

INVESTIGACIÓN ORIGINAL

DOI: <http://dx.doi.org/10.15446/revfacmed.v65n2.59379>

Validación del instrumento para determinar la prevalencia de lesiones osteomusculares en patinadores de carreras en Villavicencio

Validation of the instrument to determine the prevalence of musculoskeletal injuries in racing skaters in Villavicencio

Recibido: 01/08/2016. Aceptado: 04/10/2016.

Sandra González-Vargas¹ • Edgar Cortés-Reyes² • Felipe Marino-Isaza³

¹ Universidad de los Llanos - Facultad de Ciencias Humanas - Programa Licenciatura en Educación Física y Deportes - Villavicencio - Colombia.

² Universidad Nacional de Colombia - Sede Bogotá - Facultad de Medicina - Departamento del Movimiento Corporal Humano - Instituto de Investigaciones Clínicas - Bogotá D.C. - Colombia.

³ Indeportes Antioquia - Medicina Deportiva - Medellín - Colombia.

Correspondencia: Sandra González-Vargas. Departamento del Movimiento Corporal Humano, Facultad de Medicina, Universidad Nacional de Colombia. Carrera 30 No. 45-03, edificio 471, oficina 216. Teléfono: +57 3112103950. Bogotá D.C. Colombia. Correo electrónico: segonzalezv@unal.edu.co.

| Resumen |

Introducción. Se desconocen las condiciones físicas de los patinadores y su relación con las lesiones osteomusculares en una ciudad con alta práctica deportiva como lo es Villavicencio.

Objetivo. Validar un instrumento para determinar la prevalencia de lesiones osteomusculares en patinadores de carreras en Villavicencio.

Materiales y métodos. El instrumento incluye cinco dominios analizados por expertos en las dimensiones cualitativa y cuantitativa, aplicando el coeficiente alfa de Cronbach para análisis de confiabilidad. La validez fue analizada con las propiedades psicométricas de las preguntas, es decir, correlación pregunta-prueba, discriminación de preguntas, nivel de dificultad e índices de no-respuesta.

Resultados. A nivel cualitativo, hubo gran concordancia entre las evaluaciones de expertos con el uso de la W de Kendall (variables ordinales) al 95% de confianza y alta concordancia inter e intraevaluador en la confiabilidad del cuestionario. Por coherencia entre los ítems, a nivel cuantitativo se observó un valor $p > 0.05$. La consistencia interna fue > 0.6 con correlación biserial buena en 50% de las preguntas. 96.9% aprobaron criterios de dificultad, 87.5% tuvieron discriminación entre alta y muy alta y 93.8% presentaron índices de no respuesta adecuada o aceptable.

Conclusiones. La reproductibilidad del instrumento fue alta. Respecto a la validez, la prueba final manifestó propiedades psicométricas óptimas para su uso en el análisis de resultados.

Palabras clave: Estudios de validación; Cuestionarios; Sistema músculoesquelético; Prevalencia (DeCS)..

.....
González-Vargas S, Cortés-Reyes E, Marino-Isaza F. Validación del instrumento para determinar la prevalencia de lesiones osteomusculares en

patinadores de carreras en Villavicencio. Rev. Fac. Med. 2017;65(2):189-96. English. doi: <http://dx.doi.org/10.15446/revfacmed.v65n2.59379>.

| Abstract |

Introduction: The physical condition of skaters and its correlation with musculoskeletal injuries in a city with a high level of sports practice are unknown.

Objective: To validate an instrument to determine the prevalence of musculoskeletal injuries in racing skaters in Villavicencio.

Materials and methods: The instrument includes five domains qualitatively and quantitatively analyzed by experts using the Cronbach's alpha coefficient for reliability analysis. Validity was analyzed based on the psychometric properties of the questions, that is, question-test correlation, question discrimination, difficulty level, and non-response indexes.

Results: In qualitative terms, a high level of agreement between expert evaluations was found using Kendall's W (ordinal variables) at 95% confidence, as well as a high inter- and intra-evaluator agreement in the reliability of the questionnaire. A p value greater than 0.05 was quantitatively observed for consistency between items. Internal consistency was > 0.6 with good biserial correlation in 50% of the questions. 96.9% approved the difficulty criteria, 87.5% were differentiated between high and very high, and 93.8% were not suitable or acceptable responses.

Conclusions: The reproducibility of the instrument was high. Regarding validity, the final test showed optimal psychometric properties that can be used in the analysis of results.

Keywords: Validation Studies; Questionnaires; Musculoskeletal System; Prevalence (MeSH).

González-Vargas S, Cortés-Reyes E, Marino-Isaza F. [Validation of the instrument to determine the prevalence of musculoskeletal injuries skaters racing in Villavicencio]. Rev. Fac. Med. 2017;65(2):189-96 English. doi: <http://dx.doi.org/10.15446/revfacmed.v65n2.59379>.

Introducción

En la actualidad no existe un cuestionario validado que permita conocer las condiciones de la práctica deportiva de los patinadores de carreras, sus procesos de entrenamiento, su relación con la presentación de lesiones osteomusculares, los mecanismos para generar cambios en el entrenamiento y los elementos o estructuras asociadas con la práctica deportiva. Este conocimiento posibilitaría una orientación hacia la atención oportuna de las lesiones y la mejora del bienestar del patinador.

Tampoco se conoce la prevalencia de lesiones, pues no se han desarrollado instrumentos adecuados para medirla. Para este estudio, se consideró a Villavicencio (Colombia) por ser un sitio con alta práctica de patinaje de carreras.

