



CARTAS AL DIRECTOR

Sobre los métodos de regresión en el análisis de la composición corporal de atletas profesionales

On regression methods for the analysis of body composition of professional athletes

Sergio Santana Porbén

Médico, Especialista de Segundo Grado en Bioquímica Clínica. Profesor Asistente. Servicio de Laboratorio Clínico. Hospital Pediátrico "Juan Manuel Márquez". Marianao. La Habana. Cuba

* Autor para correspondencia.
Correo electrónico: ssergito@infomed.sld.cu (Sergio Santana Porbén).

Recibido el 23 de octubre de 2020; aceptado el 7 de noviembre de 2020.

Cómo citar este artículo:

Santana Porbén S. Sobre los métodos de regresión en el análisis de la composición corporal de atletas profesionales. JONNPR. 2021;6(n):nnn-nn. DOI: 10.19230/jonnpr.4057

How to cite this paper:

Santana Porbén S. On regression methods for the analysis of body composition of professional athletes. JONNPR. 2021;6(n):nnn-nn. DOI: 10.19230/jonnpr.4057



This work is licensed under a Creative Commons
Attribution-NonCommercial-ShareAlike 4.0 International License
La revista no cobra tasas por el envío de trabajos,
ni tampoco cuotas por la publicación de sus artículos.

Sr. Director

Se me ha llamado la atención sobre el artículo redactado por Pérez Castillo *et al.* (2020)⁽¹⁾ para explorar la probable relación entre la ocurrencia de lesiones musculoesqueléticas y la composición corporal de futbolistas profesionales⁽¹⁾. Según los autores, la composición corporal del deportista, reconstruida de la suma de los pliegues cutáneos, podría determinar el riesgo de ocurrencia de tales lesiones (definidas éstas en el artículo citado como esguinces de la rodilla y/o el tobillo)⁽¹⁾. En 28 futbolistas se contabilizaron 11 lesiones, válido para un 39,3 % del tamaño de la serie de estudio⁽¹⁾. La suma promedio de 6 pliegues cutáneos fue de 62,2 ± 47,9 mm⁽¹⁾. Los investigadores encontraron que la aparición de las lesiones se correlacionó



estrechamente con la sumatoria de los pliegues cutáneos⁽¹⁾. No solo eso: se demostró que en los atletas no lesionados la sumatoria de los pliegues fue de 32 – 69 mm; mientras que, en los lesionados, la suma de los pliegues fue mayor de 70 mm⁽¹⁾.

Existe un abundante cuerpo literario que ha provisto modelos matemáticos para la reconstrucción de la composición corporal del deportista profesional⁽²⁻⁵⁾. Todos estos modelos persiguen estimar el tamaño de la masa muscular esquelética, por cuanto es el factor biológico que determina el éxito deportivo, y es, por lo tanto, el objetivo de los programas de acondicionamiento físico. Siguiendo la línea de pensamiento esbozada, la efectividad de los programas de acondicionamiento físico resultaría en un mayor número de medallas, trofeos, campeonatos y récords individuales y colectivos. Es inmediato que lo anterior se lograría con una tasa mínima de lesiones musculoesqueléticas para asegurarle así al deportista una larga (y también exitosa) carrera profesional.

Lo contrario también sería cierto. Si a la conclusión del programa de acondicionamiento físico se observan pocos cambios en la masa muscular esquelética (o complementariamente: una mayor masa grasa corporal que la aceptada según las exigencias del deporte en cuestión), entonces cabría anticipar un pobre desempeño físico en las competencias previstas, junto con una mayor incidencia de las lesiones tan temidas, y una vida profesional acortada⁽⁶⁾.

Tal (loable) propósito tropieza no obstante con varias dificultades metodológicas, desde la definición de un protocolo de captación de las variables antropométricas de interés, la elección del mejor modelo de reconstrucción de la composición corporal (como los autores citados más arriba mencionaron en el trabajo reseñado, pero sin brindar evidencias de ello), y la correcta interpretación de los hallazgos. Y es aquí donde quiero detenerme. Los investigadores apelan a un modelo de regresión lineal simple para modelar la relación probable entre las lesiones ocurridas y la suma de los pliegues.

