Archives of
Rehabilitation

Summer 2025. Vol 26. Num 2

Research Paper

Investigating the Reliability of the Knee Joint Flexors and Extensors Muscles Length Tests Modified to Meet the Requirements of Scientific Research

Rafał Gnat[1], Anna Gogola[1], *Agnieszka Polaczek[1], Piotr Woźniak[2], Tomasz Wolny[1]

1. Institute of Physiotherapy and Health Sciences, Motion Analysis Laboratory, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland.
2. Dream Motion Medical Group, Libiąż, Poland.

Use your device to scan and read the article online

Citation Gnat R, Gogola A, Polaczek A, Woźniak P, Wolny T. Investigating the Reliability of the Knee Joint Flexors and Extensors Muscles Length Tests Modified to Meet the Requirements of Scientific Research. Archives of Rehabilitation. 2025; 26(2):230-253. https://doi.org/10.32598/RJ.26.2.3986.1

https://doi.org/10.32598/RJ.26.2.3986.1

ABSTRACT

Objective This study evaluates the intra- and inter-rater reliability of modified versions of the widely used clinical tests for assessing knee flexor and knee extensor muscle length. These tests are fundamental components of postural assessments in children and adolescents, helping to identify postural abnormalities and inform appropriate treatment or preventative strategies. The findings can determine the feasibility of using these modified versions in scientific research.

Materials & Methods A total of 76 children (37 girls) with a mean age of 13.24 years were included. The two evaluators took the measurements on the same day, as well as with breaks lasting two and five days. The following indices of reliability were calculated: Intraclass correlation coefficients (ICC), standard errors of measurement, and smallest detectable differences.

Results The outcomes of both muscle length tests demonstrated good to excellent reliability, regardless of whether the measurements were conducted by the same evaluator after 2 or 5 days, or by different evaluators on the same day and after similar intervals. The lowest calculated ICC of 0.79 was for the knee flexors length test on the left side of the body. In all remaining cases, ICCs>0.80 were obtained.

Conclusion After specific modifications, there is a possibility of obtaining a high level of reliability of the two addressed muscle length tests, which justifies the utilization of the modified versions in scientific settings.

Keywords Knee flexors, Knee extensors, Research, Healthy children, Reliability

Received: 20 Nov 2024
Accepted: 14 Apr 2025
Available Online: 01 Jul 2025

Corresponding Author:
Agnieszka Polaczek
Address: Institute of Physiotherapy and Health Sciences, Motion Analysis Laboratory, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland.
Tel: +48 (??) 792892895
E-Mail: a.polaczek11@gmail.com

Copyright © 2025 The Author(s);
This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-By-NC: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

English Version

Introduction

Over the past half-century, physiotherapy has undergone dynamic development in both scientific and clinical domains. Beyond clinical advancements, physiotherapy is gaining recognition in scientific research, as evidenced by the growing number of high-impact professional journals and influential studies worldwide.

Despite this progress, scientifically oriented physiotherapists face certain challenges, particularly regarding diagnostic and measurement tools. One issue is the lack of precise instruments that can be seamlessly integrated into research. This has led to the adoption of technologies from the field of medicine. Physiotherapists have successfully incorporated ultrasound for muscle morphology and activity assessment [1-7], electromyography in various applications [8, 9], functional magnetic resonance imaging for motor tasks [10, 11], motion analysis for spinal mobility, pelvic configuration [12-16], as well as myotonometry [17, 18] and elastography [19, 20] to evaluate tissue biomechanics and therapeutic effects [21].

Nevertheless, an equally significant issue arises when advanced technology proves insufficient. While sophisticated tools provide highly precise data, they may fail to capture broader functional aspects of human movement, an essential focus in physiotherapy. In some cases, simpler, more practical clinical assessments are needed to provide a holistic view. To make these widely used tests suitable for scientific research, they must first be refined to meet rigorous methodological standards. This raises a key question about the reliability of refined assessment techniques.

Assessing the reliability of clinical tests is crucial to ensure their accuracy, consistency, and applicability in both research and practice. Reliable tests allow clinicians to track changes in muscle flexibility and function over time, ensuring that observed differences are due to actual physiological changes rather than measurement errors. In a clinical setting, this enhances decision-making regarding diagnosis, treatment planning, and rehabilitation progress. In research, high reliability is essential for producing reproducible and comparable findings, which strengthens the validity of scientific conclusions and supports evidence-based practice. Without proper reliability testing, even widely used assessments may yield inconsistent results, leading to potential misinterpretations and ineffective interventions.

This research focuses on the adaptation of two widely utilized clinical muscle length assessment techniques, one for the knee flexors (hamstrings) and another for the knee extensors (quadriceps femoris). These muscle groups play a crucial role in postural compensation mechanisms and are often affected by sedentary lifestyles in children and adolescents [22]. Shortened muscles and strength imbalances are frequently observed in inactive youth and young athletes [23-25]. As a result, these tests are integral to postural assessments during development.

Various versions of knee flexor and extensor length tests exist in the literature. The knee flexors test, in particular, has been widely studied, with researchers implementing different measurement techniques, yielding varying degrees of reliability [22, 26-35]. However, there remains a need for simplified, time-efficient, and user-friendly versions that still meet scientific research standards. Additionally, certain reliability aspects, particularly longer intervals between test-retest measurements, have been insufficiently explored in previous studies.

Accordingly, this study assesses the intra- and inter-rater reliability of modified knee flexor and knee extensor length tests, examining their reliability within the same day as well as at 2-day and 5-day intervals in children aged 10-15 years. The findings determine whether modified versions of tests can be effectively implemented in scientific research settings.

Materials and Methods

Study participants

A total of 82 children from three local primary schools volunteered, including 42 boys and 40 girls from classes 4-8 (10-15 years). The participants were generally healthy children presenting various levels of day-to-day physical activity (rather than sitting, recreational activity, and sport activity). The inclusion criteria were as follows: age between 10 and 15 years (to cover the whole puberty spurt period); typical neurologic and motor development (no medical diagnoses of any illnesses or dysfunctions, between 25-75 centiles of the normative body mass and height according to age); and ability to comply with verbal commands. The participants were excluded if they had a history or current diagnosis of any significant orthopedic or neurological disorders (such as

fractures, congenital abnormalities, cerebral palsy, or musculoskeletal pain and or dysfunction lasting more than two weeks). Additional exclusion criteria were having a history of surgical procedures, recent episodes of minor musculoskeletal issues within one month before the study, or experiencing minor health complaints (e.g. cold, headache) on the day of assessment. Two girls were excluded due to a history of surgical interventions within the abdominal area, one girl and three boys, due to a history of fractures within the lower limb.

The chosen age range was selected for several reasons. Firstly, it encompasses the entire puberty growth spurt, a period during which clinicians routinely conduct postural assessments in children, including the two muscle length tests under investigation, and when numerous postural abnormalities are commonly identified. The variability in participants' postural characteristics contributed to diversity and posed a challenge for measurements, while also enhancing the potential for generalizing the results. For this reason, the participants with varying levels of daily physical activity were also included.

Finally, 76 children (37 girls) who qualified for the study (mean age=13.24 years [age range=10-15 years], average body mass=51.27 kg [body mass range=30-72 kg], and average body height=1.534 m [body height range=1.35-1.745 m]) together with their parents received detailed information on the objectives and procedures. The parents granted their informed consent. The measurements took place in one of the local academic centers, in the Motion Analysis Laboratory. During the procedure, no cases of resignation were recorded.

The minimal sample size was estimated using the sample size calculator [36]. The highest number of participants (n=53) was required for calculations concerning the knee extensors length test based on one repeated measurement with an assumed minimal intra-class correlation coefficient (ICC) of 0.70 and an assumed expected ICC of 0.85.

Study design

This was a technical study with repeated measurements of the variables of interest performed by the two assessors on the three measurement days separated by one-day pauses (Figure 1). To capture all considered comparisons, the narrow linear arrows (representing inter-rater reliability) connecting the blocks of rater A and rater B should be mirrored to the opposite side. Likewise, the solid block arrows (depicting intra-rater reliability) linking successive blocks for rater B should also be mirrored accordingly. To avoid redundancy during data analysis, all comparisons indicated by the same type of arrows in Figure 1 were pulled together. The intra-rater reliability 2/5 refers to reliability across measurements taken by this same rater with 2 and 5-day intervals; the inter-rater reliability 0/2/5 denotes the reliability for measurements conducted by the different raters on the same day as well as after 2 and 5 days, respectively.

Raters

To establish a more demanding testing environment, the study involved two raters with relatively limited professional experience. Both were licensed physiotherapists, each with 2.5 to 3.5 years of clinical practice. The raters completed a two-week intensive training program (3 h×3 times per week) to develop sufficient proficiency in performing the two-muscle length test. The training was supervised by an experienced specialist who was not otherwise involved in the study. Upon completion, the specialist evaluated and confirmed that the raters had acquired the necessary skills to competently perform their roles in the research. Raters' skills were also tested in the minor pilot study conducted on 15 adult volunteers, during which the lowest recorded ICC (3,3) equaled 0.91 (inter-rater ICC for knee extensor length test). For blinding reasons, an assistant was also employed whose task was to read and record measurement devices' readouts, which were hidden from the rates themselves.

