

Más allá de la reproducibilidad:

Replicabilidad y otras cuestiones relativas a la validez de la investigación

Miguel A. Martínez-Beneito - Área de Desigualdades en Salud, FISABIO-Salud Pública.

7 de junio de 2017

Estructura de la sesión:

- Un vistazo al estado actual de la **reproducibilidad** en el **ámbito científico**.
- **Replicabilidad.**
 - **Definición e ilustración** del concepto de replicabilidad.
 - **Causas** de la crisis de replicabilidad.
 - **Abordaje** de la replicabilidad.

**Un vistazo al estado actual de la reproducibilidad en el
ámbito científico.**

Reproducibilidad en Biostatistics

Reproducible research and Biostatistics

Roger D. Peng

Biostatistics (2009) 10 (3): 405-408. DOI: <https://doi.org/10.1093/biostatistics/kxp014>

Published: 01 July 2009

2.1 Dimensions of reproducibility

The Associate Editor for reproducibility (AER) will handle submissions of reproducible articles. Currently, the AER's involvement with a submission begins only when an article has been accepted for publication. **The AER will consider three different criteria when evaluating the reproducibility of an article.**

1. **Data:** The analytic data from which the principal results were derived are made available on the journal's Web site. The authors are responsible for ensuring that necessary permissions are obtained before the data are distributed.
2. **Code:** Any computer code, software, or other computer instructions that were used to compute published results are provided. For software that is widely available from central repositories (e.g. CRAN, Statlib), a reference to where they can be obtained will suffice.
3. **Reproducible:** An article is designated as reproducible if the AER succeeds in executing the code on the data provided and produces results matching those that the authors claim are reproducible. In reproducing these results, reasonable bounds for numerical tolerance will be considered.

Authors can choose to meet a subset of these criteria if they wish. For example, **an author may choose to release code showing how a particular method is implemented but may not have permission to publish the data. In such a case, the "code" criterion is satisfied, but the "data" and "reproducible" criteria are not.** For authors interested in submitting materials satisfying the "reproducible" criterion, the journal is currently limiting submissions to those whose analyses are conducted using the R software environment. This limitation may change in the future and will generally be dependent on the resources of the journal and the AER. Papers that meet any or all of the above three numbered criteria will be kite marked D, C, and/or R on their title page in the journal.

Se establece **Associated Editor for Reproducibility.**

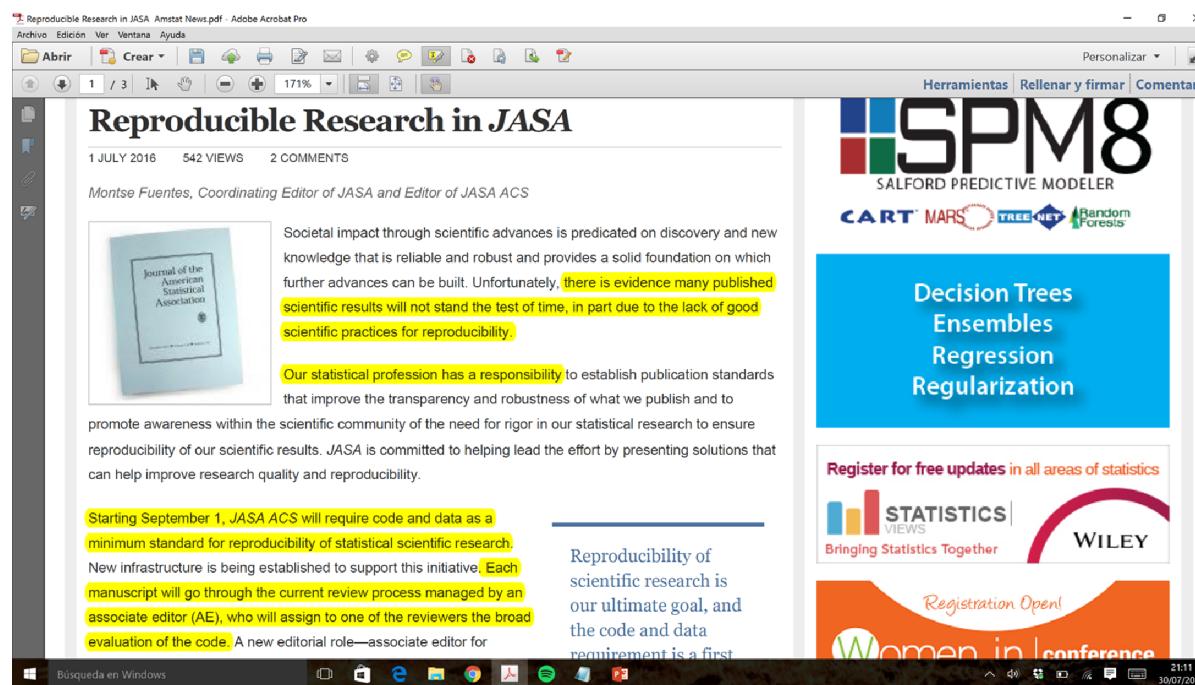
Evaluación de reproducibilidad tras aceptación de trabajos.

Distintos **sellos** de reproducibilidad de trabajos:

- D=Data.
- C=Code.
- R=Reproducibility (D + C + Verificación de resultados).

Reproducibilidad en JASA

Editorial de Journal of the American Statistical Association sobre reproducibilidad (Julio de 2016)



Se requiere aportar **código y datos**. Código será específicamente **evaluado** durante el proceso de revisión **por uno de los revisores**.

Reproducibilidad en Nature

NATURE | EDITORIAL

عربى

Code share

Papers in Nature journals should make computer code accessible where possible.

29 October 2014

A manifesto for reproducible science

Marcus R. Munafó ✉, Brian A. Nosek, Dorothy V. M. Bishop, Katherine S. Button, Christopher D. Chambers, Nathalie Percie du Sert, Uri Simonsohn, Eric-Jan Wagenmakers, Jennifer J. Ware & John P. A. Ioannidis

Nature Human Behaviour 1,
Article number: 0021 (2017)
doi:10.1038/s41562-016-0021
[Download Citation](#)

Published online: 10 January 2017

[where-are-the-data-1.20541](#)) y un manifiesto a favor ([Enero 2017](#))
(<https://www.nature.com/articles/s41562-016-0021>).

NATURE | EDITORIAL

Announcement: Where are the data?