Las maquinarias de regresión lineal¹² basadas en los mínimos cuadrados tratan de encontrar aquel juego de parámetros (léase también los coeficientes del sistema | modelo de predicción) que haga mínima el cuadrado de la diferencia entre el valor observado de una variable dependiente y el esperado⁽⁷⁾. La regresión lineal simple sería una de las posibles modalidades de la regresión de mínimos cuadrados, y ofrecen una primera aproximación al problema a resolver al solo considerar 2 parámetros: un intercepto y una pendiente. Pero debo

¹ El término “lineal” se refiere a que el cambio que ocurre en la variable predictora se traslada directamente (esto es: linealmente) a un cambio de igual magnitud en la variable dependiente. Aceptado el término lineal, existen también modelos no lineales de regresión.

² Alternativamente, se dice que el modelo es lineal si los parámetros ocupan el mismo nivel (orden) que las variables en estudio. Esto es: los parámetros del modelo no se incluyen dentro de exponentes ni tampoco dentro del denominador de una fracción.



advertir que siempre será posible encontrar otro modelo con 3 (y más también) parámetros que ofrezca una suma de los cuadrados de los residuales menor³.

Existe otro “pero” en el caso del uso de un modelo de regresión lineal de mínimos cuadrados como solución del problema de la modelación de la ocurrencia de lesiones según la suma de los pliegues cutáneos. Para que los parámetros no sean sesgados (y que el error de estimación de los mismos se mantenga “controlado”), la variable independiente (también denominada en estas lides como predictora) debe medirse sin error⁴. Por su parte, la variable dependiente muchas veces se hace corresponder con realizaciones dispuestas en una escala numérica real. Ello no es el caso aquí. Las lesiones musculoesqueléticas son realizaciones binarias. Esto es: por propia definición la lesión solo puede existir en dos estados excluyentes y a la vez complementarios: Ausente vs. Presente. Trasladado a una escala binaria: Ausente \equiv 0 vs. Presente \equiv 1.

Por estas razones, tratar de describir la relación entre la ocurrencia de las lesiones musculoesqueléticas y la suma de los pliegues mediante un modelo lineal de regresión de mínimos cuadrados es contraproducente, y solo coloca a los investigadores en una situación en la que tienen un resultado (como el que se refleja en la Gráfica 2) que no ofrece claridad sobre la relación que se desea encontrar.

¿Qué hacer?

Primero, hay que trascender el modelo de regresión lineal de mínimos cuadrados, e introducir el concepto de “riesgo”⁽⁸⁾. Esto es: en condiciones naturales los atletas salen a competir libres de lesiones. Por el propio hecho de la actividad deportiva, el deportista está siempre en riesgo de lesionarse, si bien el programa de acondicionamiento físico busca minimizar tal riesgo.

Introducido el concepto de riesgo, entonces podemos recurrir a una regresión logística como la mejor maquinaria de regresión para describir la relación que se desea⁽⁹⁾. A partir de un conjunto (mejor se diría un vector) de covariables (que en el caso que nos ocupa se limitaría a solo una, pero que pueden ser varias), el juego final de parámetros que se obtenga describe

³ Tampoco se trata de “recargar” el modelo de predicción con un número desproporcionado (tal vez exagerado) de parámetros. Existen reglas para construir un modelo de predicción, pero eso está fuera del alcance de esta carta. Solo diré que una buena práctica es buscar el menor número de parámetros con el que se logre la menor suma de los cuadrados de los residuales.

⁴ En algunas situaciones se prefiere incluso asignarle valores predeterminados a la variable predictora.



cuán propenso es el deportista a sufrir una lesión si muestra un valor determinado de la suma de los 6 pliegues cutáneos.

La probabilidad de que ocurra una lesión estaría dada por la expresión siguiente:

$$Prob(Y = y) = \frac{\exp(X' \beta)}{1 + \exp(X' \beta)}$$

En la expresión anterior: $X'\beta$: producto de la multiplicación de los vectores contentivos de las covariables y los parámetros del modelo.

