***Ensayos terapéuticos, significación estadística y relevancia clínica***

***Aguilar Valdéz Itzel Priscilla***

La medicina basada en la evidencia tiene su mejor sustento científico en los ensayos clínicos terapéuticos de grandes dimensiones. La publicación en 1983 del artículo que hoy podríamos llamar el manifiesto de los megaensayos (“Why do we need …”, traducido como ¿Por qué necesitamos estudios aleatorizados sencillos de grandes dimensiones? orientado a los problemas de salud más frecuentes significó un cambio de paradigma en la manera de diseñar y llevar adelante ensayos científicos.

Los conceptos estadísticos han ido evolucionando tanto para el diseño como la para la interpretación de los resultados, aunque su difusión y comprensión por parte de los médicos es variable. En los últimos años ha comenzado un debate creciente, producto de un momento histórico diferente.

¿Cuáles son las diferencias entre estos dos períodos y cómo están impactando en la interpretación de los ensayos clínicos? Cuento con la comprensión del lector al afirmar que la descripción que se presenta es esquemática y orientada sólo a las grandes tendencias. En la década de los ’80 no contábamos con tratamientos útiles para el infarto, salvo ante algunas complicaciones, y muy escasos recursos en la insuficiencia cardíaca. Tampoco existían conductas de rutina validadas de prevención secundaria. El surgimiento de la nueva forma de investigar y sus extraordinarios éxitos llevaron a incorporar un arsenal terapéutico valioso y a modificar notablemente la evolución natural en diferentes ámbitos de la patología cardiovascular. Era fácil y ético efectuar estudios contra placebo con diferentes intervenciones y se generó un extraordinario entusiasmo comunitario para participar en las redes de investigación.

¿Un paciente que ya recibe luego de un infarto betabloqueantes, aspirina, inhibidores de la enzima convertidora y estatinas puede beneficiarse con un nuevo fármaco que actúa sobre un mecanismo antiaterosclerótico diferente? ¿Qué régimen antitrombótico-antiplaquetario es más eficaz durante la angioplastia en términos de infarto periprocedieinto y sangrado? ¿Es equivalente-no inferior un tratamiento determinado al previo validado: antagonistas de la angiotensina vs inhibidores, diferentes betabloqueantes, diferentes grupos antihipertensivos, diferentes hipoglucemiantes? Como con la mayoría de los nuevos tratamientos no se podrá reducir la mortalidad respecto de los previos, se evalúan puntos combinados o subrogantes. Así, la mortalidad se descompone en cardiovascular y no cardiovascular para mejorar el poder del estudio, se define infarto con criterios que incrementan su incidencia y que todavía no se han podido uniformar o se obtienen beneficios en algunos puntos pero escaso impacto en otros más relevantes. Con números muy grandes de pacientes se obtienen resultados estadísticamente significativos, pero con una eventual repercusión clínica cuestionable, o sobre eventos de escaso significado clínico, como el late-loss en los estudios con stents coronarios

EL NIVEL DE P Y LA SIGNIFICACIÓN ESTADÍSTICA La introducción de la estadística en el pensamiento médico ha sido dificultosa, en parte por un conflicto esencial e irresoluble: la necesidad implícita en el acto médico de colaborar en la atención de la salud del paciente que nos consulta hoy a partir de pruebas científicas basadas en ensayos que se afirman en conceptos probabilísticos de los grandes números y que nos dicen poco sobre este paciente individual. El primer concepto complejo es comprender el nivel de p y a qué llamamos significación estadística

Hipótesis de nulidad: mortalidad A = mortalidad P; también puede expresarse como mortalidad A – mortalidad de P = 0 Hipótesis alternativa: A<>B o A-P <>0

Error alfa Así se establece el error alfa, en la práctica el nivel de p que se considerará significativo para el ensayo, habitualmente < 0,05 o < 0,01. Si en la comparación estadística resulta un valor de p menor que el error alfa establecido a priori, rechazamos la hipótesis de nulidad y abrazamos la hipótesis alternativa: existen diferencias “significativas” entre los tratamientos. El error alfa expresa la probabilidad de que al rechazar la hipótesis de nulidad podamos equivocarnos, es decir, de incurrir en considerar al estudio falsamente positivo. En otras palabras, si el valor de p fue < 0,05, estamos afirmando que rechazamos la hipótesis de nulidad, aceptamos que hay diferencias significativas entre los tratamientos y que el error en que podemos incurrir al afirmarlo es menor del 5%; hay una probabilidad menor del 5% de que una diferencia de esa magnitud pueda ocurrir por azar.

Valor de p o intervalos de confianza En la década de los ’80 varios autores difundieron las limitaciones de la convencionalidad del nivel de p, su falta de relación con la relevancia clínica y la debilidad de transformar problemas complejos e inciertos en éxitos y fracasos por un parámetro convencional. Propusieron, lo que fue rápidamente aceptado por muchas revistas médicas, informar los resultados en términos de intervalos de confianza e incluso en lo posible omitir el valor de p.

Limitaciones del valor de p y el intervalo de confianza En la década de los ’90, Alvan Feinstein y otros autores introdujeron algunas observaciones críticas sobre la limitación del valor de p y el intervalo de confianza, que han devenido actualmente en una forma diferente de enfocar los resultados de los ensayos. En una de sus publicaciones más representativas afirmaron que tanto el valor de p como el intervalo de confianza referían a la incertidumbre en el rechazo de la hipó- tesis de nulidad, pero eran sólo dos caras de la misma moneda.

La propuesta bayesiana Un aporte ha surgido desde el pensamiento bayesiano proyectado al análisis de los ensayos terapéuticos. La argumentación es que los valores de p y el intervalo de confianza dan una falsa sensación de seguridad frente a la lógica incertidumbre de los ensayos terapéuticos. El intento es transmitir al médico ese margen de incertidumbre a través de nuevas herramientas conceptuales.

La utilización del intervalo de confianza en el sentido bayesiano Hemos comentado que el intervalo de confianza del 95% expresa los valores entre los cuales estimamos que se encuentra el verdadero efecto con un 95% de confiabilidad. Otra forma de interpretar el intervalo de confianza es asumir, sobre la base de los resultados, la probabilidad de diferentes efectos porcentuales, lo que ha sido propuesto por Sackett en los años recientes. Volvamos al ejemplo del GISSI: Hemos observado una reducción de la mortalidad que redondearemos en 17% (9% a 25%).

RELEVANCIA CLÍNICA Y SIGNIFICACIÓN ESTADÍSTICA En febrero de 2010, Sanjay Kaul y George Diamond propusieron una nueva forma de graficación de los resultados, tratando de aproximarse al concepto de relevancia clínica. Queda claro que la importancia o relevancia de una intervención terapéutica está fuertemente influida por el contexto del problema, sus costos, terapéuticas alternativas, riesgos implícitos, de manera que es imposible establecer un porcentaje “ló- gico” o deseable.

**Bibliografía:**

Dr. Carlos D. Tajer. (JULIO-AGOSTO 2010). Ensayos terapéuticos, significación estadística y relevancia clínica. En REVISTA ARGENTINA DE CARDIOLOGÍA(385-390). Argentina: RAC.