Knee flexors length test

The subject lay supine on the couch (Figure 2), with the hip and knee of the tested leg flexed to 90 degrees and supported by a small stool held in place by the subject's hands for stability. Two reference points were marked on the anterior lower leg, 5 cm and 15 cm below the patella (measured with the knee extended). Flat neodymium magnets (1.5 cm diameter) were attached with adhesive tape, aligning their proximal edges with these marks to streamline measurement setup and minimize manual error. A Baseline 12-1057 digital inclinometer (Baseline Products, USA) was calibrated vertically and positioned so its proximal edge aligned with the upper magnet. The readout was kept out of the rater's view. A horizontal line was also drawn on the skin between the malleoli. Passive knee extension force was applied using a Steinberg SBS-KW-300A force gauge, placed on the posterior lower leg at the marked line. The rater held the gauge perpendicular to the leg, ensuring consistent force application. The verbal instructions were

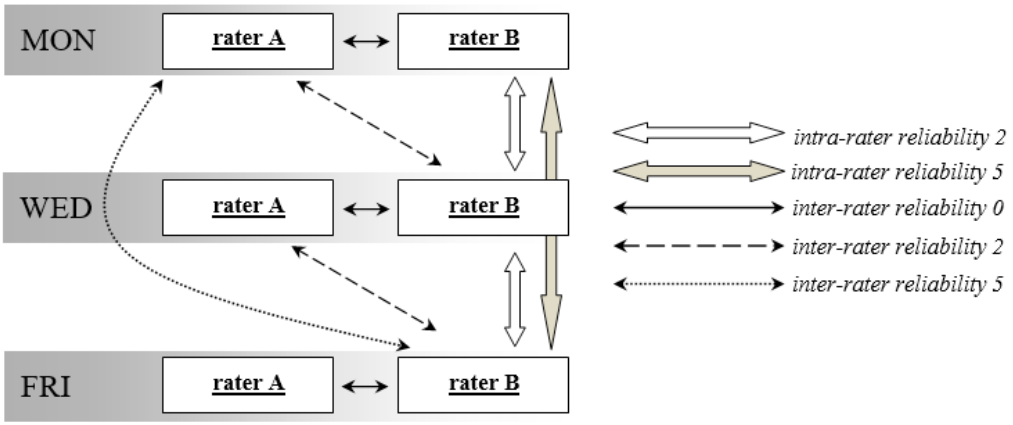


Figure 1. A restructured overview of the measurement procedure and the calculation of ICC

Archives of Rehabilitation

as follows: "Inhale, exhale, let your leg rise smoothly. Say 'Stop' when you feel a strong but tolerable stretch in the back of your thigh." An initial trial determined the force required to reach the "Stop" point. This force was replicated in three test repetitions. Control for the weight of the lower leg was neglected following the findings of Guex et al. [30]. At the following "Now" command, the assistant recorded the inclinometer reading, ensuring blinding of the rater. The procedure was repeated for the opposite leg.

Knee extensors length test

The subject lay prone, positioned diagonally on the couch (Figure 3). The untested leg hung freely off the edge of the table, while the knee (flexed to approximately 90 deg.) was supported on a stool. The couch height was adjusted for comfort. The inclinometer was set up in the same manner as described previously. Similarly, the force gauge was used as before but was applied to the anterior aspect of the lower leg to induce passive knee flexion. The rater gave the following instructions: "Inhale, exhale, and let your leg move smoothly upwards. At the end, say 'Stop' when you feel a strong but tolerable stretch, not pain, in the front of your thigh." As in the flexors test, an initial trial determined the force required to reach the 'Stop' point, ensuring consistency in subsequent repetitions. The actual muscle length test consisted of three repetitions, during which the rater replicated the previously recorded force level. At the following "Now" command, the assistant recorded the inclinometer reading, which was, again, obscured from the rater. After testing one leg, the procedure was repeated on the opposite side.

Study procedure

After the introductory interview and verification of the selection criteria were completed, the convenient week for measurements was chosen by the parents and children. They showed up in the laboratory in the afternoon after at least 3-h post-school rest and at least 2-h after the last meal. The children were asked to wear non-restrictive clothes and perform 10-min low-load warm-up using the stationary bike. Then, subjects received information on how to cooperate during the test performance, and 2-3 attempts of each test were introduced; however, they were fulfilled without exploring the full range of the subjects' motion. The order of the tests for the given participant was randomly established on the first measurement day (Monday) and stayed unchanged on the consecutive measurement days (Wednesday, Friday of the same week). The order of body sides was always right-left. Letters A and B were randomly assigned to the raters. After this, the raters' order always remained A–B. On Monday, after the subjects' proper preparation, the two muscle length tests were repeated by Rater A, three times each. Then, all markings on subjects' skin, attached magnets, etc., were removed, and subjects rested in a relaxed half-seated position in an armchair for 15 min. Subsequently, Rater B prepared subjects again for their measurements, and two tests were repeated once more, three times each. The whole procedure was identically repeated on Wednesday and Friday (Figure 1).

The participants were instructed to maintain their usual level of physical activity throughout the entire measurement week and avoid introducing any sudden changes to their routine activity.

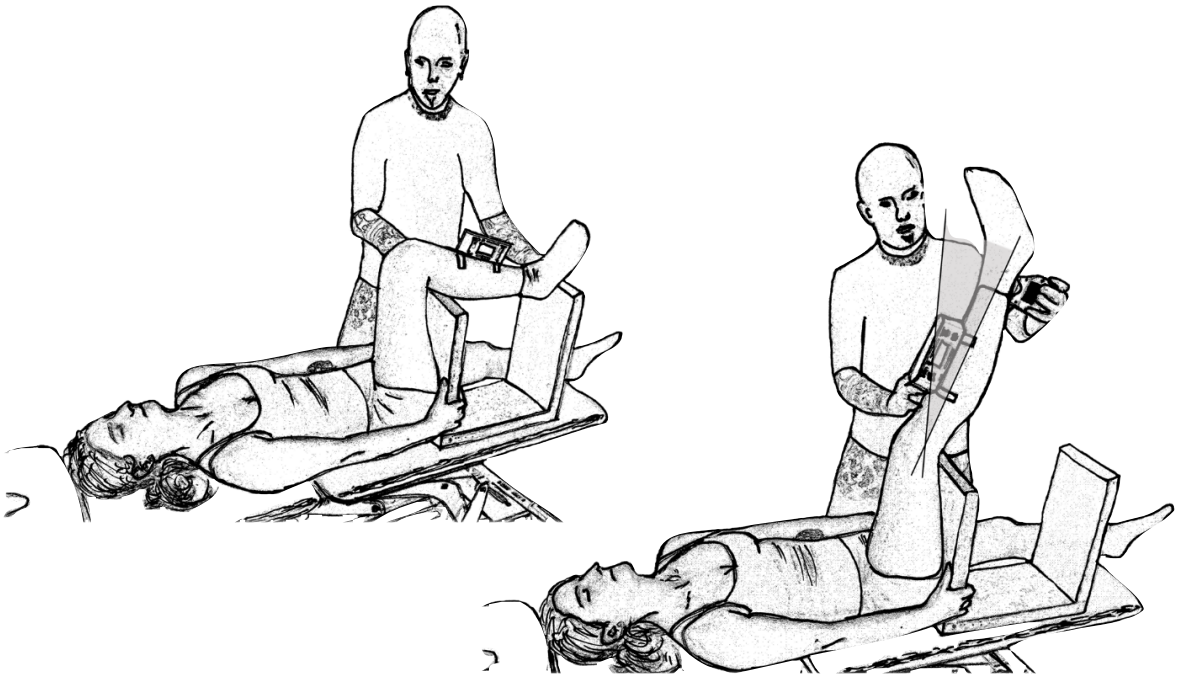


Figure 2. The modified knee flexors length test: Initial position (left) and final position (right)

Archives of Rehabilitation

Note: To preserve clarity, the assistant is not present in the figure. The right-side measurement is presented. Subsequently, the procedure was repeated on the opposite side of the body. On the right, an arc representing the outcome measure is marked in transparent grey.

Data processing

Data recorded by the two raters was gathered in the Statistica 13 (Statistica, Tulsa, USA) spreadsheet. To calculate reliability indices for different types of reliability (Figure 1) and to avoid redundancy of results, this database was specifically restructured. Planned comparisons were pooled together following the following scheme:

Intra-rater reliability, 2-day break between consecutive measurements: rater A on Monday vs rater A on Wednesday+rater A on Wednesday vs rater A on Friday+rater B on Monday vs rater B on Wednesday+rater B on Wednesday vs rater B on Friday (304 records in total); Intra-rater reliability, 5-day break between consecutive measurements: Rater A on Monday vs rater A on Friday+rater B on Monday vs rater B on Friday (152 records in total); Inter-rater reliability, measurements taken on the same day: Rater A on Monday vs rater B on Monday+rater A on Wednesday vs rater B on Wednesday+rater A on Friday vs rater B on Friday (228 records in total); Inter-rater reliability, 2-day break between consecutive measurements: Rater A on Monday vs rater B on Wednesday+rater A on Wednesday vs rater B on Friday+rater B on Monday vs rater A on Wednesday+rater B on Wednesday vs rater A on Friday (304 records in total); Inter-rater reliability, 5-day break between consecutive measurements: Rater A on Monday vs rater B on Friday+rater B on Monday vs rater A on Friday (152 records in total).

Statistical analysis

A factorial analysis of variance model was employed to compare the mean values of the muscle length tests obtained by the two raters across the successive measurement days, with measurement day, rater, and body side treated as independent factors.

