



Revista Española de Cirugía Ortopédica y Traumatología

www.elsevier.es/rot



ORIGINAL

4 Domain Sports PROM en español: adaptación transcultural en la población chilena y análisis de confiabilidad

D. Figueroa^{a,b,*}, R. Guiloff^{a,b,c}, F. Figueroa^{a,b,c}, E. Stocker^{a,b} y S. Rocha Piedade^d

^a Departamento de Ortopedia y Traumatología, Clínica Alemana, Santiago, Chile

^b Universidad del Desarrollo, Santiago, Chile

^c Hospital Sótero del Río, Santiago, Chile

^d Faculdade de Ciências Médicas, Universidade Estadual de Campinas - UNICAMP, Campinas, Brasil

Recibido el 11 de octubre de 2023; aceptado el 1 de febrero de 2024

PALABRAS CLAVE

Adaptación transcultural;
Validación;
4-Domain Sports Patient-Reported Outcome Measure;
Medidas de resultado reportadas por el paciente;
Deporte;
Ligamento cruzado anterior

Resumen

Antecedentes y objetivo: Las medidas de resultado reportadas por el paciente (PROM) son herramientas de interés creciente en la población deportiva. El propósito de este estudio fue realizar la adaptación transcultural y análisis de confiabilidad del 4-Domain Sports Patient-Reported Outcome Measure (4DSP) al español.

Métodos: Se ejecutó un protocolo de adaptación transcultural en seis etapas para obtener la versión en español del 4DSP (S-4DSP). Posteriormente, se aplicó el cuestionario a una población de 108 deportistas posoperados de reconstrucción de ligamento cruzado anterior (RLCA). Se aplicó nuevamente el cuestionario luego de 30 días. Se evaluó la aceptabilidad, efecto piso y techo, consistencia interna (alfa de Cronbach) y reproducibilidad (correlación intraclase).

Resultados: La S-4DSP fue respondida completamente por 108 participantes (edad media $34 \pm 10,75$, 26% mujeres) alcanzando una aceptabilidad de 100%. No se detectó efecto piso. El análisis estadístico entregó un alfa de Cronbach global para el cuestionario de 0,65, y desagregado por dominios de 0,88, 0,72, 0,27, 0,68 para el primer, segundo, tercer y cuarto dominio, respectivamente. El estudio de correlación intraclase alcanzó un máximo de 0,94 y un mínimo de 0,48 en la primera y quinta preguntas, respectivamente.

Conclusión: El S-4DSP es una herramienta confiable y útil para evaluar deportistas de habla hispana posterior a una RLCA.

© 2024 Publicado por Elsevier España, S.L.U. en nombre de SECOT. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

* Autor para correspondencia.

Correo electrónico: dhfigueroa@gmail.com (D. Figueroa).

<https://doi.org/10.1016/j.recot.2024.02.001>

1888-4415/© 2024 Publicado por Elsevier España, S.L.U. en nombre de SECOT. Este es un artículo Open Access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Cómo citar este artículo: D. Figueroa, R. Guiloff, F. Figueroa et al., 4 Domain Sports PROM en español: adaptación transcultural en la población chilena y análisis de confiabilidad, Revista Española de Cirugía Ortopédica y Traumatología, <https://doi.org/10.1016/j.recot.2024.02.001>

KEYWORDS

Cross-cultural adaptation;
Validation;
4-Domain Sports Patient-Reported Outcome Measure;
Patient-Reported Outcome Measures;
Sport;
Anterior cruciate ligament

4 Domain Sports PROM in Spanish: Cross-cultural adaptation in the Chilean population and reliability analysis

Abstract

Background: Patient-Reported Outcome Measures (PROMs) are tools of increasing interest in the sports population. The purpose of this study was to perform the cross-cultural adaptation and reliability analysis of the 4 Domain Sports Patient-Reported Outcome Measure (4DSP) into Spanish.

Methods: A six-stage cross-cultural adaptation protocol was executed to obtain the Spanish version of the 4DSP (S-4DSP). Subsequently, the questionnaire was administered to a population of 108 postoperative athletes with ACL (Anterior Cruciate Ligament) injuries. The questionnaire was administered again after 30 days. Acceptability, floor and ceiling effects, internal consistency (Cronbach's alpha), and reproducibility (Intraclass Correlation) were evaluated.