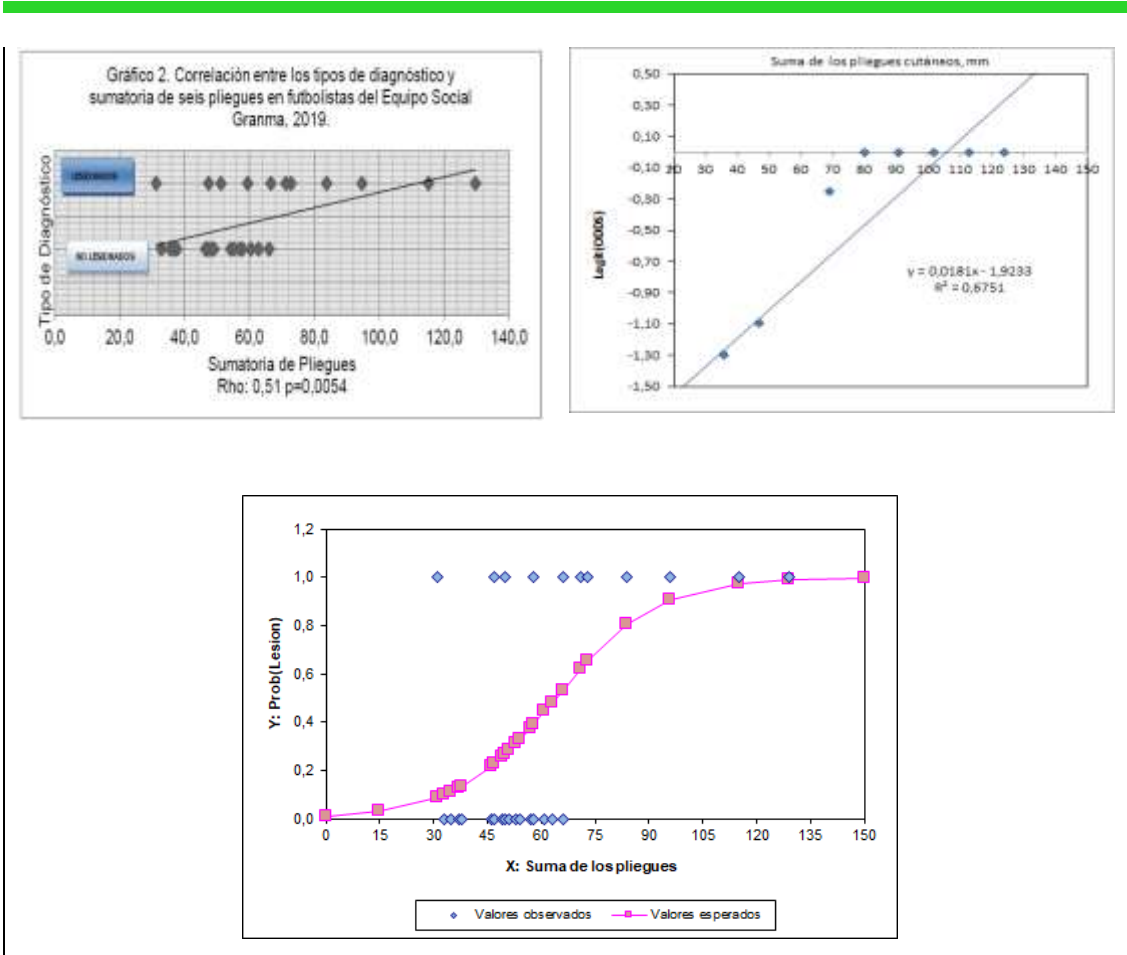


Figura 1. Representaciones gráficas de las asociaciones entre la ocurrencia de lesiones musculoesqueléticas y la suma de los pliegues cutáneos en futbolistas profesionales. Panel superior. *Izquierda:* Lesión presente ($Y = 1$) vs. Sumatoria de los pliegues. *Derecha:* Logit = Log(Odds) vs. Sumatoria de los pliegues. Panel inferior. Probabilidad de ocurrencia de la lesión vs. Suma de los pliegues. Leyenda: ODDS: Posibilidad de ocurrencia de la lesión. La ODDS se estima como la proporción de aquellos futbolistas con lesiones (Y) respecto de los incluidos en un estrato de distribución de la suma de los pliegues cutáneos (N) que están libres de lesiones (Q): $ODDS = Y/(N - Y) = Y/Q$. En condiciones naturales: $ODDS > 0$. En ocasiones se añade 0.5 en ambos miembros de la ODDS para mejorar la estimación de los parámetros logísticos: $ODDS = (Y + 0.5)/(Q + 0.5)$.

