



RINCÓN DEL INVESTIGADOR

Yo no tiro las faltas como Leo Messi, porque no rechazar la hipótesis nula no es aceptarla



I do not shoot fouls as Leo Messi, because rejecting the null hypothesis is not accepting it

Javier Pascual Huerta

Clínica del Pie Elcano, Bilbao, Vizcaya, España

Recibido el 25 de junio de 2017; aceptado el 17 de julio de 2017
Disponible en Internet el 9 de agosto de 2017

Las pruebas de razonamiento científico, y en especial el test de significación de la hipótesis nula o el test contraste de hipótesis que hemos discutido en números anteriores de esta sección, han aportado herramientas valiosas a los investigadores para medir y cuantificar los resultados de la investigación. Todo comienza al establecer la hipótesis nula inicial del estudio, que básicamente viene a decir que no pasa nada interesante: la infiltración de A no es eficaz en la fasciopatía plantar o la limitación de la flexión dorsal de tobillo no está relacionada con la tendinopatía de Aquiles, por ejemplo. A continuación, y con esta idea en mente, tomamos los datos del grupo de estudio y del grupo control y comparamos los resultados realizando los test estadísticos pertinentes. Cuando el valor p que obtenemos en la prueba estadística utilizada es muy bajo (generalmente menor de 0,05) entonces empezamos a pensar que tenemos algo interesante en nuestras manos y *rechazamos la hipótesis nula*, asegurando con cierto nivel de seguridad, que la infiltración de A sí es eficaz en la fasciopatía plantar o que la limitación de la flexión dorsal de tobillo sí que parece estar relacionada con la tendinopatía de Aquiles. Bien, hasta aquí todo claro. Sin embargo, lo que ocurre en el caso contrario en el que el valor p sea muy alto (generalmente por encima de 0,05) es algo más complejo de interpretar.

El otro día estuve con Leo Messi «echando una pachanga» y nos quedamos al final tirando unas faltas. Pusimos una barrera estática de entrenamiento de 5 jugadores y desde diferentes puntos fuera del área íbamos tirando faltas sin portero a ver quién metía más goles. Los 2 tiramos 8 faltas desde las mismas posiciones con la barrera estática y Messi metió 8 goles (todos dentro) y yo metí 4 y fallé otros 4. El caso es que me quedé pensando si estadísticamente había mucha diferencia entre él y yo tirando faltas y cuando llegué a casa por la noche metí los datos en mi ordenador y calculé el valor p mediante la prueba del test exacto de Fisher y el valor p que salió fue de 0,077. Es decir, no hay diferencias estadísticamente significativas. Así, me pude ir a dormir mucho más contento sabiendo que *no hay diferencias* tirando faltas entre Messi y yo.

Pero, hay algo erróneo en este razonamiento ¿De veras alguien se cree eso? (no me refiero a la historia en sí —que es evidentemente inventada—) ¿De veras alguien se cree que realmente no hay diferencias tirando faltas entre Messi y yo? ¿Podemos concluir que tiramos igual las faltas? Este es un error continuo y repetido en la interpretación de los resultados de una investigación cuando el valor p de un estudio es elevado. La hipótesis nula establece que no existen diferencias entre Leo Messi y yo tirando faltas. No poder rechazar la hipótesis nula con los resultados de una investigación porque el valor p es muy alto pero eso no significa aceptarla. Estos resultados no demuestran que Leo Messi y yo tiramos

Correo electrónico: javier.pascual@hotmail.com

<http://dx.doi.org/10.1016/j.repod.2017.07.001>

0210-1238/© 2017 Consejo General de Colegios Oficiales de Podólogos de España. Publicado por Elsevier España, S.L.U. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

las faltas igual. Estos resultados nos dicen que no podemos rechazar la hipótesis nula con los datos que tenemos, pero no significa que estamos afirmando que tiramos igual las faltas. En otras palabras, no poder demostrar que existe un efecto NO SIGNIFICA demostrar que no existe efecto.