El proceso de construcción y validación de un cuestionario/instrumento es complejo y requiere el conocimiento teórico claro de lo que se quiere medir. Por tanto, ha de reunir las siguientes características: a) ser adecuado para el problema de salud que se pretende medir, teóricamente justificable e intuitivamente razonable, esto es, la llamada validez de contenido; b) ser válido, es decir, ser capaz de medir aquellas características que se propone y no otras; c) ser fiable, preciso y con un mínimo de error en la medida; d) ser sensible al cambio, esto es, capaz de medir variaciones en los individuos y en la respuesta de un mismo individuo a través del tiempo; e) delimitar con claridad sus componentes (dimensiones), de manera que cada uno contribuya al total de la escala de forma independiente (validez de constructo); f) basarse en datos generados por los propios pacientes, y g) ser aceptado por pacientes, profesionales e investigadores (1). Este proceso se trata, por supuesto, de obtener una prueba que contenga el menor número de preguntas y que presente las mejores propiedades con el fin de conformar el puntaje total y sus variables o dominios.

Lo anterior se logra a través del análisis de preguntas con criterios base para la evaluación psicométrica de las mismas, como una etapa de la construcción de la prueba asociada al momento posterior del pilotaje. Se recomienda analizar las preguntas después de la aplicación definitiva, es decir, eliminar aquellas que no satisfacen un conjunto de propiedades que hacen de la prueba un instrumento apropiado e insesgado al estimar la evaluación.

Estas propiedades dependen del modelo adoptado para analizar las pruebas. Los modelos más usados son los de la teoría clásica de los test (TCT) y los modelos de la familia de la teoría de respuestas a ítems (TRI). En el análisis de la prueba, se aplicaron los criterios definidos bajo el enfoque de la TCT como comparabilidad y variabilidad de la prueba, los cuales determinan un conjunto de indicadores referenciales (2).

El objetivo de este estudio fue diseñar y validar un cuestionario que permitiera el registro veraz y confiable sobre el proceso de entrenamiento, los antecedentes de lesiones osteomusculares asociados a la práctica del patinaje de carreras, los escenarios involucrados y el manejo dado a estas lesiones.


Materiales, métodos y resultados

Se llevó a cabo el proceso de validación de un cuestionario/instrumento con el fin de determinar la prevalencia de lesiones osteomusculares en patinadores de carreras de Villavicencio,

este incluyó cinco dominios los cuales contienen las variables de análisis pertinentes (Figura 1). Se analizaron las preguntas desde dos dimensiones complementarias: cualitativa y cuantitativa.

Dimensión cualitativa

El análisis de las preguntas abarcó la revisión en su forma — enunciado y alternativas de respuesta—, las características de sus distractores y la calidad de la clave o respuesta —es decir, si está bien o mal especificada—, mediante revisión de expertos (3). Este análisis se incluyó dentro del proceso final de construcción de la prueba y estuvo a cargo de los especialistas o grupo de expertos responsables del desarrollo de las pruebas desde la etapa de pilotaje. Se pretendió certificar estas propiedades, de manera que la prueba final constituyera un ensamblaje de las mejores preguntas de las versiones piloteadas que, inclusive, recibieron modificaciones.

 UNIVERSIDAD NACIONAL DE COLOMBIA SEDE BOGOTÁ FACULTAD DE MEDICINA CUESTIONARIO PARA DETERMINAR LA PREVALENCIA DE LESIONES OSTEOMUSCULARES EN PATINADORES DE CARRERAS													
El objetivo de este instrumento es recolectar la información necesaria para establecer la frecuencia de presentación de lesiones osteomusculares y relacionar las variables que puedan favorecer su presencia con la práctica del patinaje de carreras													
1. DATOS GENERALES													
NOMBRE: SEXO: F <input type="checkbox"/> M <input type="checkbox"/> FECHA DE NACIMIENTO: (día/mes/año) PESO: ESTATURA: DIARIO DE ENTRENAMIENTO: SI <input type="checkbox"/> NO <input type="checkbox"/> AÑOS DE PRÁCTICA: CATEGORÍA ACTUAL:													
2. DESCRIPCIÓN DEL ENTRENAMIENTO													
¿CUÁNTOS DÍAS POR SEMANA ENTRENA?		1-2		3-5		>6							
¿CUÁNTAS HORAS POR DÍA ENTRENA?		1-2		3-5		>6							
¿REALIZA CALENTAMIENTO ANTES DEL ENTRENAMIENTO?		SÍ				NO							
¿CUÁNTO TIEMPO DURA SU CALENTAMIENTO EN MINUTOS?		5m	10m	20m	30m	40m							
¿REALIZA ESTIRAMIENTO ANTES DEL ENTRENAMIENTO?		SÍ				NO							
¿CUÁNTO TIEMPO DURA SU ESTIRAMIENTO EN MINUTOS?		5m	10m	20m	30m	40m							
¿REALIZA ENTRENAMIENTO DE LA TÉCNICA TODOS LOS DÍAS?		SÍ				NO							
¿CUÁNTO TIEMPO ENTRENA LA TÉCNICA?		5m	10m	20m	30m	40m							
¿REALIZA ENTRENAMIENTO DE FUERZA?		SÍ				NO							
DISTRIBUCIÓN DEL ENTRENAMIENTO DE FUERZA:													
		Días						Duración del entrenamiento					
Distribución	A	L	M	Mi	J	V	S	10m	20m	30m	40m	50m	60m
	B												
	C												