07 September 2016

Editoriales sobre diseminación de código (Octubre 2014)

[\(<http://www.nature.com/news/code-share-1.16232>\), de datos](http://www.nature.com/news/code-share-1.16232) ([Septiembre 2016](http://www.nature.com/news/announcement-where-are-the-data-1.20541))
[\(<http://www.nature.com/news/announcement-where-are-the-data-1.20541>\)](http://www.nature.com/news/announcement-where-are-the-data-1.20541)

Se requiere reportar **disponibilidad de códigos e instrucciones de uso**.

Se requiere **especificar si** los datos del estudio son **accesibles y de qué forma**.

Promueve la publicación y cita de fuentes de datos de la misma forma que se citan otros estudios previos de la literatura.

Reproducibilidad en NEJM



The NEW ENGLAND
JOURNAL of MEDICINE

HOME ARTICLES & MULTIMEDIA ISSUES SPECIALTIES & TOPICS FOR AUTHORS CME >

EDITORIAL

Data Sharing

Dan L. Longo, M.D., and Jeffrey M. Drazen, M.D.

N Engl J Med 2016; 374:276-277 | January 21, 2016 | DOI: 10.1056/NEJMMe1516564

A second concern held by some is that a new class of research person will emerge — people who had nothing to do with the design and execution of the study but use another group's data for their own ends, possibly stealing from the research productivity planned by the data gatherers, or even use the data to try to disprove what the original investigators had posited. There is concern among some front-line researchers that the system will be taken over by what some researchers have characterized as "research parasites."

Reproducibilidad es **más controvertida** en el ámbito médico.

Editorial de Enero de 2016.

Se asocia la **diseminación** de datos con la **proliferación** de "investigadores parásitos".

Se prevé "possible **robo de la productividad** planeada de los recolectores de datos".

El editorial genera enorme **controversia** y duras críticas en tan sólo horas.

Translation to plain English of selected portions of Longo and Drazen's editorial on

data sharing

January 22, 2016

Longo DL, Drazen JM (2016) Data sharing. New England Journal of Medicine 374:276-277.

A second concern held by some is that a new class of research person will emerge — people who had nothing to do with the design and execution of the study but use another group's data for their own ends

Whoever said "If I have seen further, it is by standing on the shoulders of giants" should probably be arrested for stealing those giants' shoulders for their own ends.



EDITORIAL



Data Sharing and the *Journal*

Jeffrey M. Drazen, M.D.

We want to clarify, given recent concern about our policy, that the *Journal* is committed to data sharing in the setting of clinical trials. As stated in the Institute of Medicine report from the committee¹ on which I served and the recent editorial by the International Committee of Medical Journal Editors (ICMJE),² we believe there is a moral obligation to the people who volunteer to participate in these trials to ensure that their data are widely and responsibly used. *Journal* policy will therefore follow that outlined in the ICMJE editorial and the IOM report: when appropriate systems are in place, we will require a commitment from authors to make available the data that underlie the reported results of their work within 6 months after we publish them.

In the process of formulating our policy, we spoke to clinical trialists around the world. Many were concerned that data sharing would require them to commit scarce resources with little direct benefit. Some of them spoke pejoratively in describing data scientists who analyze the data of others.³ To make data sharing successful, it is important to acknowledge and air those concerns.³ In our view, however, research-

ers who analyze data collected by others can substantially improve human health.

We need your help to move medicine forward and improve patient care. Our enemy is disease. By working in collaboration, as we have suggested,³ biologists, data scientists, and clinical trialists can advance the art, and everyone will gain. Clinical trial data are some of the highest quality data in medicine. They should be used responsibly and extensively to help alleviate suffering. We believe that we will all benefit most if this is done collaboratively, but the *Journal's* data sharing policy will apply in all settings.

Disclosure forms provided by the author are available with the full text of this article at NEJM.org.

This article was published on January 25, 2016, at NEJM.org.

1. Committee on Strategies for Responsible Sharing of Clinical Trial Data. Sharing clinical trial data: maximizing benefits minimizing risk. Washington, DC: National Academies Press, 2015.
2. Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data — a proposal from the International Committee of Medical Journal Editors. *N Engl J Med* 2016;374:384-6.
3. Longo DL, Drazen JM. Data sharing. *N Engl J Med* 2016;374:276-7.

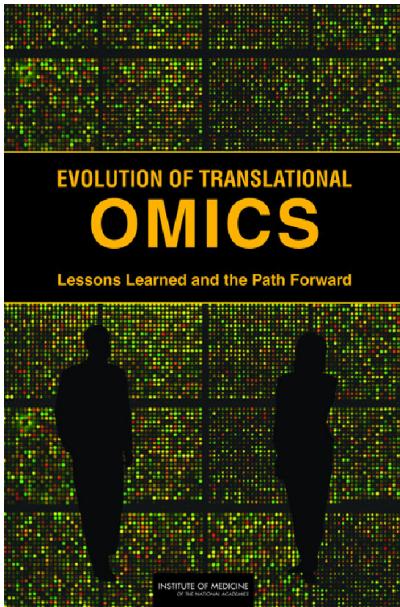
DOI: 10.1056/NEJMoa1601087
Copyright © 2016 Massachusetts Medical Society.

Las críticas obligan a NEJM a publicar un **segundo editorial** en apenas 4 días.

Comprometen a los autores a **compartir datos en 6 meses** desde publicación (medida propuesta por el ICMJE, International Committee of Medical Journal Editors).

Apela a la **obligación moral** con los sujetos de estudio de los ensayos clínicos.

Reproducibilidad en agencias financieras



La Academia Nacional de Medicina (EEUU) (<https://www.nap.edu/catalog/13297/evolution-of-translational-omics-lessons-learned-and-the-path-forward>) formula **recomendaciones** para el desarrollo de **tests** de enfermedades basados en **datos ómicos**.

Entre sus **recomendaciones**: **Publicar datos y metadatos** utilizados. Compartir **código** y procedimientos computacionales utilizados, en particular los asociados al **preproceso** de los datos.

El National Cancer Institutes (NCI) acoge este informe desarrollando una **lista** (<http://www.nationalacademies.org/hmd/Reports/2012/Evolution-of-Translational-Omics/NCI-Checklist.aspx>) de **OBLIGADO CUMPLIMENTO** para las **propuestas de financiación** basadas en datos ómicos.

