



## La investigación, la estadística y la tiranía de los valores-p

Ioseba Iraurgi e Iñaki Markez

**C**on bastante frecuencia nos encontramos con colegas clínicos que tras haber realizado un esfuerzo considerable en la ejecución de un estudio para valorar los resultados de un determinado tratamiento o terapia se muestran decepcionados porque las pruebas estadísticas utilizadas para la comparación de los grupos tratados no han mostrado significación estadística, es decir, no han aparecido valores de probabilidad ('valores-p') con el suficiente número de ceros tras la coma. En estos casos, en la mayoría de las ocasiones, la decisión es abandonar el estudio. También es bastante frecuente que cuando revisamos un artículo científico vayamos inmediatamente a las tablas de datos y miremos con expectación cuántos asteriscos o cuántos ceros asociados a los valores-p aparecen en las tablas. Si por casualidad aparecen muchos ceros o asteriscos aparece una especie de clarividencia que nos anima a pensar que ese estudio está muy bien, nos informa de resultados importantes y es recomendable leerlo y considerarlo para nuestras futuras decisiones. Si la significación estadística alcanzada está cercana al valor de referencia estándar de  $p = 0,05$ , entonces nuestro entusiasmo ya no es tan exaltado, y tendemos a pensar que, bueno, los resultados son válidos, pero no tan espectaculares como los que estarían asociados a un valor-p, por ejemplo, de 0,001. Por otra parte, es raro que en la revisión

de artículos de revistas científicas nos encontremos con resultados donde las pruebas estadísticas muestren valores-p superiores a 0,05 ya que los editores tienden a no considerar aquellos estudios 'no significativos'.

Nos encontramos, entonces, que en la valoración de nuestros hallazgos, en la elección de los artículos que merecen ser leídos y en la consideración de los resultados que merecen ser aplicados a la práctica cotidiana nos hallamos subyugados a los valores-p, lo cual resulta muy preocupante. Hacemos de ellos una interpretación un tanto exotérica, al asociar un mayor número de ceros tras la coma —o de asteriscos— a una mayor fuerza y validez de los resultados, de modo que los aceptamos como 'evidencia probada'. Y como muchos de los lectores ya saben, o empiezan a intuir, nada más lejos de la verdad. El sentido común y, sobre todo, nuestra experiencia y conocimiento nos dicen que lo clínicamente relevante es una cosa, y lo estadísticamente significativo es otra. Pero en el ámbito de la comunicación científica tendemos a confundirlos, y priorizamos la herramienta metodológica —la estadística— sobre las bases teóricas y prácticas.

Tal es el punto de utilización perversa de la estadística que han aparecido innumerables opiniones y reflexiones sobre la misma, tales como las que se reflejan en la tabla I.

**Tabla 1. Reflexiones y citas sobre ‘la estadística’ y ‘las estadísticas’**

- Las estadísticas son al mundo lo que el poste al borracho; sirven para sostenerlo, no para mejorarlo (Cantervill)
  - Las estadísticas son como las minifaldas: te dan algunas ideas, pero esconden lo más importante (Ebbe Skovdahl)
  - Las estadísticas me producen el mismo efecto que las minifaldas: muestran lo atractivo, ocultan lo vital (Doris Band)
  - Hay mentiras, malditas mentiras y estadísticas (Mark Twain)
  - La estadística es una ciencia según la cual todas las mentiras se tornan cuadros (Dino Segre Pitigrilli)
- 
- La estadística es como una mujer con bikini, muestra casi todo, pero lo fundamental lo oculta... Olvidaba decir que lo fundamental es... lo que yo quiera creer (Robert Frost)
  - La estadística es una ciencia que demuestra que si mi vecino tiene dos coches y yo ninguno, los dos tenemos uno (George Bernard Shaw)
  - La estadística ha demostrado que la mortalidad de los militares aumenta perceptiblemente durante tiempos de guerra (Alphonse Allais)
  - La estadística es la primera de las ciencias inexactas (Edmond Hout de Goncourt)