Continúa en la siguiente página

Distribución	D																				
	E																				
A: trabajo de pierna; B: trabajo de brazos; C: trabajo de abdomen; D: trabajo de pantorrilla y pie; E: trabajo de espalda.																					
¿CÓMO REALIZA SU ENTRENAMIENTO?		GRUPAL						INDIVIDUAL													
¿DE QUÉ MANERA COMPITE?		GRUPAL						INDIVIDUAL													
¿CÚAL ES SU ESPECIALIDAD?		VELOCIDAD						FONDO													
¿QUÉ TÉCNICA DE AVANCE EN CURVA UTILIZA?		TRASPIE						OTRA													
¿QUÉ TÉCNICA DE AVANCE EN RECTA UTILIZA?		CLÁSICA			PENDULAR			DOBLE EMPUJE													
¿QUÉ TÉCNICA DE LLEGADA UTILIZA?		ESPAGATA						S		N											
3. LESIONES OSTEOMUSCULARES RELACIONADAS CON LA PRÁCTICA DEL PATINAJE																					
¿CUÁNTAS LESIONES HA TENIDO HASTA EL DÍA DE HOY?																					
0		1-2		3-4		5-6		7-8		9-10		>10									
¿EN QUÉ CATEGORÍA SE ENCONTRABA CUANDO SE LESIONÓ?																					
Categoría		Menores		Transición		Junior		Prejuvenil		Juvenil		Mayores									
No. de lesiones por categoría																					
¿SE LESIONÓ EN ENTRENAMIENTO?		SÍ						NO													
¿EN QUÉ TIPO DE ENTRENAMIENTO?																					
Tipo		PP		PR		EF		EB		EP		OTRO									
No. de lesiones																					
PP: entrenamiento en patines en pista; PR: entrenamiento en patines en ruta; EF: entrenamiento de la fuerza; EB: entrenamiento en bicicleta; EP: entrenamiento en piso.																					
¿SE LESIONÓ EN COMPETENCIA?		SÍ						NO													
¿DURANTE CUÁL(ES) PRUEBA(S)?																					
Prueba		C.R.I		Sprint		Resistencia a la velocidad		Fondo		Equipos		Habilidades									
No. de lesiones por prueba																					
DE LAS LESIONES PRESENTADAS, ¿QUÉ REGION ANATÓMICA SE COMPROMETIÓ?																					
		Miembro superior				Miembro inferior				Tronco											
DE LAS LESIONES PRESENTADAS, ¿QUÉ TEJIDO COMPROMETIÓ?																					
Tejido		Hueso		Cartilago		Tendón		Ligamento		Músculo		Otro									
No. de lesiones																					
¿EN QUÉ PERÍODO OCURRIERON LAS LESIONES?																					
		P. competitivo				P. precompetitivo				Preparación específica											
4. ESCENARIO DE LA LESIÓN																					
Lugar		Pista peraltada				Pista plana				Otro											
No de lesiones																					

Continúa en la siguiente columna

5. ESCENARIO DE LA LESIÓN									
¿CUÁL ES LA FECHA DE SU ÚLTIMA LESIÓN?			Día		Mes		Año		
¿CUÁL FUE LA FECHA DE DIAGNÓSTICO DE ESA ÚLTIMA LESIÓN?			Día		Mes		Año		
¿CUÁNTOS DÍAS DE INCAPACIDAD TUVO?									
¿CUÁNTOS DÍAS ESPERÓ PARA REGRESAR A ENTRENAMIENTO? (sin tener en cuenta los días de incapacidad)									
¿UTILIZÓ ADITAMENTOS (ORTESIS)?					SÍ		NO		
¿RECIBIÓ FISIOTERAPIA?					SÍ		NO		
No. veces		1-2	3-4	5-6	7-8	9-10		>10	
¿REQUIRIÓ MEDICAMENTOS?					SÍ		NO		
¿REQUIRIÓ TRATAMIENTO QUIRÚRGICO?					SÍ		NO		
OBSERVACIONES:									
FIRMA DEL EVALUADOR									

Figura 1. Cuestionario de recolección de información.

Fuente: Elaboración propia.

Dimensión cuantitativa

En esta dimensión, el análisis de las preguntas incluyó el cálculo de los índices correspondientes a las propiedades psicométricas de las preguntas, definidas de acuerdo con el modelo de análisis adoptado. Los criterios presentados se refieren a la TCT: validez (correlación pregunta-prueba), dificultad, discriminación e índice de no respuesta (4). La evaluación de estos indicadores, con sus niveles referenciales, permite visualizar si las preguntas se elaboraron de forma óptima, en cuyo caso el cálculo de los puntajes obtenidos por los evaluados serán válidos, confiables y diferenciables para el uso que se quiere. En el siguiente paso, se presentan los indicadores de la dimensión cuantitativa del análisis a las preguntas.

Validez de la prueba

En general, la validez de una prueba es cuánto esta mide el constructo planteado. El término constructo se refiere a las características que no se pueden medir de manera directa, sino inferidas desde un conjunto de observaciones. El enfoque moderno del concepto de validez es más amplio, pues ya no se habla de diferentes tipos de validez —por ejemplo, de contenido, concurrente o de constructo—. En cambio, se trata de diferentes líneas o formas de evidenciar validez (5).

Existe un conjunto de criterios que proveen información relevante para determinar la validez del cuestionario, de modo que se pueda establecer la prevalencia de lesiones osteomusculares en patinadores de carreras de Villavicencio. Estos son:

1. El juicio de expertos, cuya finalidad es analizar la correspondencia entre el contenido de las pruebas y los constructos que estas intentan medir. Este juicio de expertos se basa en el estudio del contenido

y las tablas de especificaciones que produjeron los especialistas responsables de las pruebas.