Results: The S-4DSP was fully completed by 108 participants (mean age 34 ± 10.75 , 26% women), achieving 100% acceptability. No floor effect was detected. The statistical analysis yielded a global Cronbach's alpha for the questionnaire of 0.65, and domain-specific alphas of 0.88, 0.72, 0.27, and 0.68 for the first, second, third, and fourth domains, respectively. The Intraclass Correlation test reached a maximum of 0.94 and a minimum of 0.48 for the first and fifth questions, respectively.

Conclusions: The S-4DSP is a reliable and useful tool for evaluating Spanish-speaking athletes after ACL reconstruction.

© 2024 Published by Elsevier España, S.L.U. on behalf of SECOT. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introducción

Las medidas de resultado reportadas por el paciente (PROM) son herramientas indispensables en la práctica diaria en ortopedia y traumatología, pues permiten evaluar el impacto de las intervenciones desde la perspectiva de la persona atendida, ayudar a la toma de decisiones clínicas y a la mejora continua de la atención^{1,2}. Además, las PROM han demostrado ser útiles en el desarrollo de investigación médica, proporcionando datos empíricos sobre los efectos de las intervenciones desde la perspectiva del paciente³.

A pesar de la creciente importancia de las PROM en traumatología y ortopedia, existe una brecha considerable en la literatura con respecto a su aplicación específica en la población deportiva⁴. Este grupo, que frecuentemente enfrenta lesiones y procedimientos de rehabilitación, presenta necesidades y expectativas únicas que no se abordan adecuadamente con PROM generales o incluso con los específicos para ortopedia, siendo esto particularmente evidente para deportistas de alto rendimiento⁴⁻⁶.

En este contexto, el *4-Domain Sports Patient-Reported Outcome Measure* (4DSP) ha surgido como una herramienta valiosa para evaluar la calidad de vida relacionada específicamente con el deporte. Este cuestionario considera cuatro dimensiones claves en la evaluación deportiva: estado basal, nivel deportivo, expectativas del paciente y resultados posoperatorios; y permite conocer el nivel deportivo pre lesión, la autopercepción de compromiso debido a la lesión, y la percepción subjetiva del paciente sobre sus resultados posoperatorios⁷.

Aunque esta herramienta fue desarrollada para aplicarse a una variedad de sitios anatómicos y mecanismos de lesión,

una aplicación posible es en sujetos quirúrgicos de rodilla, particularmente aquellos con rotura de ligamento cruzado anterior (LCA), frecuente en la población deportiva y en la que la evaluación de aspectos como la rehabilitación, la percepción de funcionalidad y expectativas de rendimiento deportivo por parte del paciente es crucial para el éxito terapéutico^{8,9}.

Dada la necesidad de contar con los recursos más adecuadas tanto en el procedimiento quirúrgico como para la evaluación y manejo posquirúrgico de los pacientes operados de reconstrucción de ligamento cruzado anterior (RLCA), es esencial que las herramientas de evaluación PROM se encuentren debidamente traducidas y adaptadas a la población en la que se pretenden utilizar¹⁰⁻¹².

El objetivo de esta investigación es adaptar transculturalmente y validar el 4DSP al español, con enfoque en pacientes que han sido sometidos a cirugía de RLCA.

Materiales y métodos

Participantes

Se incluyeron pacientes posoperados de RLCA con al menos un mes cumplido desde el procedimiento, todos mayores de 18 años y que declararon algún grado de participación deportiva previo a la lesión. Aquellos menores de 18 años, con cirugía de revisión de LCA o incapacidad de responder independientemente el cuestionario, fueron excluidos. Todos los participantes firmaron un consentimiento informado antes de su inclusión en el estudio, el cual fue aprobado por el comité de ética institucional.

Tabla 1 Edad de los participantes y proporción según sexo

	N	Frecuencia relativa (%)	Media	DE	Mínimo	Máximo
Sexo						
Femenino	28	26	35	10,73	18	57
Masculino	80	74	33	10,75	18	63
Total	108	100	34	10,72	18	63

DE: desviación estándar.