For the calculation of ICC, a mixed-model analysis of variance was used, where the repeated factor was the consecutive measurements and the independent factor was the subjects. The reliability of the same-day measurements conducted by a single rater is not reported separately, as it is inherently encompassed within other forms of intra-rater reliability (Figure 1). In Model 2, k of the ICC was applied to allow for generalization into the whole population of similar raters. The ICCs were separately calculated for the results of a single repetition of the given muscle length test, two repetitions (mean value), and three repetitions (mean value). Additionally, standard errors of measurement (SEM) were calculated using the Equation 1, along with the smallest

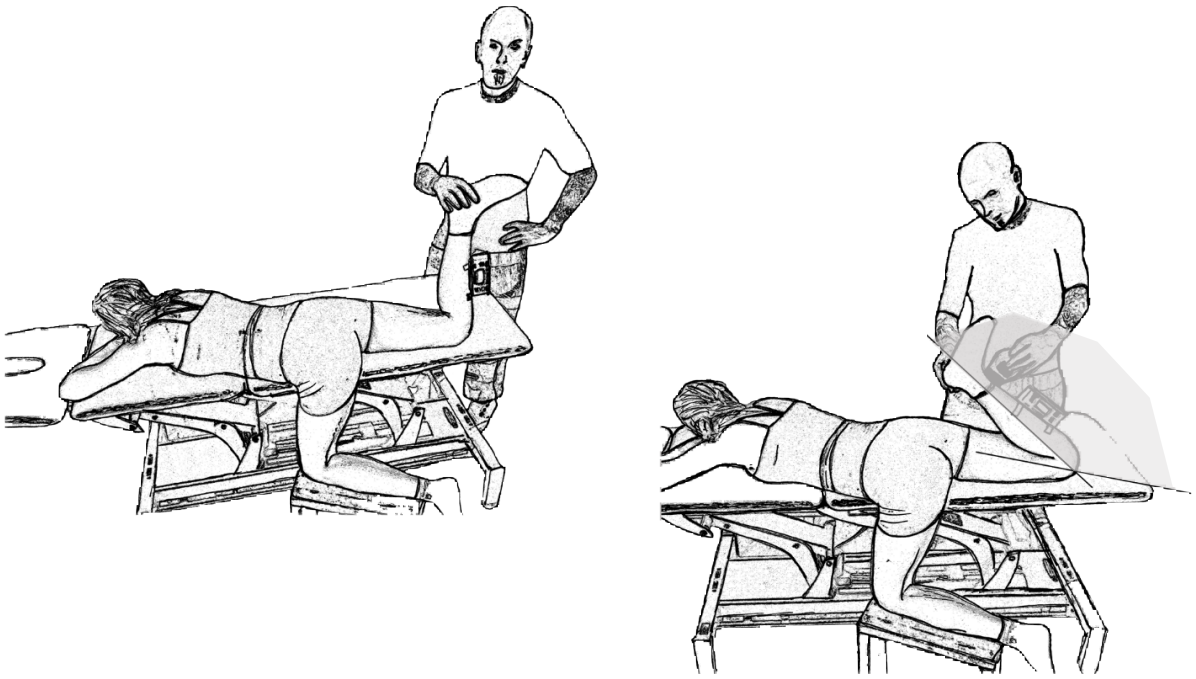


Figure 3. The modified knee extensors length test: Initial position (left) and final position (right)

Archives of Rehabilitation

Note: To preserve clarity, the assistant is not present in the figure. The right-side measurement is presented. Subsequently, the procedure was repeated on the opposite side of the body. On the right, an arc representing the outcome measure is marked in transparent grey.

detectable differences (Equation 2). Interpretation of the reliability of ICCs values was as follows: Poor=0.00-0.50; moderate=0.50-0.75; good=0.75-0.90; and excellent=0.90-1.00 [37].

1. (SEM=SD×(1-ICC)½)

2. (SDD=1.96×2½)

Results

In Table 1, presented are descriptive statistics for the results of the two muscle length tests of interest, as recorded by both raters across consecutive measurement days. For all main effects of analysis of variance (measurement day, rater, side of the body), as well as their interactions, no significant differences were revealed (all $P<0.05$).

The two muscle length tests of interest demonstrated good to excellent reliability, regardless of whether measurements were performed by the same rater with a 2- or 5-day interval (Table 2) or by different raters on the same day as well as with a 2- or 5-day interval (Table 3). The lowest calculated ICC of 0.79 was the one for the knee flexors length test on the left side of the body (SEM=5.47, SDD=15.17 degrees; inter-rater reliability based on one repeated measurement taken with a 5-day break). In all remaining cases, ICCs higher than 0.80 were obtained.

In case of intra-rater reliability (Table 2), even one repeated measurement proved enough to obtain ICCs higher than 0.87. An increasing number of repeated measurements caused a further rise in the ICC values. Across the three repeated trials, ICC values reached or exceeded 0.94. In parallel, a gradual decrease in SEMs and SDDs values was recorded.

For inter-rater reliability (Table 3), lower ICC values were obtained, especially with the longer time break between the measurements. It was frequently the case that one repeated measurement returned ICCs lower than 0.90 (lower than 0.80 in one case mentioned earlier in this section). An increasing number of repeated measurements caused the noticeable rise of the ICC values. With three repeated measurements, all ICCs, even those calculated for measurements taken with the 5-day break, proved higher than 0.90. Again, all SEMs and SDDs gradually decreased.

For each reliability type under investigation, the increasing number of repeated measurements caused the gradual narrowing of the ICCs' confidence intervals, even when the ICC value was no longer increasing markedly.

Table 1. Mean±SD values (minimum-maximum) recorded in the knee flexors and knee extensors length tests based on three repeated assessments performed by each rater across three consecutive measurement days

Rater		Mean±SD (Min-max)			
		Knee Flexors (Degree)		Knee Extensors (Degree)	
		Left	Right	Left	Right
Monday	A	28.18±12.48 (9.00-53.67)	27.83±10.40 (8.00-49.33)	117.80±4.66 (110.00-130.67)	118.07±5.48 (110.00-135.33)
	B	29.05±12.83 (7.00-54.33)	27.12±9.65 (9.33-47.67)	115.97±5.12 (108.33-132.33)	117.35±5.69 (108.00-137.33)
Wednesday	A	27.94±13.53 (6.00-53.33)	28.48±12.48 (9.00-53.67)	117.90±5.45 (110.67-134.33)	118.19±5.45 (110.67-135.33)
	B	27.15±13.89 (5.33-54.67)	29.21±13.05 (7.67-54.67)	116.98±5.78 (108.67-135.00)	119.01±5.81 (108.33-136.67)
Friday	A	28.56±12.58 (8.33-54.67)	28.13±10.81 (9.33-50.67)	118.12±4.95 (109.33-131.33)	117.35±5.78 (109.00-135.00)
	B	28.06±13.26 (5.67-53.67)	28.34±12.33 (8.00-53.33)	117.32±6.05 (109.33-134.67)	118.59±5.11 (109.67-135.33)

Note: Differences between measurement days, raters, and body sides were all non-significant (P<0.05).

Archives of
Rehabilitation

Discussion

The results of our study indicate that the two modified muscle length tests, for the knee flexors and knee extensors tests, demonstrate a very satisfying level of reliability (for 3 repeated measurements, all calculated ICCs proved higher than 0.90). This finding supports the use of these modifications in scientific research settings, providing a reliable alternative to highly sophisticated technological tools. These tools may sometimes fail to capture the broader functional aspects of human movement that clinical tests can address. The findings are particularly beneficial for researchers aiming to implement clinically-oriented measurements in their studies, where simpler, more accessible methods are preferable.

In terms of absolute reliability, our SDD for three repeated measurements was a maximum of 9.93 degrees for the knee flexors and 4.48 degrees for the knee extensors. While the clinical or scientific implications of these values depend on the context, we observed that the differences between raters or on different days were smaller than the SDDs, suggesting that the measurements remained stable and consistent over time. This highlights the robustness of our modified procedures and supports their use for reliable clinical and scientific assessments.

This study was conducted under challenging conditions that might negatively affect reliability. The subjects were in the middle of puberty, a period marked by rapid and unpredictable changes in both morphology and function. Additionally, we incorporated long intervals between tests (up to 5 days) and used two relatively inexperienced raters who underwent a thorough training procedure. Despite these potential challenges, our findings showed excellent reliability, encouraging the use of these modified tests in both clinical practice and research applications.

The knee flexor length test is frequently of interest to researchers worldwide. Our results align with those of previous studies; however, we could address aspects of test reliability that have not been explored before. While many studies focus on intra-rater reliability and measurements taken on the same day, there is a lack of data on inter-rater reliability and longer test-retest intervals. To our knowledge, Hamid et al. [38] are the only authors who have reported inter-rater reliability for this test with a one-week interval. They recorded an ICC (2, 1) of 0.81-0.87, which is comparable to our results (inter-rater at 5 days), although their confidence intervals were much broader, ranging from 0.32 to 0.92. This discrepancy is likely due to their small sample size (n=14) and the lack of repeated measurements (they only performed single measurements). As a result, we can assert that our approach provides more precise estimations of the true ICCs for this measurement, which is particularly important in scientific research settings. Furthermore, our team has previously reported the inter-rater reliability of the knee flexor test [29], though those measurements were conducted by two raters on the same day. Since then, we have refined our methodology by introducing a more "user-friendly" strain gauge and force application technique, as well as a different method for determining the motion endpoint. Additionally, we employed a simple tool to support the subjects' lower legs, which

