

[Home](#)[Contenido](#)[jameslindlibrary.org](http://jameslindlibrary.org)

## Interpretación de comparaciones no prejuiciosas:

### Cómo considerar la obra de la casualidad

Cuando se comparan dos tratamientos, cualquier diferencia en los resultados puede deberse, simplemente, a la casualidad. Tomemos, por ejemplo, la comparación de un nuevo tratamiento con un tratamiento estándar en el que 4 personas mejoraron con el primero y 6 mejoraron con el segundo. Deducir con certeza que el nuevo tratamiento fue peor que el estándar sería, claramente, un error, dado que estos resultados podrían sólo reflejar la obra de la casualidad. Si se repitiera la comparación, la cantidad de pacientes que mejoraran podría revertirse (6 contra 4) o podría ser la misma (5 contra 5) o la relación podría ser otra.

Sin embargo, si 40 personas mejoraran con el tratamiento nuevo y 60 con el estándar, sería menos probable que la casualidad pudiera explicar la diferencia. Y si 400 personas mejoraran con el nuevo tratamiento y 600 con el estándar, quedaría claro que el nuevo tratamiento tendría muchas probabilidades de ser peor que el tratamiento estándar. La manera de reducir las probabilidades de que la casualidad nos impulse al error al comparar tratamientos es, por lo tanto, asegurarse de que las pruebas auténticas incluyan cantidades suficientemente grandes de personas que muestren los criterios de valoración que uno espera que influyan, como la mejoría o el deterioro.

En determinadas circunstancias, para obtener cálculos confiables de los efectos del tratamiento, es necesario que participe un número muy grande de personas: miles y, a veces, decenas de miles. Se necesitan grandes cantidades de participantes, por ejemplo, si los criterios de valoración que nos interesan son raros, como podría ser el caso de infartos y accidentes cerebrovasculares en mujeres de edad mediana, aparentemente sanas, que utilizan la terapia de reemplazo hormonal (TRH). También se necesitan grandes números de personas si se desea detectar con certeza efectos moderados pero importantes de los tratamientos, por ejemplo, la reducción en un 20 por ciento del riesgo de muerte temprana en personas con infarto.

Para evaluar el papel que pudo haber desempeñado la casualidad en los resultados de pruebas auténticas, los investigadores utilizan "pruebas de significación estadística". Cuando los estadísticos y otras personas se refieren a "diferencias significativas" entre los tratamientos, normalmente se refieren a la significación estadística. Las diferencias estadísticamente significativas entre tratamientos no necesariamente tienen importancia práctica. Pero, de todos modos, las pruebas de significación estadística son importantes porque nos ayudan a evitar que erróneamente deduzcamos que existen diferencias reales entre los tratamientos cuando, de hecho, no las hay. Esta deducción equivocada es lo que a veces llamamos error tipo I.

También es importante tener en cuenta una cantidad suficientemente amplia de resultados del tratamiento para evitar un peligro mucho más común: deducir que no existen diferencias entre los tratamientos cuando, de hecho, sí las hay. Esta deducción equivocada es lo que a veces llamamos error tipo II. Thomas Graham Balfour era consciente de este tipo de riesgo cuando interpretó los resultados de su prueba del potencial de la belladonna para evitar que los huérfanos a su cuidado desarrollaran fiebre escarlata ([Balfour 1854](#)). En comparación con 2 de 75 niños que no recibieron *belladonna*, 2 de 76 niños asignados al grupo que sí la recibía desarrollaron fiebre escarlata. Balfour observó que "las cantidades son demasiado pequeñas para justificar las deducciones respecto del potencial profiláctico de la *belladonna*". Si hubieran sido más los niños que desarrollaron fiebre escarlata, quizá Balfour hubiera podido arribar a una conclusión más certera sobre los posibles efectos de la *belladonna*. En cambio, sólo observó que 4 casos de fiebre escarlata entre 151 niños era un número demasiado pequeño para llegar a una conclusión confiable.

Un método que reduce la probabilidad de que los efectos de la casualidad nos engañen es estimar un rango de diferencias entre los tratamientos dentro del cual deberían caer las diferencias reales ([Gavarret 1840](#); Huth 2006). Estos cálculos de los rangos se conocen como intervalos de confianza. Como se ilustra en el párrafo inicial de este ensayo, es probable que la repetición de una comparación entre tratamientos arroje cálculos variados de los efectos diferenciales de los tratamientos sobre los criterios de valoración, en particular si los cálculos se hacen en función de cantidades pequeñas de dichos criterios. Los intervalos de confianza tienen en cuenta esta variación, son más informativos que las meras pruebas de significación estadística y, por lo tanto, más útiles para reducir la probabilidad de que la intervención de la casualidad nos conduzca a error.

Las pruebas estadísticas y los intervalos de confianza —ya sea para el análisis de estudios individuales o en el [metanálisis](#) de una cantidad de estudios diferentes pero similares— nos ayudan a determinar la dimensión de la influencia de la casualidad y a no deducir que los efectos y las diferencias entre los tratamientos existen cuando no los hay o no existen cuando sí los hay.

**Cite as:** Editorial commentary (2007). Cómo considerar la obra de la casualidad. The James Lind Library ([www.jameslindlibrary.org](http://www.jameslindlibrary.org)).

**Show JLL records:** illustrating [taking account of the play of chance](#)

**Next essay:** [Identificación de los efectos no previstos de los tratamientos](#)

**Select essay:**

## Referencias

Balfour TG (1854). Quoted in West C. Lectures on the Diseases of Infancy and Childhood. London, Longman, Brown, Green and Longmans, p 600.

Gavarret LDJ (1840). Principes généraux de statistique médicale: ou développement des règles qui doivent présider à son emploi. Paris: Bechet jeune & Labé.

Huth EJ (2006). Jules Gavarret's Principes Généraux de Statistique Médicale: a pioneering text on the statistical analysis of the results of treatments.

[Home](#)

[Contenido](#)

[Comments welcome](#)