2. La confiabilidad de la prueba, la cual mide cuán consistente es esta frente a los puntajes que obtiene. Idealmente, se determina al ejecutar la misma prueba dos o más veces a un examinado y revisar si los puntajes obtenidos son idénticos o similares. En la práctica, la consistencia se observa de formas alternativas, una de las cuales se basa en la consistencia interna de la prueba, por ejemplo, cuán consistente es la mitad de una prueba respecto a su otra mitad.
3. Otras características, basadas en las propiedades psicométricas de las preguntas —como el nivel de dificultad, el grado de discriminación y los índices de no respuesta—, las propiedades derivadas de la construcción de las escalas y las transformaciones hechas para los objetivos de las pruebas.

Validez de las preguntas

Permite demostrar si las preguntas logran lo que la prueba se propone. Así, la correlación pregunta-prueba (punto biserial, cálculo de correlación biserial, índice de homogeneidad) indica el grado de validez de la pregunta. Es un índice de validez interna, ya que la correlación se calcula con el puntaje total de la prueba, a falta del verdadero criterio o constructo. La Tabla 1 presenta las categorías usadas en la evaluación de la validez de las preguntas.

Tabla 1. Clasificación del nivel de dificultad de las preguntas.

Clasificación	Índice de validez
Muy buena	0.2-1
Aceptable	0-0.19
Eliminar	<0

Fuente: Elaboración propia.

Se obtuvo la validez de las preguntas mediante el coeficiente de correlación puntual biserial (6) con la ecuación:

$$r_{bp} = \frac{m_p - m_x}{s_x} \sqrt{\frac{p_i}{q_i}}$$

Donde r_{bp} es el coeficiente de correlación puntual biserial, m_p representa la media de las puntuaciones de los sujetos que responden correctamente al ítem i , m_x indica la media de las puntuaciones de los examinados en el test, s_x es la desviación estándar de las puntuaciones obtenidas por todos los examinados en el test, p_i expresa la dificultad del ítem y $q_i=1-p_i$.

Dificultad de las preguntas

Para calcularla, se considera el porcentaje de aciertos en la pregunta. El indicador varía entre 0 y 1, en el que una pregunta con dificultad de 0.3 es más difícil que aquella con dificultad de 0.8. En el primer caso, solo el 30% acertaron la pregunta y en el segundo caso acertaron el 80%. Esto significa que el indicador se mide en forma inversa a la dificultad de la pregunta.

Aplicado a este cuestionario, el índice de dificultad señala el grado de dificultad de cada pregunta en las áreas seleccionadas. Si una pregunta tiene un índice de dificultad cercano a 0 o 1, esta debería

ser alterada o descartada por no dar cuenta de las diferencias entre los encuestados.

Una pregunta tiene una dificultad media si fue respondida correctamente por cerca del 50% de los examinados, es decir, su índice de dificultad es cercano a 0.5. Índices de dificultad entre 0.3 y 0.7 maximizan la información revelada sobre la diferencia entre los encuestados. En ese sentido, la dificultad media es la proporción de sujetos que escoge la opción correcta y se obtiene con el cociente entre el número de personas que lo aciertan y el número de personas que lo responden. La Tabla 2 presenta los diversos grados de dificultad de las preguntas de la prueba (7).

Tabla 2. Clasificación del nivel de dificultad de las preguntas.

Clasificación	Índice de dificultad
Muy fácil	0.75-1
Fácil	0.55-0.74
Intermedio	0.45-0.54
Difícil	0.25-0.44
Muy difícil	0.00-0.24

Fuente: Elaboración propia.

En adición a la consideración del porcentaje de examinados que eligieron la opción correcta, es de suma importancia analizar el porcentaje de examinados que eligieron cada opción en cada pregunta. Las opciones incorrectas de cada pregunta se llaman distractores.

Discriminación de las preguntas

Se mide por cuánto ayuda la pregunta a ampliar las diferencias estimadas entre los que obtuvieron un puntaje total relativamente alto de la prueba y quienes obtuvieron uno relativamente bajo. El índice de este indicador varía entre -1 y 1. Valores positivos señalan que la pregunta discrimina a favor del grupo superior y valores negativos que la pregunta favorece al grupo inferior.

El índice de discriminación se calcula con la siguiente ecuación:

$$D_i = \frac{GA_{Aciertos} - GB_{Aciertos}}{N_{Grupo\ mayor}}$$

Donde D_i es el índice de discriminación; $GA_{Aciertos}$, número de aciertos en el reactivo i del 25% de los sujetos con las puntuaciones más altas del test; $GB_{Aciertos}$, número de aciertos en el reactivo i del 25% de los sujetos con las puntuaciones más bajas del test, y $N_{Grupo\ mayor}$, número de sujetos en el grupo con mayor número de individuos.

La Tabla 3 muestra las categorías del índice de discriminación del análisis de las preguntas.

Tabla 3. Clasificación de discriminación de las respuestas.

Clasificación	Índice de discriminación
Muy alto	0.4-1
Alto	0.3-0.39
Moderado	0.2-0.29
Muy bajo	0-0.19
Eliminar	<0

Fuente: Elaboración propia.

La discriminación de un ítem es muy alta si su valor se encuentra entre 0.4 y 1 y es muy baja, entre 0 y 0.19. Si los valores son negativos, es recomendable que la pregunta se elimine (8).

No respuesta de las preguntas

El índice de no respuesta de una pregunta se calcula con la proporción de personas que no contestan la pregunta —no señalan ninguna opción como la correcta— o por la proporción de personas que invalidan una opción al señalarla como la correcta. En SPSS, esta situación se programa como missing (casos perdidos). El índice de no respuesta (Nr_i) se define como:

$$Nr_i = 1 - p_i - q_i^*$$

Donde p_i es la dificultad de la pregunta y q_i^* es la proporción de preguntas que no aciertan porque marcaron opciones incorrectas desde la teoría psicológica.