Table 2. Estimates of intra-rater reliability

Reliability		Measurements	Left			Right		
			ICC (±95% CI)	SEM (Degree)	SDD (Degree)	ICC (±95% CI)	SEM (Degree)	SDD (Degree)
knee flexors	Intra-rater 2	1	0.89 (0.85-0.94)	4.09	11.35	0.87 (0.82-0.93)	4.45	12.33
		2 (mean)	0.94 (0.91-0.97)	3.02	8.38	0.91 (0.88-0.94)	3.70	10.26
		3 (mean)	0.96 (0.94-0.98)	2.47	6.84	0.95 (0.92-0.98)	2.76	7.65
	Intra-rater 5	1	0.87 (0.83-0.91)	4.33	12.01	0.89 (0.85-0.93)	3.99	11.05
		2 (mean)	0.93 (0.91-0.96)	3.18	8.82	0.92 (0.90-0.95)	3.40	9.42
		3 (mean)	0.95 (0.93-0.97)	2.69	7.45	0.95 (0.93-0.96)	2.68	7.42
knee extensors	Intra-rater 2	1	0.89 (0.85-0.94)	1.73	4.80	0.90 (0.87-0.94)	1.79	4.95
		2 (mean)	0.95 (0.92-0.98)	1.09	3.03	0.96 (0.93-0.98)	1.14	3.16
		3 (mean)	0.96 (0.95-0.97)	1.18	3.28	0.96 (0.95-0.98)	1.04	2.88
	Intra-rater 5	1	0.88 (0.84-0.93)	1.80	5.00	0.87 (0.84-0.92)	1.83	5.09
		2 (mean)	0.92 (0.89-0.96)	1.47	4.08	0.92 (0.90-0.95)	1.44	3.99
		3 (mean)	0.95 (0.93-0.97)	1.16	3.23	0.94 (0.92-0.96)	1.24	3.46

Archives of Rehabilitation

Note: The reported data are ICC±95% CI, calculated using models 2,1 for single measurements, 2,2 for two repetitions, and 2,3 for three repeated measurements. Additionally, SEM and the SDD are presented. Intra-rater reliability 2 refers to the consistency of measurements taken by a single rater with a 2-day interval (e.g. Monday–Wednesday, Wednesday–Friday), while intra-rater reliability 5 corresponds to measurements taken with a 5-day interval (e.g. Monday–Friday) (Figure 1).

Abbreviations: ICC: Intraclass correlation coefficient; SEM: Standard error of the mean; SDD: Smallest detectable difference.

appears even more straightforward compared to the devices used in other studies [33, 38]. Moreover, the available literature data are frequently gathered from small samples [26, 27, 29, 30, 38]. Authors use inconvenient ways of motion endpoint determination [33] or implement complicated instrumentation for pelvic motion control [26, 31]. Such measures do not seem to be necessary to obtain a reliable angular measurement in the knee flexors length test.

The available data on the knee extensor length test is somewhat limited. Gajdosik [28] demonstrated excellent reliability using goniometric measurements; however, his study involved a single rater taking measurements on the same day with only 15 subjects. Our team [29] previously provided data on both intra- and inter-rater reliability for this test, but again, all measurements were taken on a single day with a small sample size of 14 subjects. In this study, we addressed these limitations by using a different strain gauge and, again, a refined method for determining the motion endpoint. We also opted not to use external pelvic stabilization, as we found that stabilization provided by the counter-rotation of the two pelvic bones was sufficient (Figure 3). All other data regarding the knee extensor length test presented in this study are novel.

To provide recommendations for the use of the two muscle length tests, we suggest that in scientific settings, it would be beneficial to perform three repeated measurements of the test, particularly when there is a longer time separation between tests and retests (up to 5 days in this study) or when the tests and retests are conducted by different raters. In situations where the time separation is shorter and or a single rater is involved, two repeated measurements may be sufficient. We do not recommend reducing the number of repeated measurements to fewer than two. While this approach might be acceptable in a clinical setting (assuming the clinician has received proper training and adheres strictly to the procedure),

Table 3. Estimates of inter-rater reliability

Reliability		Measurements	Left			Right		
			ICC (±95% CI)	SEM (Degree)	SDD (Degree)	ICC (±95 %CI)	SEM (Degree)	SDD (Degree)
knee flexors	Inter-rater 0	1	0.88 (0.84-0.93)	4.14	11.48	0.89 (0.83-0.95)	4.10	11.35
		2 (mean)	0.93 (0.90-0.96)	3.16	8.76	0.92 (0.89-0.95)	3.49	9.68
		3 (mean)	0.96 (0.93-0.99)	2.39	6.62	0.96 (0.92-0.99)	2.47	6.85
	Inter-rater 2	1	0.85 (0.79-0.90)	4.69	12.98	0.88 (0.83-0.93)	4.17	11.54
		2 (mean)	0.91 (0.87-0.96)	3.63	10.05	0.93 (0.89-0.96)	3.18	8.81
		3 (mean)	0.94 (0.91-0.96)	2.96	8.21	0.95 (0.92-0.98)	2.69	7.45
	Inter-rater 5	1	0.79 (0.74-0.84)	5.47	15.17	0.82 (0.77-0.86)	5.04	13.97
		2 (mean)	0.88 (0.84-0.92)	4.14	11.47	0.89 (0.86-0.92)	3.94	10.92
		3 (mean)	0.91 (0.88-0.93)	3.58	9.93	0.93 (0.91-0.95)	3.15	8.71
knee extensors	Inter-rater 0	1	0.87 (0.83-0.91)	1.96	5.44	0.90 (0.86-0.94)	1.69	4.69
		2 (mean)	0.92 (0.89-0.95)	1.54	4.27	0.93 (0.90-0.95)	1.42	3.93
		3 (mean)	0.95 (0.93-0.98)	1.22	3.38	0.94 (0.92-0.96)	1.31	3.64
	Inter-rater 2	1	0.86 (0.82-0.90)	1.96	5.42	0.85 (0.81-0.89)	2.04	5.66
		2 (mean)	0.91 (0.88-0.95)	1.59	4.42	0.91 (0.89-0.94)	1.58	4.39
		3 (mean)	0.93 (0.91-0.95)	1.38	3.83	0.94 (0.91-0.96)	1.29	3.58
	Inter-rater 5	1	0.81 (0.77-0.85)	2.23	6.17	0.84 (0.81-0.88)	2.06	5.72
		2 (mean)	0.85 (0.82-0.88)	1.98	5.48	0.89 (0.86-0.91)	1.71	4.74
		3 (mean)	0.90 (0.88-0.93)	1.62	4.48	0.92 (0.90-0.95)	1.46	4.04

Archives of
Rehabilitation

Note: The following values are presented: ICC±95% confidence intervals, based on model 2,1 for single measurements, 2,2 for two repeated measurements, and 2,3 for three repeated measurements; SEM; and SDD. Inter-rater reliability 0 refers to measurements taken by different raters on the same day, while inter-rater reliability 2 represents measurements performed with a 2-day interval (e.g. Monday–Wednesday, Wednesday–Friday), and inter-rater reliability 5 pertains to a 5-day interval (e.g. Monday–Friday) (Figure 1).

Abbreviations: ICC: Intraclass correlation coefficient; SEM: Standard error of the mean; SDD: Smallest detectable difference.

where reliability requirements are slightly more flexible, we note that for a single repeated measurement, the lowest ICC we recorded was 0.79 (for the knee flexors length test, inter-rater reliability with a 5-day interval), which is still a satisfactory result for clinical use. Meanwhile, beyond a certain point, increasing the number of repeated measurements does not result in a further increase in the ICC value but does help narrow the confidence interval of the ICC. This phenomenon is advantageous from a scientific perspective, as it enhances the precision of the reliability estimates.

Conclusion

After appropriate specific modifications, there is a possibility of obtaining a very good and excellent level of reliability of the knee flexors and knee extensors length tests. Using the three repeated measurements, all recorded ICCs proved higher than 0.90. Such a high reli-

ability level justifies the utilization of the modified versions of the tests in scientific settings. No sophisticated equipment or high time, personal, or economic costs are needed to achieve this goal.

Study limitations

One limitation of our study is its restricted external validity, as the sample only included teenagers. The knee flexors and knee extensors tests are often used in postural assessments during this developmental stage, but caution is needed when attempting to generalize these findings to other age groups. Nonetheless, the modifications we employed offer advantages, including low cost, minimal equipment (digital goniometer and force gauge), and reduced personnel requirements, which make them accessible for use in various settings. Furthermore, our results suggest that the modified tests can be implemented without the need for complex or expensive tools like X-rays or motion analysis systems. Finally, we chose to present ICCs for model 2, k, which allows for broader generalization across different raters, though model 3, k would likely show higher reliability. We did not provide intra-rater reliability for same-day measurements, as it is inherently captured in the higher-order reliabilities, where we observe higher ICC values due to reduced variance.

The promising results of this study open up several avenues for future research on the reliability and application of modified clinical muscle length tests, particularly for use in both clinical practice and scientific research. The key areas of the future research may include: broader age range of participants and different populations (e.g. young athletes), diverse clinical conditions (e.g. cerebral palsy, muscular dystrophies, etc.), longer follow-up periods, investigations on validity of the tests, and finally incorporating them into experiments in the role of measurement tools.

Ethical Considerations

Compliance with ethical guidelines

The study was approved by the institutional Biomedical Research Ethics Committee (Code of Ethics: 18/2020).

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

Conceptualization, Resources, Writing – Original Draft Preparation and Writing – Review & Editing: All authors; Methodology and Supervision: Rafał Gnat, Anna Gogola, Tomasz Wolny; Software: Rafał Gnat; Validation, Formal Analysis and Project Administration: Rafał Gnat, Anna Gogola; Investigation: Anna Gogola, ;Data Curation: Agnieszka Polaczek, Piotr Woźniak; Visualization: Rafał Gnat, Agnieszka Polaczek, Piotr Woźniak; Funding Acquisition none

Conflict of interest

The authors declared no conflict of interest.

