

Significación estadística vs significación clínica

Gabriel Cavada Ch.^{1,2}

¹Facultad de Medicina, Universidad de los Andes.

²División de Bioestadística, Escuela de Salud Pública, Universidad de Chile.

Statistical significance versus clinical significance

El diseño de un estudio, la interpretación y presentación de informes a menudo ignoran la distinción entre significación clínica y significación estadística.

La significación estadística se refiere a si el valor de una prueba estadística excede o no algún nivel especificado de antemano. Mientras que la significación clínica se refiere a la importancia médica de un hallazgo. Las dos significaciones, a menudo están de acuerdo, pero no siempre. El ejemplo más claro implica un estudio (observacional o experimental) que tiene un gran número de participantes, allí pueden observarse pequeñas asociaciones entre la exposición y la enfermedad o condición (en el caso de los estudios observacionales) o las diferencias entre intervenciones (en el caso de los estudios de intervención) que sean estadísticamente significativas. Si estas asociaciones o diferencias son clínicamente significativas dependen de la gravedad de la condición que está siendo estudiada, de la prevalencia de la condición, y los otros beneficios o riesgos de la intervención.

Un efecto de la intervención mostrado por un ensayo clínico para ser estadísticamente significativo y en la dirección del beneficio, pero de pequeña magnitud, puede ser clínicamente significativo si la intervención es relativamente no tóxica, fácil de administrar y útil para una condición que es de importancia para la salud pública.

Por el contrario, el efecto de la intervención puede ser estadísticamente superior al control, pero puede no ser clínicamente significativo si la intervención tiene efectos adversos inaceptables.

Muchas veces, una combinación de eventos clínicos es utilizada como el resultado primario del estudio. El resultado puede ser estadísticamente significativo. A menos que la combinación de resultados tenga sentido científico o clínico, sin embargo, no será clínicamente relevante o significativa, ya que el resultado combina eventos que presumiblemente reflejen un mecanismo de acción común de una intervención o un factor de riesgo o permita a un clínico para resumir fácilmente el efecto de la intervención o un factor de riesgo.

Por otra parte, la valoración de la evidencia a través de p-values, puede ser muy riesgosa y atentar contra el buen criterio, sobre todo cuando se está en presencia de una escasa evidencia, debido a que la casuística es difícil de conseguir (dolencias raras como la enfermedad de Creutzfeldt-Jakob). Como tradicionalmente se usa como nivel de significación el 5%, ¿qué ocurre si en una serie clínica de tamaño 10 pacientes se observan 7 éxitos de tratamientos y 3 fracasos? la evidencia a favor del éxito tendría un p-value igual a 0,1030, por lo tanto, no hay evidencia estadística para tratar, sin embargo, el buen criterio indicaría que es mejor tratar dado que se han observado más éxitos que fracasos.

Personalmente, me gustan las siguientes frases:

- Algunos usan las estadísticas como el ebrio usa los faroles: para apoyo más que para la iluminación.
- La única certeza que se puede obtener es una probabilidad razonable.
- Las estadísticas no son sustitutas del juicio.