De la inspección de la expresión presentada previamente se desprende que la regresión logística cae dentro del dominio de los modelos no lineales, y la solución se aparta ostensiblemente de la propuesta por la regresión lineal. Sin embargo, cualquier paquete de análisis estadístico-matemático contempla como opción la regresión logística. Incluso, la regresión logística puede ser programada dentro de una hoja EXCEL® de cálculo empleando la extensión VisualBasic® para EXCEL®.



La Figura 1 muestra los dos modelos empleados en el análisis de la misma relación. Como se aprecia en la imagen que ocupa el panel inferior, la relación entre la ocurrencia de las lesiones y la suma de los pliegues es en realidad sigmoide, con una probabilidad muy escasa de sufrir una lesión para sumas de pliegues < 60 mm, probabilidad que se incrementa rápidamente a partir de este valor hasta hacerse casi universal para sumas de los pliegues > 90 mm. La regresión logística también nos devolvería cuál sería la razón de disparidades (del inglés *odds-ratio*) implícita en la ocurrencia de la lesión de acuerdo con el valor de la suma de los pliegues como para hacer inferencias pertinentes a un jugador seleccionado. Un atractivo adicional de la regresión logística sería el cálculo de la OR de la lesión si se incluye junto con la suma de los pliegues la posición del futbolista dentro del campo. Así, se podría establecer si las lesiones suelen ser más frecuentes entre los delanteros, por ejemplo, en virtud de la posición de avanzada que desempeñan.

Referencias

1. Pérez Castillo R, Hernandez Cereijo A, Cereijo Yañez D, Pupo Verdecia R. Composición corporal y lesiones musculoesqueléticas: correlación en futbolistas del Equipo Social Granma. JONNPR. 2020;5(11):1311-22. DOI: 10.19230/jonnpr.3423. JONNPR 2020;5. Disponible en: <http://doi:10.19230/jonnpr.3423>. Fecha de última visita: 20 de Octubre del 2020.
2. Pařízková J. Physical activity and body composition. En: Human body composition. Pergamon. London: 1965. pp. 161-176.
3. Santos DA, Dawson JA, Matias CN, Rocha PM, Minderico CS, Allison DB; *et al.* Reference values for body composition and anthropometric measurements in athletes. PLoS One 2014;9 (5):e97846. Disponible en: <http://doi:10.1371/journal.pone.0097846>. Fecha de última visita: 20 de Octubre del 2020.
4. Martin AD, Spent LF, Drinkwater DT, Clarys JP. Anthropometric estimation of muscle mass in men. Med Sci Sports Exerc 1990;22:729-33.
5. Cattrysse E, Zinzen E, Caboor D, Duquet W, Roy PV, Clarys JP. Anthropometric fractionation of body mass: Matiegka revisited. J Sports Sci 2002;20:717-23.
6. Pařízková J. Interrelationships between body size, body composition and function. En: Nutrition and Malnutrition. Springer. Boston [MA]: 1974. pp. 119-149.
7. Montgomery DC, Peck EA, Vining GG. Introduction to linear regression analysis. Volumen 821. John Wiley & Sons. New York: 2012.



-
8. Santana Porbén S. Sobre el riesgo, las razones de disparidad, las estadísticas, y nuestros temores y ansiedades [Editorial]. Disponible en: <https://instituciones.sld.cu/nutricionclinica/editoriales-anteriores/riesgo-razones-disparidad-estadisticas-temores-ansiedades/>. Fecha de última visita: 20 de Octubre del 2020.
 9. Hosmer Jr, DW, Lemeshow S, Sturdivant RX. Applied logistic regression. Volumen 398. John Wiley & Sons. New York: 2013.



Contestación

Raydel Pérez Castillo

Médico. Especialista en Medicina General Integral y Medicina del Deporte. Profesor Instructor. Centro Provincial de Medicina Deportiva, Las Tunas, Cuba Email: raydelperezc@nauta.cu <https://orcid.org/0000-0001-9454-5375>

La revisión de la carta editorial esboza un análisis que maximiza el objetivo del artículo original. Son positivas las consideraciones que sustenta el autor de la Carta al Director. De antemano agradecemos los aportes que sustenta para el artículo original.