El NIH "sugiere" que los datos sean **tan accesibles** como sea **posible**.

Obligación de incluir **planes de data sharing** para propuestas de financiación de más de 500.000 dolares.

Reflexiones sobre el tratamiento actual de la reproducibilidad

Reproducibilidad consiste en facilitar **transparencia** de materiales (datos) y procesos (código) en la investigación científica.

Revistas y agencias financieradoras parecen estar **tomando ya cartas** en el asunto.

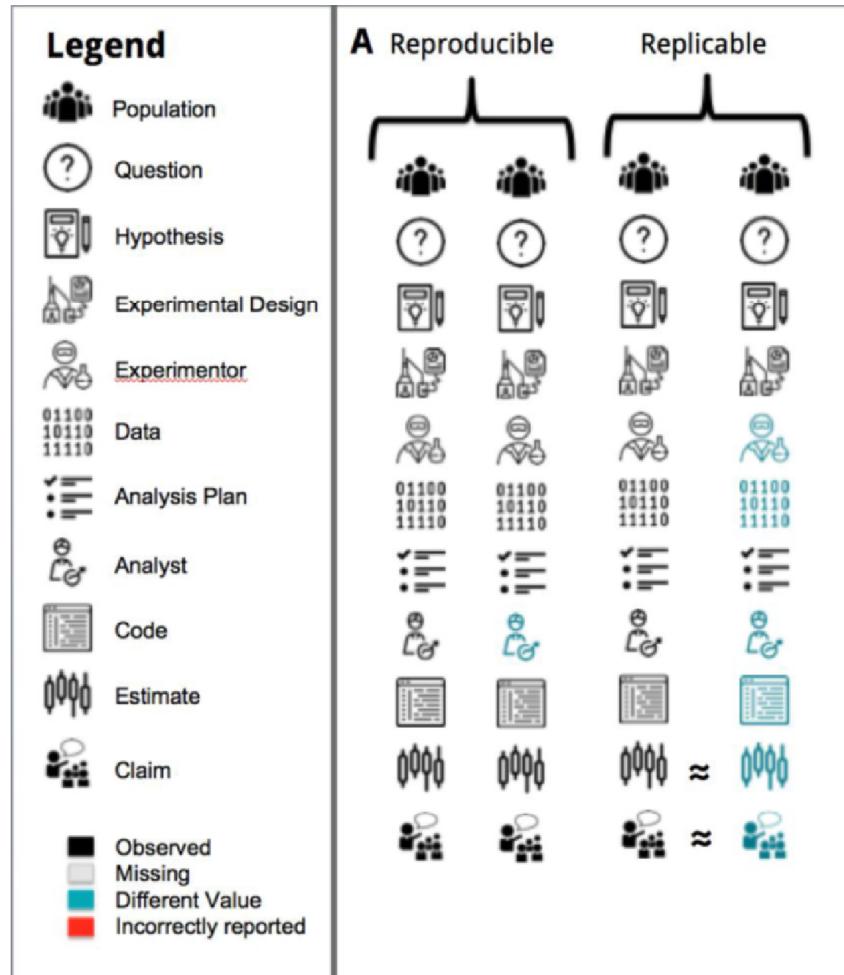
- **Revistas** no empiezan a dar por bueno **hermetismo** en cuanto a datos y códigos.
- **Agencias financieradoras** no admiten que **datos** de estudios que financien benficien **sólo a destinatarios** de las ayudas.

Reproducibilidad es un **tsunami** que se nos viene encima y va a **cambiar** nuestra **forma de trabajar** y compartir nuestro trabajo.

Adaptemos nuestras prácticas para adaptarnos a esta situación que requerirá **cambios en nuestras dinámicas** de trabajo.

Replicabilidad: Definición e ilustración del concepto de replicabilidad.

Reproducibilidad y replicabilidad



Reproducibilidad: Mismos datos, distintos experimentadores, resultados iguales (presumiblemente) -> Transparencia de procesos.

Replicabilidad: Distintos datos, similares condiciones experimentales, resultados similares (constatables) -> Generabilidad de los resultados.

Reproducibilidad y replicabilidad son dos conceptos a menudo confundidos en la literatura.

En ningún caso **reproducibilidad** indica **corrección** de los resultados. "Results that are not reproducible are **hard to verify** and results that do not replicate in new studies are **harder to trust**."

(<http://simplystatistics.org/2017/03/02/rr-glossy/>)

¿También crisis de replicabilidad?

Replication crisis

From Wikipedia, the free encyclopedia

The **replication crisis** (or **replicability crisis**) refers to a **methodological crisis** in **science** in which scientists have found that the results of many **scientific experiments** are difficult or impossible to **replicate** on subsequent investigation, either by independent researchers or by the original researchers themselves.^[1] While the crisis has long-standing roots, the phrase was coined in the early 2010s as part of a growing awareness of the problem.

[Wikipedia \(https://en.wikipedia.org/wiki/Replication_crisis\)](https://en.wikipedia.org/wiki/Replication_crisis)

SCIENCE NEWS | Wed Mar 28, 2012 | 7:09pm BST

In cancer science, many "discoveries" don't hold up



Cancer Research Is Broken

There's a replication crisis in biomedicine—and no one even knows how deep it runs.

By Daniel Engber



[http://www.slate.com/...](http://www.slate.com/)

(http://www.slate.com/articles/health_and_science/future_tense/

<http://uk.reuters.com/article/us-science-cancer-idUSBRE82R12P20120328>

Las crisis de reproducibilidad y replicabilidad han dado lugar a un campo de investigación emergente, **metaciencia**, que se encarga del **estudio científico de la ciencia** en sí misma y los factores que influyen en la validez de sus resultados.

Replicabilidad en Medicina

Contradicted and Initially Stronger Effects in Highly Cited Clinical Research

John P. A. Ioannidis, MD



Context Controversy and uncertainty ensue when the results of clinical research on the effectiveness of interventions are subsequently contradicted. Controversies are most

Ioannidis (JAMA, 2005) (jamanetwork.com/journals/jama/fullarticle/201218) lleva a cabo un **estudio de replicación** de trabajos publicados entre 1990 y 2003 en New England, JAMA, Lancet o revistas médicas especializadas con IF>7.

Criterio de inclusión: **número de citas** hasta agosto de 2004>1000 (análisis paralelo para artículos con menos citas). Se excluyen meta-análisis, revisiones, ...

