiques com aquesta no ha de sorprendre'ns que els investigadors tinguin una idea molt superficial, i sovint errònia, del que representa utilitzar grups de control.

Molts pensen, per exemple, que tots els experiments han de tenir un grup de control, quan en realitat no

és així. Imaginem per exemple que volem comparar l'eficàcia de dos fertilitzants, A i B. Si l'única cosa que ens interessa és esbrinar si hi ha diferències entre els dos productes, llavors no hi ha necessitat d'incloure un grup de control. En el millor dels casos, incloure un grup de control compost per plantes a les quals no proporcionem cap fertilitzant únicament ens diria que els productes A i B milloren el creixement de les plantes, una cosa que ja hauríem de saber atès que es tracta de fertilitzants. Al contrari, altres experiments poden requerir no un, sinó uns quants grups de control.

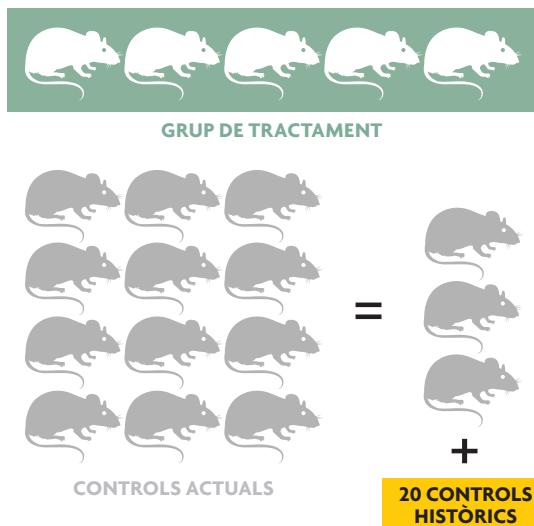
Tampoc és cert que els grups de control hagen de ser necessàriament de la mateixa grandària que els grups de tractament, una idea errònia que fomenten molts programes informàtics per al càlcul de la grandària de la mostra. Un disseny experimental en què tots els grups són igual de nombrosos es coneix com un «disseny equilibrat». En general, els dissenys equilibrats proporcionen millors resultats que els no equilibrats, però hi ha circumstàncies que poden fer preferible utilitzar grups de control més grans o més petits que els grups de tractament (Bate i Karp, 2014; Ruxton i Colegrave, 2011). A més, els grups de control tendeixen a ser intrínsecament menys variables que els grups de tractament, per la qual cosa un disseny equilibrat ben sovint donarà lloc a un problema d'heterogeneïtat de la variància.

### ■ COM PRENDRE PRESTADA INFORMACIÓ D'EXPERIMENTS PREVIS: ELS CONTROLS HISTÒRICS

Una possibilitat per a reduir el nombre d'animals d'un experiment consisteix precisament a reduir la grandària del grup o grups de control. La informació històrica és sempre important a l'hora de dissenyar un experiment: els resultats d'experiments anteriors són la base sobre la qual concebem i dissenyem nous experiments. La nostra proposta consisteix a utilitzar informació extreta d'aquells experiments anteriors, en concret dels grups de control. En el cas d'experiments que es repeteixen de manera rutinària en un laboratori, sempre amb el mateix tipus d'animals, en condicions estàndard i utilitzant els mateixos protocols experimentals, una forma molt eficaç de reduir el nombre d'animals sense sacrificar la capacitat d'arribar a conclusions vàlides consisteix a «prendre prestada» informació dels controls utilitzats en experiments previs. No es tracta d'eliminar el grup de control, sinó de reduir-ne la dimensió incorporant-hi informació de controls d'experiments anteriors, és a dir, de controls històrics.

Per exemple, en un laboratori de toxicologia és possible que es repetesca diverses vegades el mateix procediment experimental, variant únicament la subs-





Utilitzant dades simulades de característiques semblants a les que es manegen en biomedicina trobem que, si els controls històrics són bastant semblants als actuals, la capacitat per a arribar a conclusions vàlides des del punt de vista estadístic és la mateixa utilitzant dotze controls actuals que utilitzant només tres controls actuals i prenent prestada informació (mitjana i variància) de vint controls històrics. Això representa un estalvi de nou animals.

tància tòxica usada, que és distinta en cada experiment concret. És necessari incloure un control de la mateixa grandària que el grup de tractament cada vegada que es repeteix l'experiment? La resposta és un rotund no. Si ja disposem de controls històrics, podem incorporar-los al disseny experimental i això ens permetrà utilitzar un grup de control actual amb menys animals.

Alguns investigadors poden plantejar objeccions perquè pensen que reduir el nombre d'animals que utilitzen en un experiment limitarà les seues possibilitats d'aconseguir un resultat estadísticament significatiu. Però en el nostre treball demostrem, utilitzant simulacions i dades reals de la bibliografia, que aquests inconvenients no estan justificats. En general, els nostres resultats demostren que la incorporació d'informació prestada dels controls històrics augmenta la capacitat per a arribar a conclusions vàlides (i. e., incrementa la potència de les proves estadístiques) alhora que permet reduir el nombre d'animals en el grup de control actual.