This Page Intentionally Left Blank

مقاله پژوهشی

پایایی آزمون‌های طول عضلات خم‌کننده و بازکننده مفصل زانو اصلاح‌شده متناسب با تحقیقات‌علمی

رافال گنات [۱]، آنا گوگولا [۱]، *آگنیشکا پولاچک [۱]، پیوتر ووژنیاک [۲]، توماس ولنی [۱]

۱. موسسه فیزیوتراپی و علوم بهداشتی، آزمایشگاه تحلیل حرکت، آکادمی تربیت بدنی جرزی کوکوچکا، کاتوویتس، لهستان.
۲. گروه پزشکی دریم موشن، لیبیاژ، لهستان.

Use your device to scan and read the article online

Citation Gnat R, Gogola A, Polaczek A, Woźniak P, Wolny T. Investigating the Reliability of the Knee Joint Flexors and Extensors Muscles Length Tests Modified to Meet the Requirements of Scientific Research. Archives of Rehabilitation. 2025; 26(2):230-253. https://doi.org/10.32598/RJ.26.2.3986.1

https://doi.org/10.32598/RJ.26.2.3986.1

چکیده

هدف این مطالعه با هدف ارزیابی پایایی درون و بین راستی نسخه‌های اصلاح‌شده آزمون‌های بالینی پرکاربرد برای ارزیابی طول عضلات خم‌کننده و بازکننده زانو انجام شد. این آزمون‌ها اجزای اساسی ارزیابی‌های وضعیتی در کودکان و نوجوانان هستند و به شناسایی ناهنجاری‌های وضعیتی و ارائه درمان مناسب یا استراتژی‌های پیشگیرانه کمک می‌کنند. یافته‌ها ی مطالعه نشان‌دهنده امکان استفاده از این نسخه‌های اصلاح‌شده در تحقیقات علمی خواهند بود.

روش بررسی ۷۶ کودک (۳۷ دختر) با میانگین سنی ۱۳/۲۴ سال در این مطالعه شرکت کردند. اندازه‌گیری‌ها توسط دو ارزیاب در ۱ روز و همچنین با وقفه‌های ۲ و ۵ روزه انجام شد. شاخص‌های پایایی شامل ضرایب همبستگی درون‌گروهی (ICC)، خطاهای استاندارد اندازه‌گیری و کوچک‌ترین تفاوت‌های قابل‌تشخیص محاسبه شدند.

یافته‌ها نتایج هر دو آزمون طول عضله، صرف‌نظر از اینکه اندازه‌گیری‌ها توسط همان ارزیاب پس از ۲ یا ۵ روز یا توسط ارزیاب‌های مختلف در همان روز و پس از فواصل مشابه انجام شده باشد، نشان‌دهنده پایایی خوب تا عالی بود. کمترین ICC محاسبه‌شده برابر با ۰/۷۹ مربوط به آزمون طول خم‌کننده‌های زانو در سمت چپ بدن بود. در تمام موارد باقیمانده، ICC بالاتر از ۰/۸۰ به دست آمد.

نتیجه‌گیری دو آزمون طول عضله موردنظر پس از اصلاحات مشخص سطح بالایی از پایایی نشان دادند که استفاده از نسخه‌های اصلاح‌شده را در محیط‌های علمی توجیه می‌کند.

کلیدواژه‌ها خم‌کننده‌های زانو، بازکننده‌های زانو، تحقیق، کودکان سالم، قابلیت اطمینان

تاریخ دریافت: ۳۰ آبان ۱۴۰۳
تاریخ پذیرش: ۲۵ فروردین ۱۴۰۴
تاریخ انتشار: ۱۰ تیر ۱۴۰۴

* **نویسنده مسئول:**
آگنیشکا پولاچک
نشانی: لهستان، کاتوویتس، آکادمی تربیت بدنی جرزی کوکوچکا، آزمایشگاه تحلیل حرکت، موسسه فیزیوتراپی و علوم بهداشتی.
تلفن: ۷۹۲۸۹۲۸۹۵ (۴۴) ۴۸+
رایانامه: a.polaczek11@gmail.com

Copyright © 2025 The Author(s);
This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-By-NC: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

مقدمه

در طول نیم قرن گذشته، فیزیوتراپی در هر دو حوزه علمی و بالینی شاهد رشدی پویا بوده است. همان‌طور که از تعداد فزاینده مجلات حرفه‌ای با تأثیر بالا و مطالعات تأثیرگذار در سراسر جهان مشهود است، فراتر از پیشرفت‌های بالینی، فیزیوتراپی در تحقیقات علمی نیز مورد توجه قرار گرفته است.

علی‌رغم این پیشرفت، فیزیوتراپیست‌های علمی با چالش‌های خاصی، به‌ویژه در مورد ابزارهای تشخیصی و اندازه‌گیری، مواجه هستند. یکی از مشکلات، فقدان ابزارهای دقیقی است که بتوان آن‌ها را به‌طور یکپارچه در تحقیقات ادغام کرد. این امر به پذیرش فناوری‌هایی از حوزه پزشکی منجر شده است. فیزیوتراپیست‌ها به‌صورت موفقیت‌آمیز از سونوگرافی برای مورفولوژی و ارزیابی فعالیت عضلات [۱-۷]، الکترومیوگرافی برای کاربردهای مختلف [۸، ۹]، MRI عملکردی برای فعالیت‌های حرکتی [۱۰، ۱۱]، بررسی حرکت برای تحرک ستون فقرات و پیکربندی لگن [۱۲-۱۶] و همچنین میوتونومتری [۱۷، ۱۸] و الاستوگرافی [۱۹، ۲۰] برای ارزیابی بیومکانیک بافت و اثرات درمانی [۲۱] استفاده می‌کنند.

بااین‌حال، مسئله مهم این است که فناوری پیشرفته ناکافی است، درحالی‌که ابزارهای پیشرفته داده‌های بسیار دقیقی ارائه می‌دهند، ممکن است در ثبت جنبه‌های عملکردی گسترده‌تر حرکت انسان - که تمرکز اساسی در فیزیوتراپی است - شکست بخورند. در برخی موارد، ارزیابی‌های بالینی ساده‌تر و کاربردی‌تری برای ارائه یک دیدگاه جامع مورد نیاز است. برای اینکه این آزمون‌های پرکاربرد مناسب تحقیقات علمی شوند، ابتدا باید آن‌ها را به‌گونه‌ای اصلاح کرد که با استانداردهای روش‌شناختی دقیق همخوانی داشته باشند. این موضوع سؤال‌های کلیدی را درباره قابلیت اعتماد تکنیک‌های ارزیابی اصلاح‌شده مطرح می‌کند.

ارزیابی پایایی آزمون‌های بالینی برای اطمینان از دقت، ثبات و کاربردپذیری آن‌ها در تحقیق و عمل بسیار مهم است. آزمون‌های قابل‌اعتماد به پزشکان اجازه می‌دهند تا تغییرات در انعطاف‌پذیری و عملکرد عضلات را در طول زمان پیگیری کنند و اطمینان حاصل کنند که تفاوت‌های مشاهده‌شده ناشی از تغییرات فیزیولوژیکی واقعی است و نه خطاهای اندازه‌گیری. در محیط بالینی، این امر به بهبود تصمیم‌گیری درمورد تشخیص، برنامه‌ریزی درمان و پیشرفت توانبخشی منجر می‌شود. در تحقیقات، قابلیت اطمینان بالا برای تولید یافته‌های تکرارپذیر و قابل‌مقایسه ضروری است که اعتبار نتیجه‌گیری‌های علمی را تقویت کرده و از عملکرد مبتنی بر شواهد پشتیبانی می‌کند. بدون انجام آزمون‌های مناسب برای بررسی قابلیت اطمینان، حتی ارزیابی‌های پرکاربرد نیز ممکن است نتایج ناهماهنگی را ارائه دهند که به تفسیرهای نادرست و مداخلات ناموثر منجر شود.

این تحقیق با تمرکز بر تطبیق دو تکنیک ارزیابی طول عضلات بالینی که به‌طور گسترده مورد استفاده قرار می‌گیرند انجام شد: یکی برای عضلات خم‌کننده زانو (همسترینگ) و دیگری برای عضلات بازکننده زانو (چهارسر ران). این گروه‌های عضلانی نقش مهمی در مکانیسم‌های جبران وضعیت بدن دارند و اغلب تحت تأثیر سبک زندگی بی‌تحرک در کودکان و نوجوانان قرار می‌گیرند [۲۲]. کوتاهی عضلات و عدم تعادل قدرت نه‌تنها در جوانان غیرفعال، بلکه در ورزشکاران جوان نیز اغلب مشاهده می‌شود [۲۳-۲۵]. درنتیجه، این آزمایش‌ها بخش جدایی‌ناپذیری از ارزیابی‌های وضعیتی در طول رشد هستند.

نسخه‌های مختلفی از آزمون‌های طول خم‌کننده و بازکننده زانو در مقالات علمی وجود دارد. به‌طور خاص، آزمون خم‌کننده‌های زانو به‌طور گسترده مورد مطالعه قرار گرفته است و محققان تکنیک‌های اندازه‌گیری مختلفی را به کار برده‌اند که درجات مختلفی از پایایی را به همراه داشته است [۲۲، ۲۶-۳۵]. بااین‌حال، همچنان نیاز به نسخه‌های ساده‌شده، کم‌هزینه و کاربرپسند که همچنان استانداردهای تحقیقات علمی را برآورده کنند، وجود دارد. علاوه‌براین، جنبه‌های خاصی از پایایی- به‌ویژه فواصل طولانی‌تر بین اندازه‌گیری‌های آزمون بازآزمون- در مطالعات قبلی به اندازه کافی بررسی نشده‌اند.