Busqueda bibliográfica de estudios similares, con criterios de calidad (tamaño muestral, diseño, ...) **similares o superiores**, que pudieran **corroborar o refutar** dichos estudios.

Los **49** artículos encontrados como **elegibles**, se clasifican como:

- **Negative**, no presentaba resultados positivos (**4**, todos refutaban estudios anteriores).
- **Unchallenged**, no se ha valorado su replicabilidad en la literatura (**11**).
- **Contradicted**, el efecto original no se ha podido replicar (**7**).
- **Initially stronger effects**, el efecto original o su duración se reduce a menos de la mitad (**7**).
- **Replicated**, el efecto original ha sido corroborado (**20**).

41.2%(=14/34) de los estudios testados, con efectos positivos, son **refutados** por estudios de similares características.

Los **ensayos clínicos aleatorizados** mostraron **menor** tasa de **discrepancias** (Contradicted+Initially stronger effects), 9 de 39(=23.1%), que el resto de estudios, 5 de 6(=83.3%).

Table 3. Comparison of Characteristics and Citation Counts of Randomized Trials With Contradicted or Initially Stronger Effects vs Those With Replicated or Unchallenged Findings

Characteristic	Contradicted or Initially Stronger Effects (n = 9)	Replicated or Unchallenged (n = 30)	P Value
Published in 1990-1995	8	15	.06
Heart disease topic	4	13	1.00
Sample size, median (IQR)	624 (403-1500)	2165 (892-5201)	.009
All citations received, median (IQR)	1427 (1104-2046)	1542 (1255-2513)	.43
Citations in 3 y, median (IQR)	485 (421-591)	622 (393-825)	.32
Citations per year, median (IQR)	149 (105-215)	214 (146-263)	.07

Abbreviations: IQR, interquartile range.

Entre los ensayos clínicos aleatorizados:

- Estudios **más antiguos** tienen **más** probabilidad de haber sido **refutados** (8/23 vs 1/16).
- Estudios **refutados** tienen **menor tamaño** muestral que el resto (624 vs 2165 individuos en promedio).

En el estudio de trabajos con menos citas los **resultados son similares: 10 de 30(=33.3%)** estudios (corrobórables) fueron **refutados**.

Replicabilidad en Psicología

PSYCHOLOGY

Estimating the reproducibility of psychological science

Open Science Collaboration*†

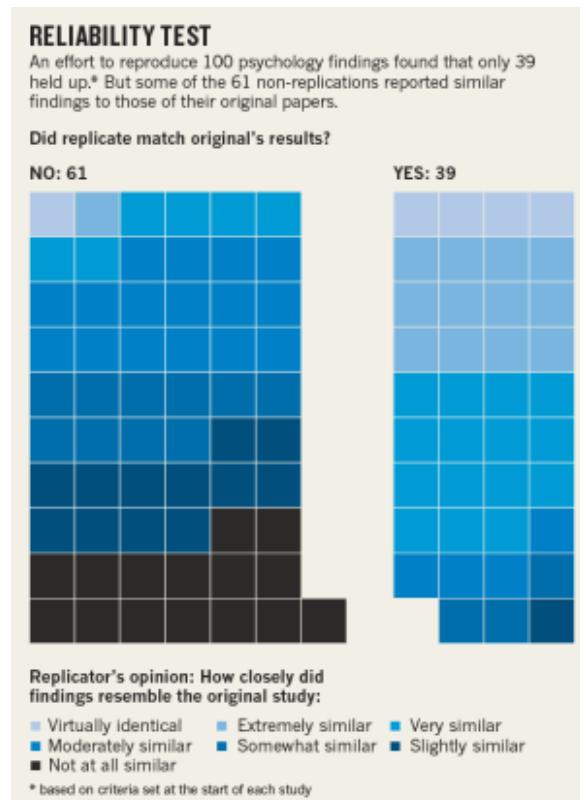
Reproducibility is a defining feature of science, but the extent to which it characterizes current research is unknown. We conducted replications of 100 experimental and correlational studies published in three psychology journals using high-powered designs and original

En 2015 ([Science](http://science.scienmag.org/content/349/6251/aac4716)) (<http://science.scienmag.org/content/349/6251/aac4716>), el Center for Open Science publica los resultados de un **estudio colaborativo de replicación** en psicología.

Mayor esfuerzo de replicabilidad llevado a cabo hasta la fecha. ["One of the top 10 scientific breakthroughs of the year \(Science\)"](https://projects.iq.harvard.edu/psychology-replications)
[\(https://projects.iq.harvard.edu/psychology-replications\)](https://projects.iq.harvard.edu/psychology-replications)

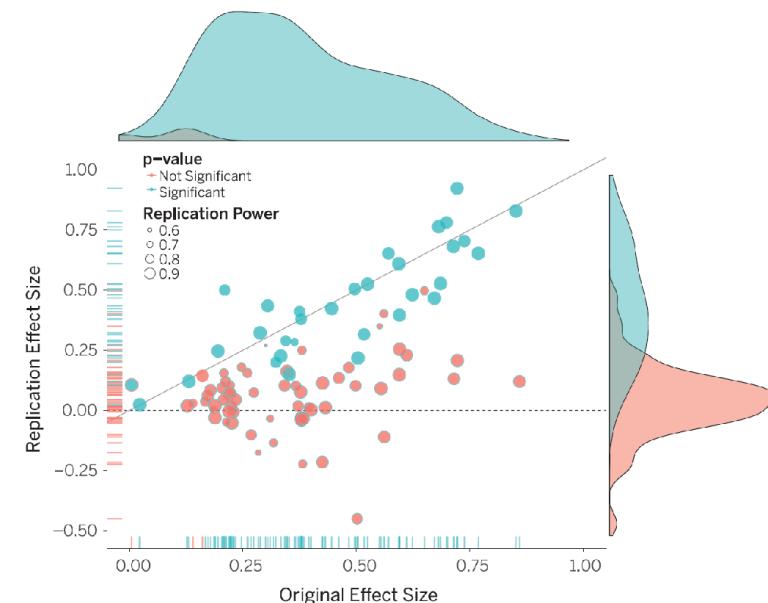
Distintos **grupos** de forma independiente **replican 100 estudios** publicados en la literatura (artículos de 2008 en revistas influyentes).

A diferencia del estudio de Ioannidis, este estudio **replica específicamente** los estudios de la literatura (**evidencia empírica**).