Adaptación transcultural

La adaptación transcultural se realizó siguiendo el protocolo de seis etapas de Beaton et al.¹⁰ La traducción inicial del cuestionario se llevó a cabo de manera independiente por dos traductores bilingües nativos en español. Posteriormente, ambas versiones fueron sintetizadas en una única consensuada. A continuación, se hizo una contratraducción al inglés por dos traductores independientes, nativos en inglés y sin conocimiento del cuestionario original. Una vez concluido este proceso, un comité de expertos formado por los autores y traductores, revisó todas las versiones y elaboró la pre final del cuestionario, la cual fue discutida con grupos de pacientes posoperados de rodilla de forma independiente, y se recogieron sus comentarios sobre la comprensibilidad y relevancia de los ítems. Las opiniones recopiladas fueron utilizadas para la elaboración de la versión final del cuestionario en español (S-4DSP).

Aplicación del cuestionario

El cuestionario fue aplicado de forma digital a través de *Google Forms* (Google LLC, California, EE. UU.) a 108 participantes a un mes de la intervención quirúrgica. Fue respondido de forma independiente sin supervisión directa del tratante o investigador. Se realizó nuevamente luego de transcurridos 30 días desde la primera aplicación para evaluar la reproducibilidad. Los datos fueron recopilados y procesados anónimamente.

Análisis estadístico

Se realizó un análisis estadístico descriptivo a la muestra para su caracterización demográfica con media, desviación estándar, mínimo y máximo.

Los resultados del cuestionario fueron analizados con el puntaje crudo de cada ítem. Aquellos con respuesta de variable cuantitativa discreta (ítems 1, 3, 5, 6, 8, 9, 10, 11) fueron descritos con media y desviación estándar para la tendencia central y dispersión, respectivamente; estos admiten respuestas desde el 0 al 10.

El análisis de confiabilidad para la evaluación de la consistencia interna del cuestionario se realizó mediante el coeficiente alfa de Cronbach. Este oscila en una escala de 0 a 1, siendo el 1 una correlación perfecta entre las respuestas del cuestionario¹³. Dicha prueba se realizó para las preguntas cuantitativas del cuestionario de forma global, y agrupadas para cada uno de los dominios de evaluación.

El análisis de reproducibilidad se llevó a cabo utilizando el estudio de correlación intraclase en un modelo de

efectos mixtos bidireccional, comparando las respuestas de orden cuantitativo para cada participante entre la primera y segunda aplicación del cuestionario. Este puntaje admite resultados entre 0 y 1, siendo el 1 una correlación perfecta entre la primera y la segunda evaluación (no existe variación en las respuestas)¹⁴.

Las pruebas se realizaron con una confianza de 95%, la significancia estadística fue considerada con un $p < 0,05$. El análisis se llevó a cabo utilizando el *software* estadístico STATA® 2023 (StataCorp LLC, Texas, EE. UU.).

Resultados

Se obtuvo una primera traducción al español consensuada por parte de dos traductores bilingües nativos (*Anexo 1*). Luego una retrotraducción al inglés consensuada por dos traductores independientes (*Anexo 2*). Posterior a la revisión por el grupo de autores, y sumado a evaluaciones independientes de pacientes, se obtuvo el cuestionario traducido final (*Anexo 3*).

Un total de 108 (100%) participantes respondieron el cuestionario de forma completa. No hubo alguno que fuera incapaz de responder. No se registraron entradas con valores faltantes, y ningún sujeto fue excluido en esta etapa. La edad y distribución según el sexo de cada uno se presentan en la *tabla 1*. Del total de estos, 26% fueron mujeres y 74% hombres. La edad media de las participantes femeninas fue de 35 años y la de los varones de 33.

En la evaluación de la consistencia interna del cuestionario, se encontró un alfa de Cronbach global de 0,65, y en el desagregado por dominios se hallaron valores alfa de 0,88, 0,72, 0,27 y 0,68 para el primer, segundo, tercer y cuarto dominio, respectivamente. En la evaluación de reproducibilidad, se registraron las respuestas de 79 participantes (73% de la muestra original). Todos los cuestionarios fueron respondidos completamente y no hubo información faltante. Se encontró un rango de correlación intraclase desde un mínimo de 0,48 para la pregunta cinco, hasta un máximo de 0,94 en la uno. Estos valores junto con los puntajes de los ítems cuantitativos se resumen en la *tabla 2*.