بنابراین، هدف این مطالعه ارزیابی قابلیت اطمینان درون و بین ارزیاب آزمون‌های اصلاح‌شده طول انعطاف‌پذیری زانو و کشش زانو، بررسی پایایی آن‌ها در همان روز و همچنین در فواصل ۲ روزه و ۵ روزه در کودکان ۱۰ تا ۱۵ ساله بود. یافته‌ها مشخص خواهد کرد که آیا نسخه‌های اصلاح‌شده آزمون‌ها می‌توانند به‌طور مؤثر در محیط‌های تحقیقاتی علمی اجرا شوند یا خیر؟

مواد و روش‌ها

شرکت‌کنندگان

۸۲ کودک از سه مدرسه ابتدایی محلی داوطلب شدند، ۴۲ پسر و ۴۰ دختر از کلاس‌های چهارم تا هشتم (۱۰ تا ۱۵ سال). شرکت‌کنندگان عموماً کودکان سالمی بودند که سطوح مختلفی از فعالیت بدنی روزانه (به‌جای نشستن، فعالیت تفریحی، فعالیت ورزشی) را انجام می‌دادند.

معیارهای ورود به مطالعه سن بین ۱۰ تا ۱۵ سال (به‌منظور پوشش کل دوره جهش بلوغ)؛ رشد عصبی و حرکتی طبیعی (بدون تشخیص پزشکی هیچ بیماری یا اختلال عملکردی، بین ۲۵ تا ۷۵ صدک توده بدنی و قد طبیعی براساس سن) و توانایی پیروی از دستورات کلامی بودند. شرکت‌کنندگان در صورت داشتن سابقه یا تشخیص فعلی هرگونه اختلال ارتوپدی یا عصبی قابل‌توجه (مانند شکستگی، ناهنجاری‌های مادرزادی، فلج مغزی، یا درد و/یا اختلال عملکرد اسکلتی-عضلانی که بیش از ۲ هفته طول بکشد) از مطالعه حذف شدند.

معیارهای خروج دیگر شامل هرگونه سابقه عمل جراحی، موارد اخیر مشکلات جزئی اسکلتی-عضلانی در ۱ ماه قبل از مطالعه، یا تجربه شکایات جزئی سلامتی (مانند سرماخوردگی و سردرد) در روز ارزیابی بود. ۲ دختر به‌دلیل سابقه مداخلات جراحی در ناحیه شکم، و ۱ دختر و ۳ پسر به‌دلیل سابقه شکستگی در اندام تحتانی از مطالعه حذف شدند.

محدوده سنی انتخاب‌شده به چند دلیل انتخاب شد. اولاً، این محدوده شامل کل جهش رشد بلوغ می‌شود -دوره‌ای که طی آن پزشکان به‌طور معمول ارزیابی‌های وضعیتی را در کودکان انجام می‌دهند، ازجمله دو آزمایش طول عضله موردبررسی و زمانی که ناهنجاری‌های وضعیتی متعددی معمولاً شناسایی می‌شوند. تنوع در ویژگی‌های وضعیتی شرکت‌کنندگان به تنوع نتایج کمک کرد و چالشی برای اندازه‌گیری‌ها ایجاد نمود، درحالی‌که همچنین پتانسیل تعمیم نتایج را افزایش داد. به همین دلیل، شرکت‌کنندگانی با سطوح مختلف فعالیت بدنی روزانه نیز در نظر گرفته شدند.

در نهایت، ۷۶ کودک (۳۷ دختر) که واجد شرایط مطالعه بودند (میانگین سنی ۱۳/۲۴ (۱۰-۱۵) سال، جرم بدن ۵۱/۲۷ (۳۰-۷۲) کیلوگرم، قد بدن ۱/۵۳۴ (۱/۳۵-۱/۷۴۵) متر) به همراه والدینشان اطلاعات دقیقی در مورد اهداف و روش‌ها دریافت کردند. والدین همچنین رضایت آگاهانه خود را اعلام کردند. اندازه‌گیری‌ها در یکی از مراکز دانشگاهی محلی، در آزمایشگاه آنالیز حرکت، انجام شد. در طول انجام اندازه‌گیری‌ها، هیچ موردی از انصراف ثبت نشد.

حداقل حجم نمونه با استفاده از محاسبه‌گر حجم نمونه [۳۶] تخمین زده شد. بیشترین تعداد شرکت‌کننده (۳۵ نفر) برای محاسبات مربوط به آزمون طول عضلات بازکننده زانو براساس یک اندازه‌گیری مکرر با حداقل ضریب همبستگی درون‌گروهی (ICC=۰/۷۰) و موردانتظار (ICC=۰/۸۵)مورد نیاز بود.

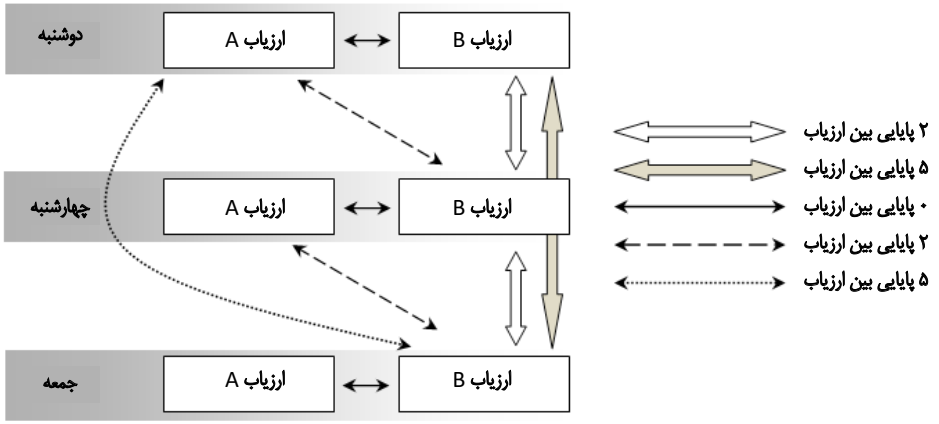
طرح مطالعه

این تحقیق مطالعه‌ای فنی با اندازه‌گیری‌های مکرر متغیرهای موردنظر بود که توسط ۲ ارزیاب در ۳ روز اندازه‌گیری که با وقفه‌های ۱ روزه انجام شد **(تصویر شماره ۱)**.

برای ثبت تمام مقایسه‌های در نظر گرفته‌شده، پیکان‌های خطی باریک (که نمایانگر قابلیت اطمینان بین ارزیاب هستند) که بلوک‌های ارزیاب A و ارزیاب B را به هم متصل می‌کنند، باید به سمت مقابل آینه‌وار شوند. به همین ترتیب، پیکان‌های بلوکی توپر (که قابلیت اطمینان درون ارزیاب را نشان می‌دهند) که بلوک‌های متوالی برای ارزیاب B را به هم متصل می‌کنند نیز باید به‌طور مناسب آینه‌وار شوند. به‌منظور جلوگیری از تکرار در حین تحلیل داده‌ها، تمام مقایسه‌هایی که با نوع مشابهی از پیکان‌ها در **تصویر شماره ۱** مشخص شده بودند، گردآوری شدند. قابلیت اطمینان درون ارزیاب ۲/۵ به قابلیت اطمینان بین اندازه‌گیری‌های انجام‌شده توسط همین ارزیاب با فاصله‌های ۲ و ۵ روز اشاره دارد، درحالی‌که قابلیت اطمینان بین ارزیاب ۰/۲/۵ به قابلیت اطمینان اندازه‌گیری‌هایی که توسط ارزیابان مختلف در همان روز و همچنین پس از ۲ و ۵ روز انجام شده‌اند، اشاره می‌کند.

ارزیابان

برای ایجاد محیط آزمایشی سخت‌تر، این مطالعه شامل دو ارزیاب با تجربه حرفه‌ای نسبتاً محدود بود. هر دو فیزیوتراپیست دارای مجوز بودند و هر کدام ۲/۵ تا ۳/۵ سال سابقه کار بالینی داشتند. ارزیابان یک برنامه آموزشی فشرده ۲ هفته‌ای (۳ ساعت×۳ بار در هفته) را برای کسب مهارت کافی در انجام آزمون طول دو عضله به پایان رساندند. این آموزش توسط متخصصی باتجربه که در مطالعه دخیل نبود، نظارت شد. پس از اتمام، متخصص ارزیابی و تأیید کرد که ارزیابان مهارت‌های لازم را برای انجام شایسته نقش خود در تحقیق کسب کرده‌اند. مهارت‌های