39 de los 100 estudios corroboran los resultados originales.

El problema podría ser **todavía peor** ya que se trata de estudios en revistas "**de prestigio**".



Original study effect size versus replication effect size (correlation coefficients). Diagonal line represents replication effect size equal to original effect size. Dotted line represents replication effect size of 0. Points below the dotted line were effects in the opposite direction of the original. Density plots are separated by significant (blue) and nonsignificant (red) effects.

Sólo parte de los estudios replicados parecen reproducir los efectos originales.

El **efecto promedio** de los estudios replicados fue sólo la **mitad** de lo publicado originalmente.

El "Decline Effect"

Pero ni siquiera la publicación de replicas positivas de trabajos es **garantía** de nada (<http://www.newyorker.com/magazine/2010/12/13/the-truth-wears-off>).

Facciones asimétricas se consideran signo de **mutaciones genéticas**.

Nature **357**, 238-240 (21 May 1992) | doi:10.1038/357238a0; Accepted 17 March 1992

Female swallow preference for symmetrical male sexual ornaments

Anders Pape Møller*

ARTICLE

Send to
 Export citation
 Export reference
...

Møller (Nature, 1991)

(<https://www.nature.com/nature/journal/v357/n6375/pdf/357238a0.pdf>) descubre que **hembras de golondrinas prefieren machos con plumaje más simétrico**.

Simetría de plumaje sería indicador indirecto de **calidad genética**.

Hembras **aplican** este criterio (mecanismo de selección natural) de forma **inconsciente**.

En los 3 años siguientes **9 (de 10)** artículos corroboran la teoría anterior.

Dicha teoría se estudia en **humanos**:

- Mujeres prefieren el **olor** de hombres con facciones **simétricas**, aunque sólo durante su periodo **fértil**.
- Mujeres tenían más **orgasmos** con hombres **simétricos**.
- **Bailarines** simétricos son consistentemente **evaluados** como mejores.

Pero:

- En 1994 sólo **8 de 14** artículos corroboran el efecto original.
- En 1995, **4 de 8** artículos lo corroboran.
- En 1998, **4 de 12** artículos lo corroboran.
- De 1992 a 1997, el **efecto** originalmente encontrado **decae en un 80%**.

Decline effect, estructura general:

Tras proponerse un paradigma, los **procesos de publicación científica tienden a corroborarlo** (sesgo de publicación). **Años después** los incentivos editoriales cambian difundiendo aquellos **resultados que desaprueban** el paradigma establecido.

Decline effect ha sido documentado en muchas **más situaciones**:

- **Efectividad** de fármacos **antipsicóticos** de segunda generación.
- **Verbal overshadowing**: Efecto de la descripción verbal de una sensación o percepción sobre el recuerdo/memorización de dicha percepción. Descripción verbal empeora el recuerdo posterior.
- **Percepción extra-sensorial**: Capacidad de predecir hechos futuros.
- ...

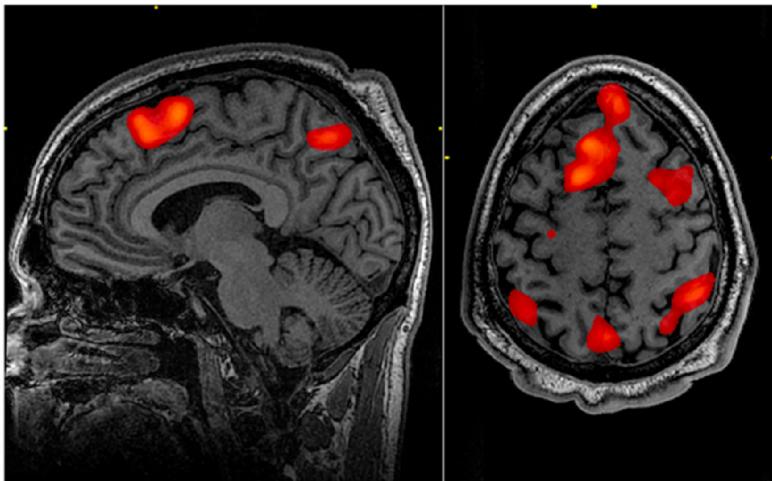
Decline effect es una **expresión más de la crisis** de replicabilidad.

Replicabilidad: causas de la crisis

1.- Deficiencias en los procedimientos

fMRI

fMRI ha sido la herramienta principal para estudiar la **funcionalidad** de cada región cerebral.



An fMRI scan during working memory tasks.

Habitualmente, **individuos** se someten a **estímulos** y, mediante resonancia magnética, se determinan las **regiones del cerebro** con mayor consumo de hemoglobina tras dicho estímulo.

Dichas **áreas** teóricamente serían las **encargadas de procesar** y asimilar el estímulo.

Habitualmente este tipo de estudios tiene

diseño de casos y controles.

Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates

Anders Eklund^{a,b,c,1}, Thomas E. Nichols^{d,e}, and Hans Knutsson^{a,c}

PNAS, 12/07/2016 (<http://www.pnas.org/content/113/28/7900>)

"fMRI is 25 years old, yet surprisingly its most common statistical methods have not been validated using real data. Here, we used **resting-state fMRI data** from 499 healthy controls to conduct 3 million task group analyses. **Using this null data** with different experimental designs, **we estimate the incidence of significant results**. In theory, **we should find 5% false positives** (for a significance threshold of 5%), but instead **we found** that the most common software packages for fMRI analysis (SPM, FSL, AFNI) can result in **false-positive rates of up to 70%**. These results question the validity of a number of fMRI studies and may have a large impact on the interpretation of weakly significant neuroimaging results."

El tratamiento de la **dependencia espacial** empleado en los 3 paquetes de software estudiado no ajusta de forma adecuada la correlación espacial, disparando el número de falsos positivos.

According to many of the headlines that greeted “Cluster failure”, the paper is a devastating bombshell that could demolish the whole field of functional magnetic resonance imaging (**fMRI**):

[Discover](#)

Bug in fMRI software calls 15 years of research into question
(Wired)

A bug in fMRI software could invalidate 15 years of brain research. This is huge. (ScienceAlert)

New Research Suggests That Tens Of Thousands Of fMRI Brain Studies May Be Flawed (Motherboard)

<http://blogs.discovermagazine.com/neuroskeptic/2016/07/07/false-positive-fmri-mainstream/#.WQZCLNryiM8>

Este trabajo ha supuesto un auténtico **terremoto** dentro del campo de fMRI.