La experiencia nos ha enseñado que los investigadores y los profesionales, especialmente en la investigación biomédica tendemos a confundir este aspecto de forma muy común¹. Es lo que se llama «aceptar la hipótesis nula» y al hacerlo lo más probable es que estemos cometiendo un error (a no ser que también queramos creernos que un podólogo español de la calle tira las faltas igual que Leo Messi). El concepto de no aceptación de la hipótesis nula implica que no hay equilibrio en la toma de decisiones, no hay simetría¹. Cuando el resultado del valor p es muy bajo podemos afirmar con cierto nivel de confianza que existen diferencias y rechazamos así la hipótesis nula. Sin embargo, si el valor p es elevado y los datos no nos permiten rechazar la hipótesis nula, *no aceptamos la hipótesis nula como cierta*, simplemente mantenemos la postura de que podría ser cierta.

Este ejemplo de Messi y yo tirando faltas es un ejemplo adaptado (mejor dicho, copiado) de Vickers² explicando este aspecto de la no aceptación de la hipótesis nula, pero existen en la literatura muchos estudios que ejemplifican este tipo de error o conducta inadecuada del investigador. Un metaanálisis de Tsikopoulos et al. publicado en 2015³ sobre el efecto de las infiltraciones del plasma rico en plaquetas (PRP) en tendinopatías mostró que estas no mostraban reducción del dolor en comparación con placebo o punción seca a 2-3 meses (aunque sí a los 6 meses) y los autores concluyeron que no producían beneficios clínicos. El valor p del estudio con respecto a la reducción del dolor a los 2-3 meses fue de 0,07 (intervalo de confianza al 95%: -0,60 a 0,02). Es decir, cerca del archiconocido valor 0,05, aunque no lo suficiente y con un intervalo de confianza prácticamente en completo mostrando reducción del dolor. A pesar de todo esto, este estudio ha sido utilizado para «demostrar» que las infiltraciones de PRP no aportan ningún beneficio más que el placebo en diversos foros, debates e incluso conferencias magistrales.

Estamos acostumbrados a interpretar los resultados de los trabajos científicos bajo un prisma dicotómico que es el que tendemos que aplicar en nuestros pacientes (Sí/No; este tratamiento funciona/este tratamiento no funciona; es importante medirlo/no es importante medirlo; este factor influye/este factor no influye...). Sin embargo, en la ciencia no es todo blanco o negro y el abanico de grises que se maneja es realmente amplio. Mi opinión personal sobre las infiltraciones de PRP en tendinopatías es que los resultados del estudio de Tsikopoulos et al. son realmente alentadores. Muestran beneficio en la reducción del dolor a los 2-3 meses con las infiltraciones de PRP sobre placebo, aunque no lo suficiente como para alcanzar una significación estadística por debajo de 0,05 (recordemos que la diferencia entre $p=0,05$ y $p=0,07$ es mínima estadísticamente y ridícula desde el punto de vista clínico). Mi predicción es que según pase el tiempo y se acumulen más estudios, más datos, se definan mejor los grupos de estudio y se estandaricen mejor los parámetros de preparación del PRP, muy probablemente se llegará a alcanzar una significación estadística y se demostrará la utilidad clínica real que tiene el PRP en el tratamiento de las tendinopatías. Si eso llega a ocurrir habrá quien entonces diga que esto de los estudios es una locura y que no te puedes fiar porque cada día dicen una cosa...

Bibliografía

1. Prieto Valiente L, Herranz Tejedor I, editores. ¿Qué significa «estadísticamente significativo»? La falacia del criterio del 5% en la investigación científica. Madrid: Ed. Díaz de Santos; 2004.
2. Vickers AJ. Michael Jordan Won't Accept the Null Hypothesis: Notes on Interpreting High P values - Medscape [consultado 15 May 2016]. Disponible en: <http://www.medscape.com/viewarticle/531928>
3. Tsikopoulos K, Tsikopoulos I, Simeonidis E, Papathanasiou E, Haidich AB, Anastasopoulos N, et al. The clinical impact of platelet-rich plasma on tendinopathy compared to placebo or dry needling injections: A meta-analysis. Phys Ther Sports. 2016;17:87-94, <http://dx.doi.org/10.1016/j.pts.2015.06.003>