Categorías de no respuestas

La Tabla 4 expone una clasificación de niveles de tasa de no respuesta de las preguntas según los índices correspondientes.

Tabla 4. Clasificación de la no respuesta de las preguntas.

Clasificación	Índice de no respuesta
Adecuada	0-0.15
Aceptable	0.16-0.2
Tolerable	0.21-0.29
Eliminable	0.3-1

Fuente: Elaboración propia.

Este indicador informa si el evaluado ha desarrollado o no contenidos relacionados con la pregunta. Así, si el indicador es alto, es muy probable que el evaluado no conteste la pregunta por desconocimiento del contenido que esta evalúa (9).

A su vez, el tamaño de la prueba puede observarse en la relación entre este indicador con el orden de la pregunta. La posibilidad de que el evaluador no haya tenido suficiente tiempo para terminar la prueba, en especial cuando el índice es más alto en las últimas preguntas, justificaría un tamaño menor de esta. Si las últimas preguntas presentan una tendencia a mayores índices de no respuesta, se puede suponer que el evaluado no las revisó. En este caso, se subestima la dificultad de las preguntas, porque no se sabe si los sujetos que trabajan más despacio acertarían esta pregunta.

Al seleccionar las preguntas de las pruebas definitivas, se priorizaron preguntas con índices de no respuesta dentro de la categoría adecuada. En menor medida, se tuvieron en cuenta preguntas en las categorías tolerable y aceptable. No se consideraron preguntas con índices de no respuesta por encima de 0.30.

Además, cabe resaltar que, para los análisis de las pruebas, la respuesta de no acierto consideró la tasa de no respuesta. Si bien en el análisis de preguntas fue necesario distinguir la no respuesta, en los resultados finales la no respuesta fue tomada como no acierto.

Se reportó la evaluación psicométrica de las preguntas de la prueba construida con el fin de identificar la validez de contenido y constructo del instrumento referido. Se revisaron los criterios con que

se determinó la utilidad psicométrica de las preguntas para calcular un puntaje total de la prueba.

Proceso del análisis de las preguntas

Para esto, se formó un equipo de diez especialistas con experiencia en el ámbito deportivo mayor a cinco años que revisó el *output* (salida) generado por el programa SPSS desde la formación y perspectiva de cada uno. Estos especialistas fueron: un profesional en cultura física deportiva, magister en ciencias del deporte; un nutricionista dietista, magister en fisiología; un médico con formación en administración deportiva; un fisioterapeuta con estudios en deporte, y seis licenciados en educación física y deporte, todos con experiencia y conocimiento en patinaje de carreras.

Ellos revisaron las propiedades óptimas para cada pregunta de las pruebas. Con el SPSS, obtuvieron la información de la dimensión cuantitativa —índices de discriminación, dificultad, no respuesta y validez (correlación biserial)— de las preguntas. El programa permitió clasificar (recodificar variables) los índices en las categorías antes señaladas.

Análisis cualitativo

En esta dimensión se analizó la información sobre distractores competitivos, respuestas correctas mal especificadas y aspectos que suscitaban las perspectivas de los especialistas respecto a la inclusión o eliminación de preguntas para el puntaje final.

Juicio de expertos

El cuestionario fue sometido a juicio de tres expertos, especialistas con amplia experiencia en manejo de atletas —un médico especialista en ortopedia y traumatología, un deportólogo y ortopedista y un licenciado en educación física especializado en entrenamiento deportivo—, quienes dieron su opinión de este a través de un formato de evaluación.

Resultados

Se evidenció un alto nivel de concordancia entre las evaluaciones de los tres expertos en las cinco dimensiones de la prueba al utilizar la *W* de Kendall (para variables ordinales) al 95% de confianza. Lo anterior manifiesta una concordancia inter e intraevaluador alta e importante en la medición de la confiabilidad o reproducibilidad del cuestionario, en especial respecto a la coherencia entre los ítems. Esto, teniendo en cuenta que el coeficiente de *W* Kendall (Tabla 5) se puede encontrar entre 0 y 1 y, por lo general, se espera que se halle entre 0.41 y 0.60 o más; a su vez, el valor $p > 0.05$.

En todas las dimensiones se observó concordancia en las evaluaciones de los especialistas respecto a los dominios del cuestionario, pues todos los valores del *WK* fueron mayores que 0.4 y los valores $p > 0.05$. (Tabla 6).

Análisis cuantitativo

Confiabilidad de la prueba

El criterio de consistencia interna de la prueba fue calculado por el coeficiente alfa de Cronbach (10). En la Tabla 7 se observa consistencia interna de cada factor del cuestionario (> 0.61) después de

reorganizar y cambiar las preguntas que presentaron inconvenientes en el pilotaje. Cabe aclarar que como el tamaño de la muestra fue de 10, el alfa para todo el cuestionario puede tener un error estándar un poco alto ($ee < 0.07$), razón por la cual se presentan los alfa para cada componente del cuestionario.

Tabla 5. W de Kendalle por dimensión del cuestionario.