Se estima que unos **3500 artículos** podrían estar **afectados** (alrededor del 9% de la literatura del campo).

http://blogs.warwick.ac.uk/nichols/entry/bibliometrics_of_cluster/

2.- Conflictos de intereses

Conflictos de intereses **distorsionan la literatura** científica, sesgando la literatura científica en **direcciones interesadas**.

No se publica en función de la **evidencia** sino de la **conveniencia**.

Conflictos de interés están, en parte, **detrás de la crisis** de replicabilidad.

Tipos de conflictos de interés en publicaciones científicas:

- Intereses **comerciales, políticos y sociales**.
- Presión por publicar de los **científicos**.
- Interés de las **revistas** por publicar resultados novedosos, rompedores.

Intereses comerciales, políticos y sociales.

Investigación científica ofrece un estupendo **argumento a decisiones o intereses** arbitrarios que de otra forma serían difíciles de justificar.

La **aureola de veracidad** de la ciencia **legitima decisiones** políticas y sociales ("evidence based policy") o intereses comerciales.

Grupos de presión, asociaciones, lobbies ... buscan **sustento** en **literatura científica**.

El **problema** surge cuando la ciencia **no** se usa para **guiar** las decisiones sino para **justificarlas**.

"**Scientists** working in the policy arena are often **naïve** about the impact of their findings" (https://en.wikipedia.org/wiki/Science_on_the_Verge)

"Instituto Danone para la Nutrición y la Salud"
(http://www.institutodanone.es/ayudas_proyectos_investigacion_cientifica.html)
ofrece **ayudas a la investigación**:



AYUDAS A PROYECTOS DE INVESTIGACIÓN CIENTÍFICA

Con el fin de **fomentar la actividad científica, y apoyar particularmente a los grupos de investigación más jóvenes**, el Instituto Danone convoca desde 2014 dos Ayudas de 40.000€ cada una para promover la labor de investigación científica en las áreas de la Alimentación, Nutrición y Salud.

Se trata de **ayudas de dos años de duración** para la realización de Proyectos de Investigación Científica sobre Alimentación, Nutrición y Salud en sus **aspectos básico, clínico y aplicado**.

[BASES DE LA CONVOCATORIA 2017](#)

[FORMULARIO DE SOLICITUD](#)

Leyendo la **convocatoria** con más detalle uno se de cuenta:

1. Se aceptarán **aquellas propuestas sobre Alimentación, Nutrición y Salud** en sus aspectos básico, clínico y aplicado a realizar en universidades, instituciones o institutos de investigación públicos o privados en España, y en relación con cualquiera de los siguientes campos de actividad: nutrición clínica, metabolismo, epidemiología, genética y biología molecular, y ciencia y tecnología de los alimentos. **En la concesión de una de las dos ayudas se dará preferencia a las propuestas que, teniendo en cuenta lo anterior, tengan como objetivo el estudio de las propiedades y efectos positivos para la salud del consumo de yogur.**

Financiación se destina a estudiar **propiedades beneficiosas** de sus productos (para que posiblemente engrosen la literatura científica).

La **intencionalidad** de estos estudios es evidente. Se busca **sustento** no conocimiento.

+ Danone

Actimel es uno de los productos estrellas de Danone con una facturación anual de alrededor de 5300 millones de dolares (**25% de la facturación del grupo**) (<https://scientiablog.com/2012/06/14/la-verdadera-historia-del-actimel-i/>).

Se trata de un **alimento funcional** con unas supuestas propiedades beneficiosas para la salud ("mejora tus defensas").

Su principal propiedad, reducción de diarreas, se atribuye a la presencia de una cepa patentada de **Lactobacillus Casei Imunitass**.

Sus **efectos positivos** se sustentan (supuestamente) en **literatura** científica generada al efecto. De ahí la **importancia comercial** de disponer de literatura científica que permita aducir estas propiedades.

La European Food Safety Administration (**EFSA**) emite en 2010 un **informe** sobre los **efectos** de este producto sobre la salud.

El informe "**rechazó** los más de **20 trabajos** de investigación que pretendían avalar los beneficios del Actimel" (<https://scientiablog.com/2012/06/14/la-verdadera-historia-del-actimel-i/>).

"The Panel concludes that the **evidence** provided is **insufficient to establish a cause and effect relationship** between the consumption of Actimel and a reduction of the risk of *C. difficile* diarrhoea" (<http://www.efsa.europa.eu/en/efsajournal/pub/1903>)

La EFSA destaca errores repetidos de procedimiento en los estudios evaluados.

Las **propiedades beneficiosas** publicitadas de Actimel, y otros prebióticos, se sostienen (y por tanto permiten) sólo por la presencia de vitamina B6 (<https://scientiablog.com/2012/06/18/la-verdadera-historia-del-actimel-ii-me-siento-enganado-por-danone/>) en su composición (aunque su dosis es, por ejemplo, un tercio inferior a la de un plátano).



En cualquier caso el halo ilusorio de **credibilidad de la ciencia** sigue presente en el producto.

Presión por publicar de los científicos

La ciencia vive una **proliferación "industrial"** de datos y de publicaciones que los explotan. Producción científica tratada como **producto de consumo** disminuye su calidad.

El sistema académico considera el **número de publicaciones** como uno de sus principales parámetros para valorar a sus integrantes. Científicos son en gran parte **evaluados por su volumen** de publicaciones.

Sería conveniente poner el **foco en la calidad** de las publicaciones más que en su cantidad.

