La estadística en los ensayos clínicos: conceptos clave

Introducción

Los métodos estadísticos contabilizan formalmente las fuentes de variabilidad en las respuestas de los pacientes al tratamiento. El uso de la estadística permite al investigador clínico llegar a conclusiones razonables y precisas a partir de la información recogida, y sondear decisiones cuando escasean las certezas. La estadística es clave a la hora de prevenir errores y sesgos en la investigación médica. Este artículo abarca algunos conceptos clave de la estadística y sus aplicaciones en los ensayos clínicos.

Contraste de hipótesis

Una hipótesis es una suposición (o un conjunto de suposiciones) que o bien a) afirma algo con una base provisional para orientar la investigación científica; o b) confirma algo como altamente probable a la vista de los hechos establecidos.

Para nuestros fines, nos interesa la hipótesis que afirma algo; por ejemplo, que un nuevo tratamiento para una enfermedad es mejor que el estándar existente de tratamiento. Si el nuevo tratamiento se denomina «B» y el estándar de tratamiento se denomina «A», entonces la hipótesis afirma que «B» es mejor que «A».

Puede que piense que los científicos se dedicarán a demostrar esta hipótesis, pero no es así. La aproximación a este objetivo, en cambio, es indirecta. En lugar de intentar demostrar la hipótesis B, el método científico asume que, de hecho, A es verdad; que no hay diferencia entre el estándar de

tratamiento y el nuevo tratamiento. Es lo que se conoce como la hipótesis «nula». Los científicos intentan entonces demostrar que A es falso. También se conoce como demostrar que la hipótesis nula es falsa. Si pueden hacerlo (demostrar que la hipótesis A es falsa y que el estándar de tratamiento no es mejor que el nuevo tratamiento), se deducirá que B será verdad y que el nuevo tratamiento es mejor que el tratamiento estándar.

¿Por qué se hace así?

No hay una respuesta sencilla a esta pregunta. Se trata del método ampliamente aceptado que ha evolucionado en la ciencia moderna, pero puede ayudar usar una analogía del ámbito jurídico. La hipótesis nula abarca nuestra/o situación/conocimiento actual (usando una analogía de juzgado, que «el acusado es inocente»'), en la/lo que debemos confiar a menos que tengamos pruebas suficientes para no hacerlo. Pero si buscáramos demostrar la «hipótesis alternativa» (como se conoce, lo contrario de la «hipótesis nula»), entonces, en efecto, «el acusado es culpable».

Otra forma (quizá más fácil) de llegar al mismo punto es citar a Albert Einstein:

"No existe cantidad de experimentación que pueda demostrar que tengo razón, pero basta un solo experimento para demostrar que estoy equivocado."

Esto parece sugerir que intentar demostrar que la hipótesis nula es falsa o está equivocada es un objetivo más riguroso y alcanzable que intentar demostrar que la hipótesis alternativa es correcta. Tenga en cuenta que esto NO explica debidamente por qué la ciencia adopta este enfoque; pero quizá pueda ayudarnos a comprender y aceptar más fácilmente un concepto complicado.

Errores de tipo I y de tipo II

Echando un vistazo a la siguiente tabla, podrá ver la diferencia entre errores de tipo I (falsos positivos) y errores de tipo II (falso negativo).

	La hipótesis nula es verdadera	**La hipótesis nula es falsa**
Rechazar la hipótesis nula	Error de tipo I «Falso positivo»	Resultado correcto «Positivo verdadero»
No se puede rechazar la hipótesis nula	Resultado correcto «Negativo verdadero»	Error de tipo II «Falso negativo»

Sigue siendo muy confuso, así que para expresarlo de forma más sencilla este es un ejemplo muy claro:

- Los errores de tipo I podrían matar a un paciente. Imagine un estudio que determinase, incorrectamente, que el estándar de tratamiento no era mejor que el nuevo tratamiento, y que, por lo tanto, conllevara que se dieran nuevos tratamientos a los pacientes (con resultados catastróficos). Si se cometen errores de tipo I, se detectará incorrectamente un efecto que no existe.
- Los errores de tipo II significan que se desperdicia investigación potencialmente valiosa. Puede que dicha investigación hubiera sido realmente útil; pero, dado que no se continúa con el estudio, los pacientes no sufren ningún daño. Si se cometen errores de tipo II, no se detectará un efecto que existe.

