



Revista Andaluza de Medicina del Deporte

Rev Andal Med Deporte. 2009;2(4):109

www.elsevier.es/ramd



Editorial

Evitando errores básicos en la investigación en ciencias del deporte

La tarea del editor no es sencilla. En primer lugar, con respecto a los artículos que se reciben, el editor debe hacer una evaluación inicial para ver si son adecuados para enviarlos a los *referees* para su revisión. Como editor de la *Revista Andaluza de Medicina del Deporte* es mi deber velar por que los artículos recibidos reúnan unas condiciones mínimas antes de enviarlos a nuestros revisores y hacerles perder innecesariamente su valioso tiempo.

Los autores deben entender tal necesidad de llevar a cabo un proceso inicial para prevenir la noticia no deseada de un rechazo directo sin pasar por revisores.

Uno de los puntos clave en los que se fija el editor es si el estudio realizado está bien diseñado y posee una metodología sólida que dé apoyo a los hallazgos del estudio.

Para que esta metodología sea sólida, el estudio debe tener un tamaño muestral apropiado, que permita mostrar diferencias en caso de que éstas existan. Quizá, la cuestión más complicada a este respecto es ¿cuál es este tamaño muestral?, ¿cuántos sujetos necesito para realizar mi estudio? Esta pregunta es sumamente complicada de responder, ya que depende de varios factores tales como el planteamiento/hipótesis que queremos resolver, el resultado primario, el tamaño del efecto (*effect size*) esperado y la población que va a ser estudiada. Así pues, si la población es muy heterogénea, se requerirá una muestra mayor; los síndromes y condiciones raras, por regla general precisan un tamaño muestral más pequeño. Esta última situación también es habitual con los deportistas de élite. Si el tamaño del efecto es grande, se requerirá una muestra menor, y viceversa.

El tamaño muestral debe calcularse durante la etapa de planificación de la investigación y debería basarse en la medida del resultado que se espera obtener. Una de las primeras cuestiones que se plantean los investigadores es cuántos sujetos van a necesitar. El establecer el tamaño muestral requerido ayudará a vislumbrar si el estudio será factible. Se pueden usar datos de la literatura para calcular el tamaño de muestra requerido. Esto debería incluirse en la sección "Métodos" del manuscrito sometido a publicación en la *Revista Andaluza de Medicina del Deporte*. Si no se ofreciese en este apartado, es muy probable que los revisores soliciten esta información durante el proceso de revisión.

Se debe tener especial atención con el poder estadístico (o potencia estadística), que representa la probabilidad de rechazar la hipótesis nula cuando es realmente falsa. Es decir, representa la capacidad de una prueba para detectar como estadísticamente significativas diferencias o asociaciones de una magnitud determinada. En un estudio de investigación, el error de tipo II (o error de tipo beta), se comete cuando el investigador no rechaza la hipótesis nula siendo ésta falsa en la población. Es equivalente a la probabilidad de un resultado falso negativo, ya que el investigador llega a la conclusión de que ha sido incapaz de encontrar una diferencia que existe en la realidad.

Ahora bien, los investigadores no deben confundir la significación estadística (*p*) que es una probabilidad, un valor continuo que va desde 0 hasta 1, con una diferencia que puede ser clínicamente muy significativa o importante.

Quizá todo esto quede más claro si se plantea un ejemplo hipotético y didáctico para un mejor entendimiento. Supongamos que se pretende investigar el efecto de dos tipos de entrenamiento (tratamientos A y B) de 6 meses de duración para la ganancia de fuerza muscular. Para tal propósito, cada estudio se compondrá de dos grupos de sujetos masculinos de características similares (la media de la fuerza antes de los tratamientos era igual en ambos grupos, 1.000 N, con una desviación estándar de 100 N).

El investigador 1 plantea su trabajo utilizando dos grupos de 10 individuos mientras que el investigador 2 lo plantea utilizando 22 individuos.

Tras el periodo de entrenamiento, el investigador 1 observa que ambos tratamientos mejoran la fuerza muscular. El grupo A mejoró su media un 18% y el grupo B un 27%. Así pues, la media tras finalizar el tratamiento fue de 1.180 N para el grupo A y 1.270 N para el grupo B; la desviación estándar conjunta se mantuvo en ± 100 N. En este estudio no se halló diferencia estadísticamente significativa entre los grupos después del entrenamiento.

Por su parte, el investigador 2 obtiene como resultado el mismo incremento porcentual, misma media final y misma desviación estándar; no obstante, observa que sí existe diferencia estadísticamente significativa entre ambos tratamientos.

Uno puede preguntarse por qué ocurre esto, pero la respuesta es sencilla: el investigador 2 había hecho un cálculo inicial teniendo en cuenta que esperaba una diferencia mínima de 90 N entre los tratamientos, con una *p* de 0,05 y una potencia estadística de 0,80; además asumió que podría perder 10% de la muestra durante los 6 meses de entrenamiento.

Además se aconseja un análisis exhaustivo de la literatura disponible (preferentemente revistas arbitradas), con especial énfasis en el número de sujetos utilizados en estudios similares. Se sobrentiende que esto debe hacerse además de analizar las características del entrenamiento utilizado (por ejemplo: duración, volumen, intensidad, densidad) y de los sujetos (estatus de entrenamiento inicial, historial, etc.), buscando las similitudes y diferencias con nuestra propuesta. En el caso de no hallar datos en la literatura científica disponible, se justifica la realización de un estudio piloto.

Ahora bien, ¿es posible que un entrenamiento (o tratamiento) que mejore la fuerza muscular 90 N más que el estándar tenga relevancia clínica? Esta pregunta debería hacerse antes de calcular la muestra necesaria. Además, este hecho debería discutirse considerando los factores ya mencionados, es decir, las características del entrenamiento y de los sujetos estudiados. No obstante, ¿esta ganancia de fuerza no es importante si estamos estudiando deportistas de alto nivel?

Finalmente, si se considera todo lo expuesto antes de empezar cualquier estudio y los hallazgos no resultasen significativos, los autores deberían reconocer la posibilidad de un error tipo II e incorporar este hecho a las limitaciones del estudio.