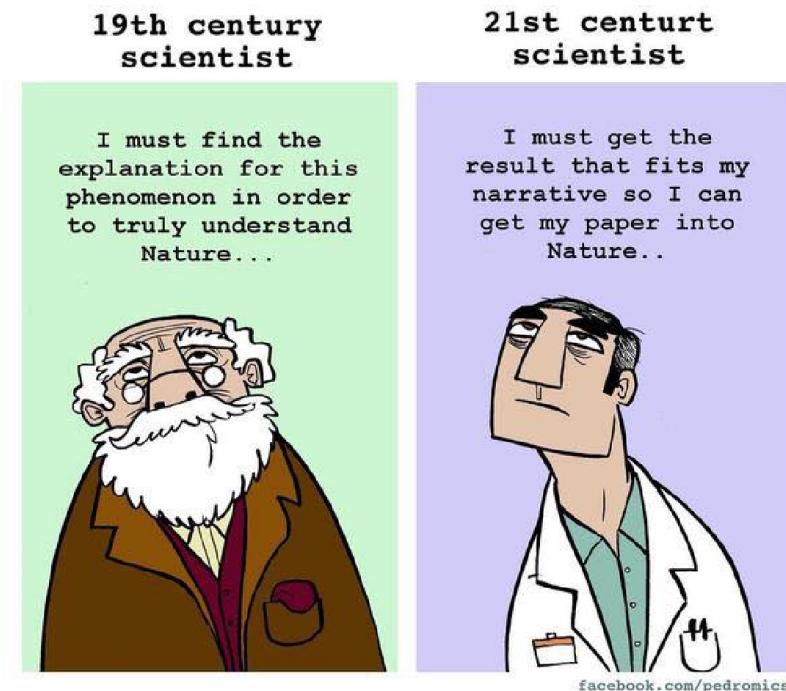
La **presión**, o simplemente el interés, por **publicar** supone un claro **conflicto de interés** (más allá de lo económico) de todas las publicaciones.

Presión de las revistas por publicar resultados novedosos

Además, las revistas exigen resultados novedosos: "Results need to be exciting, eye-catching, even implausible ... When we review papers we're often making authors prove that their findings are novel or interesting. We're not making them prove that their finding are true" (<http://www.nature.com/news/replication-studies-bad-copy-1.10634>)

Esta presión se vive fundamentalmente en aquellas **revistas de mayor impacto**, aquellas que influyen en mayor medida en la literatura.

Si los **datos** pueden hablar por sí solos también **pueden ser torturados** hasta que confiesen ...



facebook.com/pedromics

3.- Sesgo de publicación

No todos los estudios ven la luz de la **publicación científica** con la misma probabilidad.

Resultados **no significativos**, o en contra de la hipótesis inicial de los investigadores, tienden a **no ser publicados**. Este efecto se conoce como "**file drawer problem**" (https://en.wikipedia.org/wiki/Publication_bias).

Este efecto puede **alterar el corpus bibliográfico** ocultando parte de la evidencia científica disponible y por tanto **sesgando el conocimiento acumulado** en ciertas áreas de investigación.

Sesgo de publicación y el **decline effect**: "Another reason may be the publication bias: scientists and scientific journals prefer to publish positive results of experiments and tests over null results, especially with new ideas. As a result, the journals may refuse to publish papers that do not prove that the idea works. Later, when an idea is accepted, journals may refuse to publish papers that support it." (https://en.wikipedia.org/wiki/Decline_effect)

Sesgo de publicación y acupuntura

Revisión sistemática de la literatura sobre la efectividad de acupuntura, de 1966 a 1995, para un grupo de países (Vickers et al., 1998)

Table 1 Results of Controlled Clinical Trials of Acupuncture by Country of Research

Country	Total Trials Analyzed	Favoring Test Treatment	
		Number	Percentage
USA	47	25	53
China	36	36	100
Sweden	27	16	59
UK	20	12	60
Denmark	16	8	50
Germany	16	10	63
Canada	11	3	27
Russia/USSR	11	10	91
Austria	9	8	89
Italy	9	8	89
Australia	6	1	17
France	6	5	83
Taiwan	6	6	100
Japan	5	5	100
Finland	4	2	50
Hong Kong	3	3	100
Netherlands	3	1	33
New Zealand	3	2	67
Poland	3	2	67
Switzerland	3	1	33
Bulgaria	2	2	100
Brazil	1	1	100
Croatia	1	1	100
Israel	1	1	100
Nigeria	1	1	100
Sri Lanka	1	0	0
Vietnam	1	1	100
Total	252	171	68

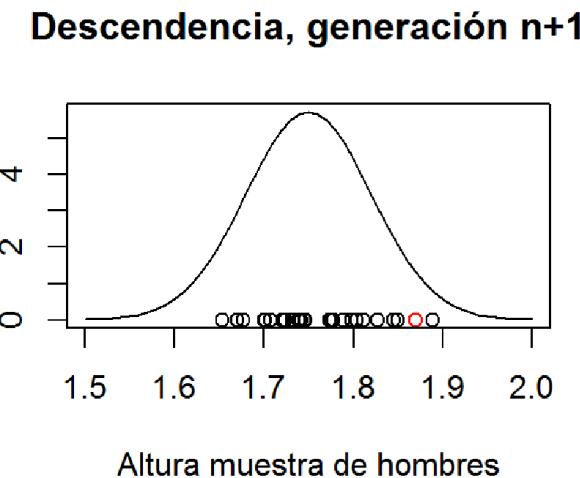
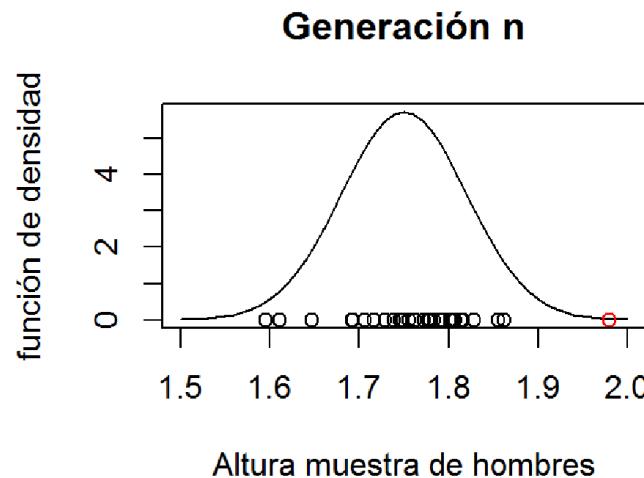
(<https://pdfs.semanticscholar.org/2b01/5f2b2e09f6b06ab236a30adf6eda72c5f6c2.pdf>).

Los países **asiáticos** reportan resultados significativamente **favorables** a la acupuntura en **61 de 63** estudios (96.8%).

Paises **no asiáticos** reportan resultados significativamente **favorables** a la acupuntura en **110 de 191** estudios (57.6%).

4.- Regresión a la media

Es el fenómeno estadístico por el que los **individuos con observaciones extremas** tenderán a estar **más cerca a la media** de la población en posteriores observaciones.