Queda claro, por tanto, que los errores de tipo I son más graves que los de tipo II por lo que respecta a los pacientes.

Nivel de significación

El nivel de significación es la probabilidad de cometer un error de tipo I. Este nivel se verá afectado por el tamaño de la muestra y por el «poder estadístico» de la prueba.

Poder estadístico

El «poder» de una prueba estadística es la probabilidad que conducirá, correctamente, al rechazo de una hipótesis nula. O, en otras palabras, la capacidad de una prueba para detectar un efecto (si es que existe). Otra forma de describirlo es decir que el «poder» de una prueba es la probabilidad de NO cometer un error de tipo II.

P-valores

Los p-valores (o valores de «probabilidad») sopesan la fuerza de la evidencia en una escala del 0 al 1. Un p-valor bajo (normalmente menos de 0,05, o 5 %) indica que hay evidencia sólida *contra* la hipótesis nula, lo que podría llevarle a rechazarla; mientras que un p-valor alto (superior a 0,05) indica lo contrario.

Correlación frente a causación

Cuando analizamos los resultados de un ensayo, es importante recordar que no es lo mismo correlación que causación. La correlación es cuando dos variables están relacionadas de alguna forma; sin embargo, eso no significa que una cause la otra (existe una asociación entre ambas variables). Un ejemplo sería el de la terapia hormonal sustitutiva (THS) y la enfermedad coronaria (EC), en el que las mujeres a las que se administraba THS presentaban un menor riesgo de EC. Sin embargo, esto no se debía al proceso de THS en sí, sino al hecho de que el grupo de mujeres que recibían THS solían

pertenecer a un grupo superior en términos socioeconómicos, con dietas y planes de ejercicio mejores que la media.

La causación se observa cuando un factor causa un resultado. Un factor causal es a menudo la causa parcial de un resultado. Para diferenciar entre correlación y causación, es importante registrar tanta información como sea posible sobre los participantes en ensayos. También es necesario aplicar con cuidado la metodología científica en el diseño de los ensayos clínicos y evaluar el posible sesgo en el ensayo.

Manipulación de datos

La manipulación de datos es la práctica de informar incorrectamente de datos (de forma selectiva) o de crear resultados falsos. Un ejemplo sería cuando se descartan intencionadamente datos que no concuerdan con el resultado esperado, para aumentar la proporción de resultados que confirmarán la hipótesis formulada. Cuando un investigador elimina los valores atípicos (un resultado que es muy inferior o superior al siguiente resultado más cercano) de los resultados, es importante verificar que se trate realmente de valores atípicos y no simplemente de resultados que sean diferentes de los resultados esperados o deseados. Otro ejemplo de manipulación de datos sería cuando un recopilador de datos genera aleatoriamente todo un conjunto de datos a partir de la medición realizada a un único paciente.

Transformación de datos

La transformación de datos es la aplicación de una fórmula matemática a algunos datos obtenidos gracias a un ensayo. A menudo se utiliza para que la presentación de los datos sea más clara o fácil de entender. Por ejemplo, si se está midiendo la eficiencia del combustible para los coches, es normal medir la eficiencia en forma de «kilómetros por litro». Sin embargo, si estuviera evaluando cuánto combustible

adicional sería necesario para aumentar la distancia recorrida, se expresaría en «litros por kilómetro». La aplicación de una fórmula incorrecta para obtener la nueva presentación de estos datos en este caso afectaría a los resultados globales del ensayo.

Fusión de datos

La fusión de datos es la acción de combinar los datos de múltiples estudios para poder entender mejor una situación. Uno de sus formatos más habituales es el metaanálisis, en el que se reúnen los resultados de diversos ensayos publicados para agruparlos y compararlos. Al realizar un metaanálisis, es importante comprobar meticulosamente que las metodologías de los ensayos sean las mismas (o, al menos, comparables). Deberá tenerse en cuenta cualquier diferencia en el diseño, de forma que no haya variables diferentes subyacentes (variables que causen confusión). Un ejemplo de fusión incorrecta de datos sería la agrupación de datos de diversos ensayos con diferentes especies de ratón como animal de ensayo.

A2-4.33.3-V1.1