EV	Jurado	Rango promedio	WK	p-value
Datos generales	J1	2.13	0.5	0.368
	J2	2.13		
	J3	1.75		
Descripción del entrenamiento	J1	1.80	0.6	0.368
	J2	2.10		
	J3	2.10		
Manejo de la lesión	J1	2.10	0.4	0.368
	J2	2.10		
	J3	1.80		
Lesiones osteomusculares presentadas en relación a la práctica de patinaje	J1	2.10	0.6	0.368
	J2	2.10		
	J3	1.80		
Escenario donde ocurre la lesión	J1	2.10	0.5	0.368
	J2	1.80		
	J3	2.10		

J1: jurado 1; J2: jurado 2; J3: jurado 3.

Fuente: Elaboración propia.

Tabla 6. W de Kendalle por dominio del cuestionario.

EV	Jurado	Rango promedio	WK	p-value
Alcance y objetivo	J1	2.13	0.41	0.368
	J2	1.75		
	J3	2.13		
Aplicabilidad	J1	2.33	0.67	0.135
	J2	1.33		
	J3	2.33		
Lesiones osteomusculares presentadas en relación a la práctica de patinaje	J1	2.10	0.4	0.368
	J2	2.10		
	J3	1.80		
Escenario donde ocurre la lesión	J1	2.10	0.6	0.368
	J2	1.80		
	J3	2.10		

J1: Jurado 1; J2: Jurado 2; J3: Jurado 3.

Fuente: Elaboración propia.

Tabla 7. Consistencia interna del test

Factor	Alfa de Cronbach	Número de elementos
Datos generales	0.69	4
Descripción del entrenamiento	0.61	19
Lesiones osteomusculares	0.67	11
Escenario donde ocurre la lesión	0.72	2
Manejo de la lesión	0.69	7
Total	0.79	43

Fuente: Elaboración propia.

Criterios basados en índices psicométricos de las preguntas

Se usaron los promedios de los índices psicométricos de las preguntas: correlación pregunta-prueba, discriminación, nivel de dificultad e índices de no respuesta. En la Tabla 8 (a) se puede observar que la validez de las preguntas medidas por el coeficiente de correlación pregunta-prueba (correlación biserial) fue muy buena en la mitad de estas, mientras que el 23.4% tuvieron un índice biserial aceptable, por lo que no parecen visualizarse ítems para eliminar según este criterio.

Fue ideal no considerar preguntas que fueran muy fáciles o muy difíciles en la prueba piloto, los rangos extremos de dificultad —índice de dificultad mayor de 0.74 o menor de 0.25, respectivamente— fueron presentados en la Tabla 8 (b). Se recomendó revisar dos preguntas: fecha de nacimiento y categoría actual, las cuales parecen no haber sido respondidas gracias a inconvenientes en la recolección de la información. El 96.9% de los ítems aprobaron los criterios de dificultad.

La discriminación de una pregunta es muy alta si su valor está entre 0.4 y 1 y si el índice se halla entre 0.3 y 0.4. Se aprecia en la Tabla 8 (c) que la mayoría de las preguntas cayeron en estas dos categorías. El porcentaje de preguntas que estuvo en el rango de alta y muy alta fue alrededor de 87.5%.

En la Tabla 8 (d) se muestra que el 93.75% de los ítems tuvieron índices de no respuesta entre adecuada y aceptable. Cuando el deportista manifestó no haber tenido lesiones, las preguntas del tipo de lesiones osteomusculares, escenario en donde ocurrió la lesión y manejo de la lesión no se consideraron para este cálculo.

En líneas generales, el índice agregado de discriminación de las preguntas de la prueba fue alto (0.45), tomando el promedio de los coeficientes de discriminación de estas. El nivel de dificultad de la prueba fue fácil (0.64) y se estimó con el promedio de los niveles de dificultad de cada componente de la prueba. El nivel de no respuesta fue adecuado (0.047) y correspondió al promedio de no respuesta de las preguntas.

Discusión

La prevalencia de una enfermedad aumenta como consecuencia de mayor duración de esta, la prolongación de la vida de los pacientes sin que estos se curen, el aumento de casos nuevos, la inmigración de casos (o de susceptibles), la emigración de sanos y la mejoría de las posibilidades diagnósticas (11). Puede colegirse, entonces, que la prevalencia de lesiones osteomusculares en patinadores aumenta con la recidiva de una lesión en quienes han tenido más años de práctica deportiva, la manifestación de nuevas lesiones y nuevos patinadores lesionados, el ingreso de nuevos patinadores provenientes de grupos con esquemas y condiciones de entrenamiento diferentes y la salida de patinadores sin lesiones durante su vida deportiva.

Tabla 8. Clasificación de validez (a), dificultad (b), discriminación (c) y no respuesta de las preguntas (d).

(a)		(b)		
Clasificación	Índice de validez	Clasificación	Frecuencia	Porcentaje
Muy buena	32	Fácil	36	56.3
Buena	17	Intermedio	20	31.3
Aceptable	15	Difícil	3	4.7
Total	64	Muy difícil	2	3.1
		Total	64	100
(c)		(d)		
Clasificación	Índice de discriminación	Clasificación	Índice de no respuesta	
Muy alta	11	Adecuada	56	
Alta	45	Aceptable	4	
Moderada	6	Tolerante	2	
Muy baja	0	Eliminable	2	
Eliminar	2	Total	64	
Total	64			

Fuente: Elaboración propia.