La nostra proposta no és en absolut original. Els controls històrics s'utilitzen des de fa molt de temps en la recerca clínica (Viele et al. 2014). No obstant això, no tots els especialistes estan convençuts de la seua utilitat. Alguns rebutgen explícitament l'ús de controls històrics argumentant que en general serà molt difícil

garantir que, si s'observen diferències entre els grups de tractament i de control, aquestes no siguin degudes al fet que les circumstàncies en el moment en què es van arregar les dades del control històric eren diferents de les actuals (Bailey, 2008). Per exemple, un medicament per al tractament de l'asma pot semblar molt eficaç senzillament perquè en el moment en què es va realitzar l'experiment els nivells de contaminació eren molt inferiors als que prevalien en períodes anteriors quan es van arregar les dades històriques.

Aquestes objeccions justifiquen, almenys en part, la mala reputació que per a molts tenen els controls històrics. Però, en els assajos clínics, els controls històrics tradicionalment consisteixen en dades de pacients que els investigadors extrauen de treballs publicats o dels arxius dels hospitals. De vegades ni tan sols es tracta de dades de pacients individuals, sinó de dades agregades procedents d'un conjunt de pacients. La informació així aplegada és molt heterogènia, i sovint no té en compte diferències importants en el diagnòstic o en el tractament dels dits controls històrics, el que podria portar a conclusions errònies. La nostra proposta és més conservadora, ja que recomanem prendre prestada

informació de controls històrics utilitzats en el mateix laboratori, pels mateixos investigadors, amb els mateixos animals i utilitzant protocols experimentals idèntics o molt semblants, un tipus de situació que pensem que és habitual en molts centres d'investigació.

**«UNA POSSIBILITAT PER A REDUIR EL NOMBRE D'ANIMALS D'UN EXPERIMENT CONSISTEIX A REDUIR LA GRANDÀRIA DEL GRUP O GRUPS DE CONTROL.»**

#### ■ QUAN PODEM UTILITZAR CONTROLS HISTÒRICS?

La clau per a utilitzar correctament els controls històrics consisteix a procurar que els controls històrics i els actuals siguin tan semblants com siga possible. En el nostre treball proporcionem recomanacions per a prendre prestada informació de controls històrics en set escenaris distints que corresponen a situacions comunes en la recerca biomèdica. Els escenaris preveuen un ampli ventall de possibilitats, des de situacions en què podem assumir que els controls històrics i els actuals són pràcticament intercanviables (tenen la mateixa mitjana i variància), fins a d'altres en què les diferències entre els controls històrics i els actuals són majors, la qual cosa limita la quantitat d'informació històrica que podem prendre prestada. Els nostres resultats demostren que la inclusió de dades històriques sempre millora la capacitat de l'investigador per a arribar a conclusions vàlides des del punt de vista estadístic (és a dir, incrementa la potència de les proves estadístiques).



Alguns investigadors poden plantejar objeccions perquè pensen que reduir el nombre d'animals que utilitzen en un experiment limitarà les seues possibilitats d'aconseguir un resultat estadísticament significatiu. Però en aquest treball es demostra, utilitzant simulacions i dades reals de la bibliografia, que aquestes objeccions no estan justificades.

Pel que fa a l'estalvi d'animals, aquest depén del tipus d'escenari considerat. Com podíem sospitar, les situacions més restrictives (és a dir, menys semblants entre controls històrics i actuals) permeten un menor estalvi d'animals en el grup de control actual. Però si podem assumir que els controls històrics són molt semblants als actuals, l'estalvi en termes d'animals pot ser considerable. Per exemple, en algunes de les nostres simulacions trobem que tres controls actuals més vint d'històrics proporcionen la mateixa potència estadística que dotze controls actuals, és a dir, la incorporació de dades històriques permet un estalvi d'animals del 75% sense cap minva en la capacitat per a arribar a conclusions vàlides.

Reduir al mínim indispensable el nombre d'animals que participen en un experiment és, a més d'un imperatiu legal, un objectiu desitjable des del punt de vista ètic. A més, reduir la grandària mostral d'un experiment redueix també el cost econòmic i el temps i l'esforç necessari per a dur-lo a terme. Desgraciada-

ment, de les tres erres, la de la reducció és la que menys avanços ha experimentat en les últimes dècades. Les recomanacions que fem en el nostre treball podrien contribuir a millorar la qualitat de la recerca, reduint al mateix temps el nombre d'animals necessaris per a obtenir conclusions vàlides.

Bàsicament, es tracta que els investigadors presten més atenció als grups de control. És necessari que, en compte de considerar-los com un detall purament cosmètic, coneguen els tipus de controls i les funcions que exerceixen en relació amb el disseny experimental. En molts experiments els grups de control són enterament prescindibles. En particular, quan se sotmeten els animals del grup de control a procediments que provoquen patiment i dolor i que tenen resultats enterament previsibles, l'investigador hauria de plantejar-se la possibilitat d'eliminar o reduir dràsticament, per motius ètics, el grup de control. Si és necessari incloure un o més grups de control en l'experiment, llavors els investigadors haurien de buscar estratègies que permeten reduir-ne el nombre, incloent-hi –quan siga raonable fer-ho– prendre prestada informació de controls històrics. ☺

#### REFERÈNCIES

- 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.** Laboratori d'Etologia, Institut Cavanilles de Biodiversitat i Biologia Evolutiva. Universitat de València (Espanya).

**Matthew Kramer.** Statistics Group, Agricultural Research Service (Beltsville). Departament d'Agricultura dels Estats Units.