Discusión

El 4DSP fue traducido, adaptado y validado al español siguiendo un proceso estructurado y ampliamente utilizado en la literatura¹⁰. La población objetivo fue seleccionada para ofrecer una aplicación práctica de esta herramienta en una situación clínica frecuentemente encontrada en la traumatología deportiva, como son las lesiones de LCA en deportistas¹⁵.

Tabla 2 Media de respuestas, alfa de Cronbach y correlación intraclase

	Media (DE)	Piso (%)	Techo (%)	Mínimo	Máximo	α de Cronbach	ICC*
<i>1er dominio</i>						0,88	
11	8,7 (1,6)	0	40,1	3	10		0,94
13	9 (1,4)	0	56,5	3	10		0,91
<i>2do dominio</i>						0,72	
15	7,7 (2,2)	0	31,5	1	10		0,48
16	8,3 (2,0)	0	47,2	1	10		0,61
<i>3er dominio</i>						0,27	
18	9,4 (1,1)	0	74	3	10		0,91
19	8,4 (1,9)	0,93	44,4	0	10		0,81
<i>4to dominio</i>						0,68	
110	8,8 (1,9)	0	54,6	1	10		0,74
111	7,7 (1,9)	0,93	20,4	3	10		0,77
<i>Global</i>						0,65	

DE: desviación estándar; ICC: puntaje de correlación intraclase; Piso: porcentaje de respuesta mínima del ítem; Techo: porcentaje de respuesta máximo del ítem.

* Cada valor es el estudio de correlación entre el puntaje de la primera y la segunda evaluación para cada pregunta.

En este estudio se obtuvo un resultado alfa de Cronbach de 0,65 (considerando todas las preguntas). Dicho valor es estadísticamente significativo, se considera aceptable y es similar al reportado por los autores en el artículo de desarrollo del cuestionario original^{7,16}.

Al analizar los resultados por dominio específico se observó una variación considerable en los valores del alfa, que oscilaron entre 0,27 y 0,88. Esto es congruente con la estructura del cuestionario, donde cada uno de los cuatro dominios evalúa aspectos diferentes de la calidad de vida relacionada con el deporte, por lo que cada uno entrega un nivel distinto de congruencia interna según el grado de relación de las preguntas que contiene.

El primer y segundo dominio de evaluación contienen preguntas estrechamente relacionadas. La uno y la tres del primer dominio evalúan la influencia del deporte en la vida del individuo y el grado de motivación que este provoca. Las cinco y seis del segundo dominio indagan sobre el impacto de la lesión en la calidad de vida y el rendimiento deportivo, respectivamente. Es esperable que un individuo con una vida muy influenciada por el deporte tenga un alto grado de motivación en su realización (primer dominio), a su vez con un mayor impacto de la lesión en la calidad de vida, por lo que es lógico que exista al mismo tiempo una mayor repercusión en el rendimiento deportivo (segundo dominio). Con este razonamiento, las respuestas intradominio se encuentran conceptualmente muy relacionadas por lo que deberían mostrar una correlación interna mayor, que es lo que efectivamente sucede y es evidenciado por puntajes alfa de Cronbach más altos (0,88 y 0,72)¹⁷. El tercer y cuarto dominio contienen interrogantes conceptualmente menos asociadas.

En el tercer dominio, la pregunta ocho apunta a si el paciente entendió la información entregada por parte de su tratante, y la nueve evalúa las expectativas que tiene de retornar al nivel de actividad deportiva previo a la lesión. En el cuarto dominio, la interrogante 10 apunta a evaluar qué le parece al sujeto el tratamiento recibido, y la 11 a su estado psicológico al final del tratamiento. Un mismo individuo puede puntuar 0 en la pregunta 8 (no entendió

nada de lo explicado por su tratante), y a la vez puntuar 10 en la nueve (confianza máxima en retornar al nivel deportivo previo). Debido a que las interrogantes están evaluando aspectos independientes del proceso terapéutico, una correlación baja (traducido en un coeficiente alfa más bajo) entre las preguntas para esos dominios específicos es el resultado intuitivo y esperable¹⁸.