Es fundamental medir y conocer la prevalencia de las lesiones, pues la vigilancia de la salud de atletas no solo proporciona información epidemiológica importante, sino que también permite hacer un seguimiento de los cambios a largo plazo en la frecuencia y las circunstancias de la lesión (12). Estudios como el de Fuller *et al.* (13) brindan recomendaciones con el fin de recolectar información de síntomas asociados al sobreuso en diferentes deportes, mencionan la necesidad de desarrollar instrumentos válidos y sensibles para medir el dolor y otros síntomas y consideran el nivel funcional y no la pérdida de tiempo de los deportes. Palmer-Green *et al.* (14) sugieren que, para el registro de los datos en los cuestionarios, es necesario que médicos o profesionales de la salud relacionados participen, permitan determinar diagnósticos específicos y eviten problemas de fiabilidad y validez vinculados con el libre registro en los cuestionarios de parte de los atletas, lo que fue acogido en esta investigación con la colaboración de profesionales del deporte como jurados del test propuesto.

El cuestionario, validado por el Centro de Investigaciones de Traumatismos Deportivos de OSLO (15) para lesiones por sobreuso se centró en tres áreas comunes de lesión: rodilla, espalda baja y hombro y concluyó con la prevalencia media semanal de todas las lesiones encontradas. En la presente investigación interesó conocer el entrenamiento en términos de duración, frecuencia, intensidad y distribución del entrenamiento de fuerza y resistencia.

A su vez, Clarsen *et al.* (16) determinaron la consistencia interna del cuestionario (α de Cronbach) al evaluar 1 074 atletas noruegos y calcular el promedio semanal de prevalencia de punto y de período para cada área anatómica, lo que generó datos sobre las lesiones por sobreuso y sitios más frecuentes de problemas por uso. Esto les permitió concluir que los estudios de prevención de lesiones en futuros triatletas Ironman deben centrarse en estas áreas. En el presente cuestionario, se incluyó el número de lesiones a la fecha, la categoría del deportista cuando sufrió el daño y si la lesión fue durante el entrenamiento, fuera de él o en competencia.

Cuestionarios como el de Harringe *et al.* (17) midieron la prevalencia de síntomas relacionados con lesión en gimnastas durante

competencia: en su primera parte, este cuestionario utilizó preguntas de caracterización del sujeto —edad, peso corporal, estatura, años de experiencia en el deporte, sesiones de entrenamiento por semana, horas de entrenamiento por semana, tiempo de calentamiento, de estiramiento, de entrenamiento de fuerza, entrenamiento de técnicas propias del deporte y período de entrenamiento— y tuvo una prueba-reprueba con el uso de R de Spearman, la cual mostró una correlación entre 0.92 y 1.0 en todas las preguntas, excepto el entrenamiento de fuerza, donde la correlación fue de 0.72. La segunda parte del contó con preguntas acerca de los síntomas de lesiones en el día de la competición, sin depender de la gravedad, para lo cual utilizaron una lista de localizaciones corporales que marcaran la ubicación de los síntomas. El test-retest de la segunda parte mostró 100% de correlación en todas las preguntas, excepto cuando se produjo la lesión (92%), el tiempo que habían estado practicando la técnica (92%) y la tasa de repetición de la lesión (83%). Los resultados se presentaron sobre la base de un modelo multivariado, a un intervalo de confianza del 95% con significancia estadística del 5% ($p=0.05$). Además, se añadieron las preguntas sobre el tratamiento de la lesión o la falta de este, la participación de algún profesional en este y la necesidad de tratamiento quirúrgico.

Bahr (18) afirma que las lesiones por sobreuso solo se registran en el sistema de vigilancia si ocasionan tiempo perdido, a pesar de que estas pueden limitar el rendimiento. Palmer-Green *et al.* (19) exponen una metodología a fin de registrar lesiones y enfermedades que alteran la disponibilidad del deportista y el rendimiento en varios deportes durante las preparaciones del equipo olímpico en los Juegos de Verano e Invierno en Gran Bretaña. Jacobsson *et al.* (20) muestran un método para registrar alteraciones del entrenamiento normal debido a lesiones en el atletismo y describen otro método de registro epidemiológico con base en el impacto de la lesión o la enfermedad sobre la capacidad para entrenar o competir de un atleta. En este caso no se incluyó, pues los atletas eran menores de edad y no cabían dentro de esta categoría.

Aunque en la epidemiología del deporte los métodos de registro de lesiones y enfermedades solían contemplar en su mayoría la pérdida de tiempo como la medida neta de las consecuencias, en el entorno actual del deporte élite han surgido nuevas discusiones sobre la necesidad de diseñar instrumentos de registro que cuantifiquen el impacto de las lesiones y enfermedades en la función del atleta y en la restricción de su formación, que describan la relación causal de elementos relacionados con características propias del deporte y características particulares de un grupo y que permitan describir también los momentos del entrenamiento y el período de formación del atleta y sus posibles relaciones con la presentación de lesiones. Estos aspectos fueron incluidos en la presente investigación.

Es importante mencionar que una de las limitaciones del estudio fue, dadas las características de este, la dependencia en el grado de veracidad con que los atletas contestaron, por lo que la validez de los datos podría estar amenazada. Por tanto, se confía en que las respuestas se hayan hecho con la suficiente sinceridad.

Conclusiones

A nivel general, el análisis de los ítems basados en la metodología presentada arrojó resultados satisfactorios desde la perspectiva de los indicadores cualitativos y la de los cuantitativos. En cuanto a la validez de la prueba, los resultados obtenidos del análisis de las preguntas a través de una variedad de tipos de evidencias —que van desde aspectos cualitativos, como la opinión de expertos sobre los contenidos de las pruebas, hasta indicadores más cuantitativos en el análisis de las preguntas— sugieren que la prueba final

exhibió propiedades psicométricas óptimas para su uso en el análisis de resultados.

Las características adicionales reflejaron un conjunto de ítems en general con buena discriminación, apesar de haberse centrado en una dificultad baja que no garantizaba las comparaciones por estratos de la muestra. En adición, las tasas de no respuesta resultaron poco significativas.