توانبخشی

تصویر ۱. مدل اندازه‌گیری پایایی بین ارزیاب‌ها

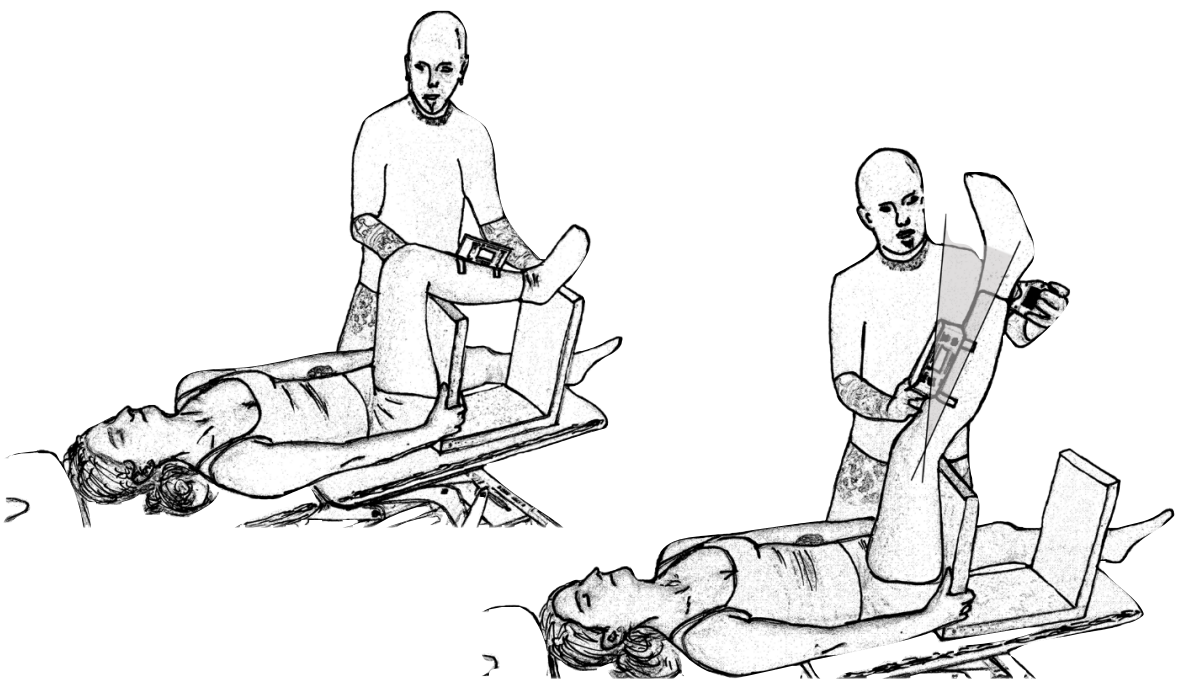
ارزیابان همچنین در یک مطالعه آزمایشی جزئی که بر روی ۱۵ داوطلب بزرگسال انجام شد، مورد بررسی قرار گرفت. در این مطالعه، کمترین ضریب همبستگی درون‌گروهی (ICC=۳/۳) ثبت‌شده برابر با ۰/۹۱ بود (ICC بین ارزیابان برای آزمون طول عضلات بازکننده زانو). به دلایل نامشخص، یک دستیار نیز استخدام شد که وظیفه‌اش خواندن و ثبت مقادیر خوانده‌شده توسط دستگاه‌های اندازه‌گیری بود که از ارزیابان پنهان بودند.

آزمون طول خم‌کننده‌های زانو

فرد موردآزمایش به حالت طاقباز روی کاناپه دراز کشید (**تصویر شماره ۲**)، درحالی‌که لگن و زانوی پای موردآزمایش تا ۹۰ درجه خم شده بود و برای حفظ تعادل، توسط یک چهارپایه کوچک که با دستان فرد در جای خود نگه داشته شده بود، پشتیبانی می‌شد. دو نقطه مرجع روی قسمت قدامی ساق پا، ۵ سانتی‌متر و ۱۵ سانتی‌متر پایین‌تر از کشکک زانو (با زانوی کشیده) علامت‌گذاری شدند. آهن‌رباهای نئودیمیوم تخت (با قطر ۱/۵ سانتی‌متر) با نوار چسب به هم متصل شدند و لبه‌های پروگزیمال آن‌ها با این علامت‌ها هم‌تراز شد تا تنظیمات اندازه‌گیری ساده شده و خطای دستی به حداقل برسد. یک شیب‌سنج دیجیتال Base-line ۱۲-۱۰۵۷ (محصولات Baseline، ایالات متحده) به‌صورت عمودی کالیبره شد و به‌گونه‌ای قرار گرفت که لبه پروکسیمل آن با آهن‌ربای بالایی هم‌راستا باشد. نمایشگر از دید ارزیاب خارج نگه داشته شد. همچنین یک خط افقی بر روی پوست بین ماله‌ها ترسیم شد. نیروی گسترش زانوی غیرفعال با استفاده از یک گیج نیرو Steinberg SBS-KW-300A اعمال شد که در قسمت خلفی پای پایین در خط علامت‌گذاری‌شده قرار گرفت. ارزیاب گیج را به‌صورت عمود بر پا نگه داشت تا اطمینان حاصل شود که نیروی اعمال‌شده به‌طور یکنواخت است. دستورالعمل‌های کلامی به این صورت بود: «نفس بکشید، نفس را بیرون دهید، بگذارید پای شما به آرامی بالا بیاید. وقتی احساس کشش قوی اما قابل‌تحملی در پشت ران خود کردید، بگویید 'توقف'». یک آزمایش اولیه، نیروی موردنیاز برای رسیدن به نقطه "متوقف" را تعیین کرد. این نیرو در ۳ تکرار آزمون تکرار شد. کنترل وزن پای پایین مطابق با یافته‌های گوکس و همکاران [۳۰] نادیده گرفته شد. در فرمان "حالا" ، دستیار خواندن شیب‌سنج را ثبت کرد و از دید ارزیاب جلوگیری شد. این روش برای پای مقابل نیز تکرار شد.

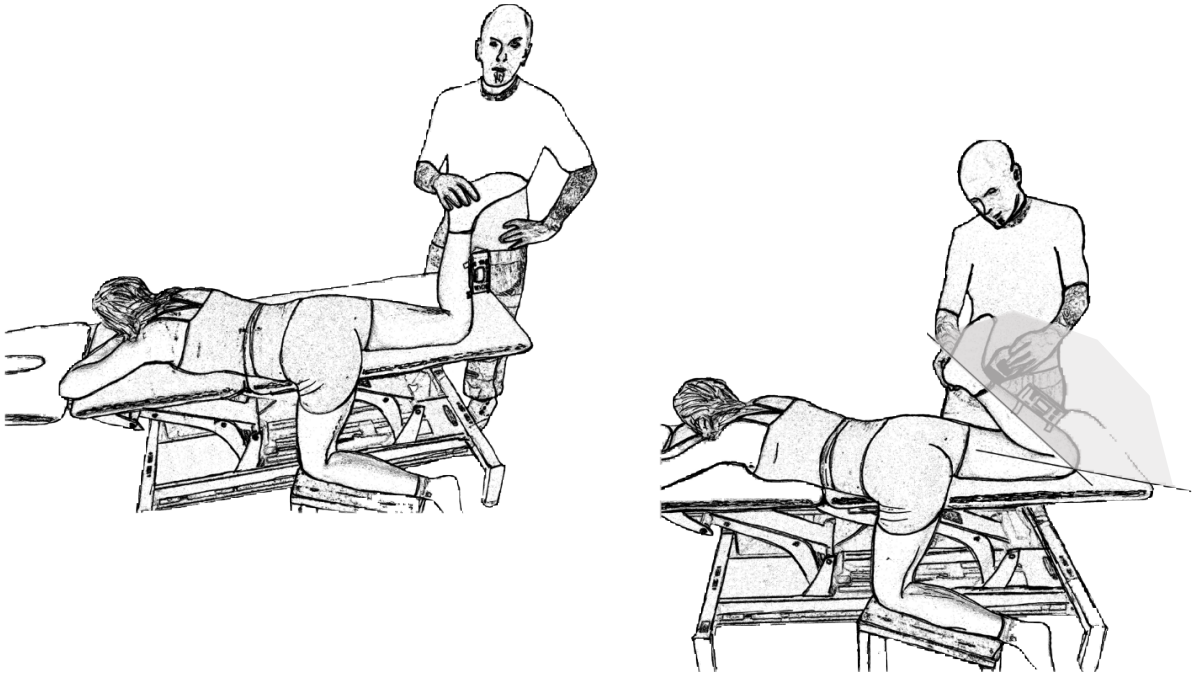
آزمون طول عضلات بازکننده زانو

فرد موردآزمایش به‌صورت مورب روی مبل دراز کشید (**تصویر شماره ۳**). پای موردآزمایش آزادانه از لبه میز آویزان بود، درحالی‌که زانو (تقریباً ۹۰ درجه خم شده بود) روی یک چهارپایه قرار داشت. ارتفاع مبل برای راحتی تنظیم شد. شیب‌سنج به همان روشی که قبلاً توضیح داده شد، تنظیم شد. به‌طور مشابه، نیروسنج مانند قبل استفاده شد، اما برای ایجاد خم شدن غیرفعال زانو، به قسمت قدامی ساق پا اعمال شد. ارزیاب دستورالعمل‌های



توانبخشی

تصویر ۲. تست اصلاح‌شده طول عضلات خم‌کننده زانو.
وضعیت اولیه (شکل چپ) و وضعیت نهایی (شکل راست). تست ابتدا برای پای راست و سپس برای پای چپ انجام شد. در شکل سمت راست یک قوس که نشان‌دهنده معیار اندازه‌گیری می‌باشد با رنگ خاکستری نشان داده شده است.



توانبخشی

تصویر ۳. تست اصلاح‌شده طول عضلات بازکننده زانو.
وضعیت اولیه (شکل چپ) و وضعیت نهایی (شکل راست). تست ابتدا برای پای راست و سپس برای پای چپ انجام شد. در شکل سمت راست یک قوس که نشان‌دهنده معیار اندازه‌گیری می‌باشد با رنگ خاکستری نشان داده شده است.

زیر را ارائه داد: «دم، بازدم کنید و اجازه دهید پای شما به آرامی به سمت بالا حرکت کند. در پایان، وقتی کشش قوی اما قابل‌تحملی، نه درد، در جلوی ران خود احساس کردید، بگویید «توقف». همانند آزمون خم‌کننده‌ها، یک آزمایش اولیه نیروی موردنیاز برای رسیدن به نقطه «توقف» را تعیین کرد و ثبات در تکرارهای بعدی را تضمین نمود. آزمون طول عضله واقعی شامل سه تکرار بود که طی آن، ارزیاب سطح نیروی ثبت‌شده قبلی را تکرار می‌کرد. با دستور «حالا»، دستیار، عدد شیب‌سنج را ثبت کرد که باز هم برای ارزیاب مبهم بود. پس از آزمایش یک پا، این روش در سمت مقابل تکرار شد.