Al resumir los hallazgos de alfa de Cronbach, se observa una diferencia entre la puntuación global del cuestionario y los puntajes desagregados por dominio, debido a las diferencias en la orientación y el enfoque de las preguntas dentro de cada dominio. Aquellas con enfoques más diferentes presentan valores más bajos de alfa, mientras que las que son similares dentro de un dominio muestran una correlación interna mayor.

En la evaluación de los resultados del cuestionario no hubo efecto piso, sin embargo, considerando un límite de 15% de respuestas máximas por ítem, se observó un efecto techo en todas las preguntas. Este hallazgo sugiere que el S-4DSP podría no ser lo suficientemente sensible para detectar mejoría en la calidad de vida en pacientes que ya están experimentando un nivel alto de funcionamiento en la esfera deportiva¹⁹.

Con respecto a la reproducibilidad, los valores reportados varían entre el 0,48 y el 0,94, estando cuatro de estos sobre el 0,8, y siete de ellos (de un total posible de ocho) sobre 0,6. Estos valores corresponden a un modelo de efectos mixtos bidireccional (consistente con una evaluación de reproducibilidad intraobservador), y se encuentran dentro del rango considerado como bueno y excelente en una evaluación de reproducibilidad¹⁸. La pregunta que obtuvo una reproducibilidad más baja corresponde a la cinco, «¿Cuánto ha afectado esta lesión en tu calidad de vida?», que puede estar siendo influenciada por la sensación subjetiva en el momento puntual en la que se responde la evaluación, y por lo tanto variar entre aplicaciones (se espera un grado importante de mejoría durante la rehabilitación kinesiológica posoperatoria). Es importante considerar que, en la evaluación de reproducibilidad, siempre existe un sesgo asociado al tiempo de aplicación, en el que se esperan puntajes más dispares a

tiempos más prolongados, esta es una explicación plausible a los resultados reportados¹⁸.

Una adaptación transcultural asume de forma inherente que existen distancias no tan solo idiomáticas sino culturales para la aplicación de un instrumento, por lo que es importante reconocer como limitación el hecho de que este cuestionario haya sido respondido por un grupo reducido de pacientes con patología quirúrgica y centro clínico en común. Por lo tanto, para fortalecer la validez y la generalización de los resultados de la versión adaptada, es necesario realizar estudios adicionales en diferentes grupos de deportistas, en sujetos sometidos a distintos procedimientos y en diversos contextos culturales en nuestro país.

Además, existe una característica importante en el desarrollo de este cuestionario que genera algunas limitaciones metodológicas, que es la incorporación de variables categóricas de respuesta múltiple (en la pregunta cuatro es posible considerar varios gestos deportivos, y en la siete distintos síntomas), si bien está orientado para una mejor tipificación del paciente en un deporte y en una constelación determinada de síntomas, imposibilita la incorporación de estas preguntas dentro del análisis intradominio de confiabilidad con alfa de Cronbach, y de uno de reproducibilidad con herramientas como el estudio de Kappa²⁰.

Por último, es necesario reconocer como limitación que esta investigación no realizó pruebas de validez comparativa con otros instrumentos, y que es necesario considerar en el futuro la realización de una comparación de este cuestionario con herramientas ya validadas en el área.

En resumen, si bien el coeficiente alfa de Cronbach nos proporciona información útil sobre la consistencia interna del cuestionario 4DSP, y la correlación intraclase encontrada nos indica que este es reproducible, es recomendable complementar el análisis realizado con otras pruebas estadísticas y aplicarlo a una población más extensa y diversa de pacientes. Estas medidas adicionales fortalecerán la evidencia de validez del cuestionario y mejorarán su utilidad en la evaluación de la calidad de vida relacionada con el deporte de nuestros pacientes.

Conclusiones

Dichos resultados establecen una base firme para continuar con el desarrollo, evaluación y construcción de validez futura de esta herramienta. Es posible utilizar este cuestionario en el contexto de pacientes posoperados de RLCA.