Este fenómeno de "regresar" hacia la media **dio nombre** originalmente a los modelos de regresión.

Estudios son **muestras de tamaño 1** del universo de posibles estudios. Sacamos **conclusiones a partir de una única muestra**.

Tomando una **muestra lo suficientemente extrema** podremos demostrar virtualmente **cualquier resultado**. Un 5% de las muestras posibles, simplemente por azar, nos deberían dar un resultado significativo (aunque no existiera efecto).

Así, muchos **hallazgos "significativos"** pueden ser simplemente **outliers estadísticos** (muestras suficientemente anómalas) que al intentar ser replicados **pierden su excepcionalidad**.

Regresión a la media podría estar también **detrás del decline effect**:

-**Conflictos de intereses** hacen que outliers estadísticos no pasen desapercibidos y **acaben publicados**.

-**Sesgo de publicación evita** que dichos resultados anómalos sean **contrarestados** inicialmente, aunque sí posteriormente cuando ya no son tan novedosos.

Abordaje de la crisis de replicabilidad

Pre-registro de estudios

Procedimiento adoptado de **ensayos clínicos**: todos los procedimientos de análisis han de ser **declarados de antemano** a la recolección de datos de análisis. Center for Open Science (<https://cos.io/rr/>), por ejemplo, ofrece esta posibilidad.

PUBLISHED: 10 JANUARY 2017 | VOLUME: 1 | ARTICLE NUMBER: 0034

editorial

Promoting reproducibility with registered reports

Authors who wish to publish their work with us have the option of a registered report. With this format, acceptance in principle happens before the research outcomes are known. As a result, publication bias is neutralized, as are incentives for practices that undermine the validity of scientific research.

Across the social and biomedical sciences, there is a pressing realization that, in order to increase the reliability of scientific results, we need to change how research is carried out, evaluated, reported, and incentivized. Ten leading advocates for the reform of scientific practices set out 'A manifesto for

```
graph LR; A[Develop idea] --> B[Design study]; B --> C["Write up introduction, methods, analysis plan"]; C --> D[Collect and analyse data]; D --> E["Write up results and discussion"]; E --> F[Publish report]; C -- "Stage 1 peer review" --> D; E -- "Stage 2 peer review" --> F;
```

<https://www.nature.com/...>

(<https://www.nature.com/articles/s41562-016-0034>)

Una vez **aprobado** el registro previo del estudio por un grupo de evaluadores **se garantiza la publicación** de los resultados del estudio.

Triple objetivo:

- Limita **intervención interesada** de experimentadores en el análisis.
- Fomento publicaciones con **métodos rigurosos y bien documentados**.
- **Disminución de sesgo** de publicación (fomenta publicación de resultados negativos).

Atención específica a la replicabilidad en revistas

Estudios de replicación son **menos atractivos** que estudios originales.

Revistas deberían promover estudios de replicación para vigilar la **calidad de sus contenidos**.

"Perspectives in Psychological Science" ha creado una **sección exclusiva** para estudios de replicación. (<http://www.newyorker.com/tech/elements/the-crisis-in-social-psychology-that-isnt>)

Social Psychology, **número monográfico** sobre replicación de estudios de importancia particular en este ámbito.

(<http://www.newyorker.com/tech/elements/the-crisis-in-social-psychology-that-isnt>)

Algunas revistas (Psychological science) **anima** específicamente a la publicación de **estudios con protocolos pre-registrados**.

(https://en.wikipedia.org/wiki/Replication_crisis)

Atención específica a resultados negativos

Published online 23 February 2011 | Nature 470, 437 (2011) |
doi:10.1038/470437a

[https://www.nature.com/...](https://www.nature.com/)

Column: World View

Unpublished results hide the decline effect



Some effects diminish when tests are repeated.
Jonathan Schooler says being open about findings that don't make the scientific record could reveal why.

Jonathan Schooler

(<https://www.nature.com/news/2011/110223/full/470437a.html>)

Propuesta de **base de datos** open-access de **resultados negativos**.

PsychFileDrawer.org: **Web** donde psicólogos pueden mandar **resultados no publicados**, sean significativos o no. (<http://psychfiledrawer.org/>)

Revistas específicas dedicadas a la publicación de **resultados negativos**:

- Journal of Negative Results.
- Journal of Negative Results in Biomedicine.

Financiación específica a estudios de replicación

La introducción de **replicabilidad** en la agenda de las **agencias de financiación** puede hacer significar una gran diferencia. "A mere 3% of scientific funding devoted to replication could make a big difference. The current amount is, he says, near zero" (Nature, 2015). (<http://www.nature.com/news/over-half-of-psychology-studies-fail-reproducibility-test-1.18248>)

The Netherlands Organisation for Scientific Research en 2016 ha lanzado una **línea de financiación específica** para **replicar** estudios de gran impacto en ciencias sociales y salud.

Sociedades científicas (o las propias revistas científicas) podrían también **incentivar/premiar** estudios de **replicación de sus actividades**, revistas que auspician ... Serían las más interesadas en validar la literatura del campo correspondiente.

Formación en replicabilidad

Replicación poco agradecida. Todos **preferimos un esfuerzo creativo** más que confirmatorio de los resultados de otros investigadores. Este hecho **debería ser combatido** en bien de la propia ciencia.

Hay pocos incentivos para practicar replicabilidad: estudios de replicabilidad invitan a sus investigadores a **ser criticados** y ayudan poco a mejorar su historial de publicaciones (<http://www.nature.com/news/replication-studies-bad-copy-1.10634>)

Resulta necesario **transmitir** a futuros investigadores la **importancia de replicar** estudios frente al desarrollo de investigación. Replicación es tarea que **nos compete a todos**.

Sugerencias de publicación de algún estudio de replicación en el ámbito de la **tesis doctoral** antes de que esta pueda ser leída.
(<http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01152/full>)

Sugerencias de promoción de trabajos de replicabilidad en **Trabajos de Fin de Máster** con carácter multicéntrico o multidisciplinar.

Para concluir

En materia de replicabilidad, como en tantos otros temas, "**lo difícil no es llegar si no mantenerse**".

Se pretende la generación de **resultados sólidos** que "pasen el **test del tiempo**" ("Many scientific results will not stand the test of time", Editorial de JASA).

El **método científico** debería dar las **bases** para que así sea aunque, por desgracia, en el entente actual es **menos efectivo de lo deseable**.