En el aspecto metodológico, fue útil recoger una serie de sugerencias que servirán para aplicar el instrumento en la población completa o una muestra representativa.

En futuros trabajos, los equipos de revisión de las preguntas observadas deben considerar los indicadores psicométricos y los propios criterios psicológicos con el fin de decidir la conveniencia de la eliminación de una pregunta. Se recomienda implementar una etapa de verificación de cambios entre la etapa piloto y la aplicación definitiva, de manera que se anticipe el comportamiento final de la prueba en términos de validez, confiabilidad e indicadores psicométricos agregados.

Conflicto de intereses

Ninguno declarado por los autores.

Financiación

Ninguna declarada por los autores.

Agradecimientos

Al Instituto de Recreación y Deporte del Meta (IDERMETA) por permitir el acercamiento al área de Medicina Deportiva y a los clubes de patinaje de la Villa Olímpica de Villavicencio, a los entrenadores y profesionales que realizaron el proceso de análisis de las preguntas y a los especialistas que aportaron su opinión experta al cuestionario de recolección de información. Además, a los patinadores y padres de familia de los clubes de patinaje Speed Cats, Activo, Luna Roja y Camaritas.

Referencias

1. **Arribas MC.** Diseño y validación de cuestionarios. *Matronas Profesión*. 2004;5(17):23-9.
2. **Muñiz J.** Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*. 2010;31(1):57-66.
3. **Bazán J.** Evaluación Psicométrica de las Preguntas y Pruebas CRECER 96. Lima: Ministerio de Educación, Vice Ministerio de Gestión Institucional, Unidad de Medición de Calidad Educativa; 2000.
4. **Ferrando-Piera P, Demestre J.** Características de forma y contenido que predicen la capacidad discriminativa en ítems de personalidad: un análisis basado en la Teoría de respuesta a los ítems. *Psicothema*. 2008;20(4):851-6.
5. **Pérez-Gil JA, Chacón S, Moreno R.** Validez de constructo: el uso del análisis factorial exploratorio-confirmatorio para obtener evidencias de validez. *Psicothema*. 2000;12(2):442-6.
6. **Lamas J, Valle-Inclán F, Albo A.** Datos normativos de la escala de susceptibilidad hipnótica de stanford, forma C, en una muestra española. *Psicothema*. 1996;8(2):369-73.
7. **Cárdenas F.** Dificultades de aprendizaje en química: caracterización y búsqueda de alternativas para superarla. *Cienc. educ. (Bauru)*. 2006;12(3):333-46. <http://doi.org/bz5t4d>.
8. **Ebel R, Frisbie A.** Essentials of Education Measurement. 4th ed. Englewood Cliffs: Prentice Hall; 1986. p.35-40.
9. **Tirado F, Backhoff E, Larrazolo N, Rosas M.** Validez predictiva del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Mexicana de Investigación Educativa*. 1997;2(3):67-84.
10. **Cronbach LJ.** Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;(16):297-334. <http://doi.org/cc5>.
11. **Moreno-Altamirano A, López-Moreno S, Corcho-Berdugo A.** Principales medidas en epidemiología. *Salud Pública Méx.* 2000;42(4):337-3.
12. **Junge A, Engebretsen L, Alonso JM, Renström P, Mountjoy M, Aubry M, et al.** Injury surveillance in multi-sport events: the International Olympic Committee approach. *Br J Sports Med*. 2008;42(6):413-21. <http://doi.org/dfhdwt>.
13. **Fuller CW, Bahr R, Dick RW, Meeuwisse WH.** A framework for recording recurrences, reinjuries, and exacerbations in injury surveillance. *Clin J Sport Med*. 2007;17(3):197-200. <http://doi.org/c9q23w>.
14. **Palmer-Green D, Fuller C, Jaques R, Hunter G.** The Injury/Illness Performance Project (IIPP): A Novel Epidemiological Approach for Recording the Consequences of Sports Injuries and Illnesses. *J Sports Med (Hindawi Publ Corp)*. 2013;2013:1-9. <http://doi.org/b8ff>.
15. **Clarsen B, Myklebust G, Bahr R.** Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) Overuse Injury Questionnaire. *Br J Sports Med*. 2013;47(8):495-502. <http://doi.org/f4vz24>.
16. **Clarsen B, Rønsen O, Myklebust G, Flørenes TW, Bahr R.** The Oslo Sports Trauma Research Center Questionnaire on Health Problems: a new approach to prospective monitoring of illness and injury in elite athletes. *Br J Sports Med*. 2014;48(9):754-60. <http://doi.org/f532b5>.
17. **Harringe ML, Lindblad S, Werner S.** Do team gymnasts compete in spite of symptoms from an injury? *Br J Sports Med*. 2004;38(4):398-401. <http://doi.org/dfvmkq>.
18. **Bahr R.** No injuries, but plenty of pain? On the methodology for recording overuse symptoms in sports. *Br J Sports Med*. 2009;43(13):966-72. <http://doi.org/drqjdb>.
19. **Palmer-Green D, Hunter G, Jaques R, Fuller C, Price C.** Methods for epidemiological study of injury and illness in Great Britain Olympic sport: the injury/illness performance project (IIPP). *Br J Sports Med*. 2011;45(4):367. <http://doi.org/fskp47>.
20. **Jacobsson J, Timpka T, Ekberg J, Kowalski J, Nilsson S, Renström P.** Design of a protocol for large-scale epidemiological studies in individual sports: the Swedish Athletics injury study. *Br J Sports Med*. 2010;44(15):1106-11. <http://doi.org/bgwd8qh>.