Nivel de evidencia

Nivel de evidencia IV.

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Anexo. Material adicional

Se puede consultar material adicional a este artículo en su versión electrónica disponible en [doi:10.1016/j.recot.2024.02.001](https://doi.org/10.1016/j.recot.2024.02.001).

Bibliografía

1. Gagnier JJ. Patient reported outcomes in orthopaedics. *J Orthop Res*. 2017;35:2098-108.
2. Docter S, Fathalla Z, Lukacs MJ, Khan MCM, Jennings M, Liu SH, et al. Interpreting Patient-Reported Outcome Measures in Orthopaedic Surgery: A Systematic Review. *J Bone Joint Surg Am*. 2020;103:185-90.
3. Basch E, Barbera L, Kerrigan CL, Velikova G. Implementation of Patient-Reported Outcomes in Routine Medical Care. *Am Soc Clin Oncol Educl Book*. 2018;38:122-34.
4. Rocha Piedade S, Hutchinson MR, Maffulli N. Presently PROMs are not tailored for athletes and high-performance sports practitioners: a systematic review. *JISAKOS*. 2019;4:248-53.
5. Filbay SR, Culvenor AG, Ackerman IN, Russell TG, Crossley KM. Quality of life in anterior cruciate ligament-deficient individuals: a systematic review and meta-analysis. *Br J Sports Med*. 2015;49:1033-41.
6. Snyder Valier AR, Huxel Bliven KC, Lam KC, Valovich McLeod TC. Patient-reported outcome measures as an outcome variable in sports medicine research. *Front Sports Act Living*. 2022;4:1006905.
7. Rocha Piedade S, Hutchinson MR, Ferreira DM, Ferretti M, Maffulli N. Validation and Implementation of 4-domain Patient-reported Outcome Measures (PROMs) Tailored for Orthopedic Sports Medicine. *Int J Sports Med*. 2021;42:853-8.
8. Smith TO, Davies L, Hing CB. Early versus delayed surgery for anterior cruciate ligament reconstruction: a systematic review and meta-analysis. *Knee Surg Sports Traumatol Arthrosc*. 2010;18:304-11.
9. Nwachukwu BU, Adjei J, Rauck RC, Chahla J, Okoroa KR, Verma NN, et al. How Much Do Psychological Factors Affect Lack of Return to Play After Anterior Cruciate Ligament Reconstruction? A Systematic Review. *Orthop J Sports Med*. 2019;7, 2325967119845313.
10. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186-91.
11. Porter I, Gonçalves-Bradley D, Ricci-Cabello I, Gibbons C, Gangannagaripalli J, Fitzpatrick R, et al. Framework and guidance for implementing patient-reported outcomes in clinical practice: evidence, challenges and opportunities. *J Comp Eff Res*. 2016;5:507-19.
12. Acquadro C, Patrick DL, Eremenco S, Martin ML, Kuliš D, Correia H, et al. Emerging good practices for Translatability Assessment (TA) of Patient-Reported Outcome (PRO) measures. *J Patient Rep Outcomes*. 2017;2:8.
13. Tavakol M, Dennick R. Making Sense of Cronbach's Alpha. *Int J Med Educ*. 2011;2:53-5.
14. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15:155-63.
15. Montalvo AM, Schneider DK, Webster KE, Yut L, Galloway MT, Heidt RS Jr, et al. Anterior Cruciate Ligament Injury Risk in Sport: A Systematic Review and Meta-Analysis of Injury Incidence by Sex and Sport Classification. *J Athl Train*. 2019;54:472-82.
16. El-Hajjar ST. Statistical analysis: Internal-consistency reliability and construct validity. *Int J Quant Qual Res*. 2018;6:46-572056.
17. Sainani KL. Reliability Statistics. *PM R*. 2017;9:622-8.
18. Taber KS. The use of cronbach's alpha when developing and reporting research instruments in science education. *Res Sci Educ*. 2018;48:1273-96.
19. Harrison CJ, Plessen CY, Liegl G, Rodrigues JN, Sabah SA, Beard DJ, et al. Overcoming floor and ceiling effects in knee arthroplasty outcome measurement. *Bone Joint Res*. 2023;12:624-35.
20. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22:276-82.