Ninguno de los autores de estas citas era estadístico o matemático, ni tampoco médico o psicólogo, pero sí encontramos desde ilustrados dramaturgos, hasta periodistas e incluso entrenadores de fútbol. En las cinco primeras citas, las opiniones se refieren básicamente a los resultados (‘las estadísticas’) y no al método (‘la Estadística’) basado en teorías matemáticas. Y aquí vendría a cuento otra cita anónima: *‘las cifras no mienten, pero los mentirosos también usan cifras’*. No hay que confundir el método con el uso que se hace de él; la estadística es una herramienta, una buena herramienta, que bien utilizada puede ayudarnos en la toma de decisiones allí donde hay incertidumbre. Recordemos a Descartes en su ‘Discurso del método’: *‘si no está en nuestro discernir las mejores opciones, debemos elegir las más probables’*; y ahí es donde puede asistirnos la estadística. Siguiendo con el juego de citas, en las cuatro últimas hay quienes se atreven a opinar sobre el propio método, planteando poco menos que se trata de una ciencia de la obvedad. Quizá la opinión menos errática es la que hace el escritor francés del siglo XIX E. Hout de Goncourt, quien al describirla como ciencia inexacta, sin

saberlo bien, estaba acertando con su objeto: donde no hay certeza (exactitud) bien cabe aproximarse desde lo que es posible, probable; y es que la estadística tiene su fundamento en la teoría de la probabilidad.

En cualquier caso, respecto al valor que hay que conceder a las opiniones recogidas, responderemos con otra cita del gran Dylan: *‘No critiques lo que no puedes entender’*. Y, precisamente, para poder criticar con entendimiento, para que aquellos clínicos que no siendo expertos en metodología tengan la necesidad de utilizar la estadística y puedan manejarse sabiendo lo que ésta implica, se inicia en este número de la revista *Norte de salud mental* una serie de artículos sobre la Evaluación de Resultados Clínicos, siendo el primero de ellos un artículo de revisión que trata de diferenciar lo que representa un resultado estadísticamente significativo en contraposición a lo que es clínicamente significativo, relevante o importante. Los hallazgos de una investigación pueden ser ciertos y notables, y, sin embargo, no alcanzar significación estadística; pero también pueden ser irrelevantes y resultar estadísticamente muy signi-



ficativos. La estadística es un instrumento para la determinación de cuán probable es que los resultados encontrados no sean debidos al azar sino a la acción de una variable que actúa como causa de otra (un agente infeccioso respecto a una enfermedad, un tratamiento respecto al logro de salud, etc.) y que ha sido puesta a prueba en el estudio. En ese artículo se repasa lo que está implícito en una investigación y en el proceso de inferencia estadística; los errores en la interpretación a los que nos podemos someter y cómo minimizarlos; lo que diferentes protagonistas del ámbito clínico interpretan como 'resultado clínicamente relevante' y cómo diferenciarlo del concepto 'resultado estadísticamente significativo'. Se trata de un artículo cuyo objetivo es aclarar conceptos y sentar algunas bases que posteriormente reaparecerán cuando afrontemos algunos métodos específicos en el análisis de resultados clínicos.

La investigación clínica es un requerimiento necesario e ineludible, tanto para avanzar en el conocimiento de las mejores acciones clínico-terapéuticas como para actualizar su aplicación en los recursos asistenciales ordinarios. A este respecto, en aras a evaluar de forma precisa y válida la eficacia de un determinado tratamiento, el diseño de elección es el ensayo clínico aleatorio, el cual compara grupos de participantes sometidos a diferentes opciones de tratamiento y donde una de ellas es la que se evalúa como novedosa o preferente a las alternativas. Este tipo de estudios, por su com-

plejidad y costes, no está al alcance del clínico que ejerce su labor asistencial en lo que coloquialmente conocemos como 'las trincheras'. Requiere la formación de grupos muy controlados en su selección y meticulosamente seguidos en su evolución, cuestiones raramente habituales en la práctica cotidiana; atendemos lo que nos toca a demanda y, en muchas ocasiones, las altas no son producidas por haberse cubierto los objetivos terapéuticos sino por desenlaces de lo más variopintos. No obstante, el clínico de trincheras no tiene por que perder su interés por evaluar los resultados obtenidos.

Concluimos esta editorial con una última cita, que en esta ocasión pertenece al físico Stephen W. Hawking: *'Alguien me dijo que cada ecuación que incluyese en mi libro reduciría las ventas a la mitad. He puesto una ecuación, la famosa ecuación de Einstein  $E=mc^2$ . Espero que esto no asuste a la mitad de mis potenciales lectores'*. Los artículos metodológicos pueden ser materia árida que requieran una lectura reposada, o quizá más de una, pero consideramos que pueden ser una guía para la valoración científica y rigurosa de nuestros resultados clínicos. Asimismo, hay que valorar con mayor conocimiento el alcance de los estudios publicados, diferenciando entre las decisiones tomadas con ayuda de la estadística y las conclusiones devenidas de la relevancia clínica. Como Hawking, esperamos que los lectores no se atemoricen y den acogida a esta y próximas iniciativas.