پس از مصاحبه مقدماتی و تأیید معیارهای انتخاب، هفته مناسب برای اندازه‌گیری‌ها توسط والدین و کودکان انتخاب شد. آن‌ها بعدازظهر پس از حداقل ۳ ساعت استراحت پس از مدرسه و حداقل ۲ ساعت پس از آخرین وعده غذایی در آزمایشگاه حاضر شدند. از کودکان خواسته شد لباس‌های بدون محدودیت بپوشند و با استفاده از دوچرخه ثابت به‌مدت ۱۰ دقیقه با شدت کم بدن خود را گرم کنند. سپس، به آزمودنی‌ها اطلاعاتی در مورد نحوه همکاری در طول اجرای آزمون داده شد و ۲-۳ تلاش از هر آزمون معرفی شد. بااین‌حال، این تلاش‌ها بدون بررسی کامل دامنه حرکتی آزمودنی‌ها انجام شد. ترتیب آزمون‌ها برای شرکت‌کننده موردنظر به‌طور تصادفی در اولین روز اندازه‌گیری (دوشنبه) تعیین شد و در روزهای متوالی اندازه‌گیری (چهارشنبه، جمعه همان هفته) بدون تغییر باقی ماند.

ترتیب قرارگیری طرفین بدن همیشه از راست به چپ بود. حروف A و B به‌صورت تصادفی به ارزیابان اختصاص داده شدند. پس از این، ترتیب ارزیابان همیشه از A تا B باقی ماند. روز دوشنبه، پس از آماده‌سازی مناسب آزمودنی‌ها، دو آزمون طول عضله توسط ارزیاب A، هر کدام ۳ بار، تکرار شد. سپس، تمام علامت‌های روی پوست آزمودنی‌ها، آهن‌رباهای متصل و غیره، برداشته شدند و آزمودنی‌ها به‌مدت ۱۵ دقیقه در حالت نیمه نشسته و آرام روی صندلی راحتی استراحت کردند. متعاقباً، ارزیاب B دوباره آزمودنی‌ها را برای اندازه‌گیری‌های خود آماده کرد و ۲ آزمایش ۱ بار دیگر، هر کدام ۳ بار، تکرار شد. کل این مراحل به‌طور یکسان در روزهای چهارشنبه و جمعه تکرار شد **(تصویر شماره ۱)**. به شرکت‌کنندگان گفته شد سطح فعالیت بدنی معمول خود را در طول کل هفته اندازه‌گیری حفظ کنند و از ایجاد هرگونه تغییر ناگهانی در فعالیت معمول خود خودداری کنند.

پردازش داده‌ها

داده‌های ثبت‌شده توسط دو ارزیاب در صفحه گسترده Sta-tistica ۱۳ (Statistica، تولسا، ایالات متحده آمریکا) جمع‌آوری شد. به‌منظور محاسبه شاخص‌های پایایی برای انواع مختلف پایایی **(تصویر شماره ۱)** و برای جلوگیری از افزونگی نتایج، این پایگاه داده به‌طور خاص بازسازی شد. مقایسه‌های مشخص‌شده براساس طرح زیر تجمیع شدند:

قابلیت اطمینان درون ارزیاب، با فاصله ۲ روز بین اندازه‌گیری‌های متوالی

ارزیاب A در روز دوشنبه در مقابل ارزیاب A در روز چهارشنبه + ارزیاب A در روز چهارشنبه در مقابل ارزیاب A در روز جمعه + ارزیاب B در روز دوشنبه در مقابل ارزیاب B در روز چهارشنبه + ارزیاب B در روز چهارشنبه در مقابل ارزیاب B در روز جمعه؛ مجموع ۳۰۴ رکورد.

قابلیت اطمینان درون ارزیاب، با فاصله ۵ روز بین اندازه‌گیری‌های متوالی

ارزیاب A در روز دوشنبه در مقابل ارزیاب A در روز جمعه+ارزیاب B در روز دوشنبه در مقابل ارزیاب B در روز جمعه؛ مجموع ۱۵۲ رکورد.

قابلیت اطمینان درون ارزیاب، اندازه‌گیری‌ها در همان روز انجام شده است

ارزیاب A در روز دوشنبه در مقابل ارزیاب B در روز دوشنبه + ارزیاب A در روز چهارشنبه در مقابل ارزیاب B در روز چهارشنبه + ارزیاب A در روز جمعه در مقابل ارزیاب B در روز جمعه؛ مجموع ۲۲۸ رکورد.

قابلیت اطمینان بین ارزیاب، با فاصله ۲ روز بین اندازه‌گیری‌های متوالی

ارزیاب A در روز دوشنبه در مقابل ارزیاب B در روز چهارشنبه+ارزیاب A در روز چهارشنبه در مقابل ارزیاب B در روز جمعه+ارزیاب B در روز دوشنبه در مقابل ارزیاب A در روز چهارشنبه+ارزیاب B در روز چهارشنبه در مقابل ارزیاب A در روز جمعه؛ مجموع ۳۰۴ رکورد.

قابلیت اطمینان بین ارزیاب، با فاصله ۵ روز بین اندازه‌گیری‌های متوالی

ارزیاب A در روز دوشنبه در مقابل ارزیاب B در روز جمعه +ارزیاب B در روز دوشنبه در مقابل ارزیاب A در روز جمعه؛ مجموع ۱۵۲ رکورد.

تحلیل آماری

از مدل آنووا[1] عاملی برای مقایسه میانگین‌های آزمون‌های طول عضله که توسط دو ارزیاب در روزهای اندازه‌گیری متوالی به دست آمدند، استفاده شد. در این مدل، روز اندازه‌گیری، ارزیاب و سمت بدن به‌عنوان عوامل مستقل در نظر گرفته شدند.

برای محاسبه ICC، از آنووا مدل مختلط استفاده شد که در آن عامل تکراری اندازه‌گیری‌های متوالی و عامل مستقل، افراد بودند. قابلیت اطمینان اندازه‌گیری‌های انجام‌شده در همان روز توسط یک ارزیاب به‌طور جداگانه گزارش نشده است، زیرا به‌طور ذاتی در سایر اشکال قابلیت اطمینان درون ارزیاب گنجانده شده است **(تصویر شماره ۱)**. از مدل k،۲ از ICC به کار گرفته شد تا امکان تعمیم نتایج به کل جمعیت ارزیابان مشابه فراهم شود. ICCها به‌طور جداگانه برای نتایج یک تکرار از آزمون طول عضله داده‌شده، ۲ تکرار (مقدار میانگین) و ۳ تکرار (مقدار میانگین) محاسبه شدند. علاوه‌براین، خطاهای استاندارد اندازه‌گیری (SEM) با استفاده از **فرمول شماره ۱** همراه با کوچک‌ترین تفاوت‌های قابل‌شناسایی محاسبه شدند **(فرمول شماره۲)**. تفسیر مقادیر ICC به این ترتیب بود: ۰/۰۰-۰/۵۰=ضعیف، ۰/۵۰-۰/۷۵=متوسط، ۰/۷۵-۰/۹۰=خوب و ۰/۹۰-۱/۰۰=قابلیت اطمینان عالی **[۳۷]**.

1. $SEM = SD \times (1 - ICC)^{1/2}$

2. $SDD = 1/96 \times 2^{1/2}$

یافته‌ها

در **جدول شماره ۱**، آمار توصیفی نتایج دو آزمون طول عضله موردنظر که توسط هر دو ارزیاب در روزهای اندازه‌گیری متوالی ثبت شده، ارائه شده است. برای تمامی اثرات اصلی آنووا (روز اندازه‌گیری، ارزیاب، سمت بدن) و همچنین تعاملات آن‌ها، هیچ تفاوت معنی‌داری مشاهده نشد (همه $P>0/05$). تفاوت بین روزهای اندازه‌گیری، ارزیاب‌ها و طرفین بدن، همگی غیرمعنی‌دار بودند ($P>0/05$).

دو آزمون طول عضله موردنظر، قابلیت اطمینان خوب تا عالی را نشان دادند، بدون توجه به اینکه اندازه‌گیری‌ها توسط همان ارزیاب با فاصله ۲ یا ۵ روز انجام شده باشد **(جدول شماره ۲)** یا توسط ارزیابان مختلف در همان روز و همچنین با فاصله ۲ یا ۵ روز **(جدول شماره ۳)**. پایین‌ترین ICC محاسبه‌شده برابر با ۰/۷۹ مربوط به آزمون طول خم‌کننده‌های زانو در سمت چپ بدن بود (SEM=۵/۴۷، SDD=۱۵/۱۷) درجه؛ قابلیت اطمینان بین ارزیاب براساس یک اندازه‌گیری تکراری که با فاصله ۵ روز انجام شد). در تمامی موارد باقیمانده، ICCهای بالاتر از ۰/۸۰ به‌دست آمد.

در مورد قابلیت اطمینان درون ارزیاب **(جدول شماره ۲)**، حتی یک اندازه‌گیری مکرر نیز برای به‌دست آوردن ICCهای بالاتر از ۰/۸۷ کافی بود. افزایش تعداد اندازه‌گیری‌های مکرر باعث افزایش بیشتر مقادیر ICC شد. در طول ۳ آزمایش تکراری، مقادیر ICC به ۰/۹۴ رسید یا از آن فراتر رفت. کاهش تدریجی و هم‌زمان در مقادیر SEM و SDD نیز ثبت شد.

ضرایب همبستگی درون‌گروهی (فاصله اطمینان ICC ± ۹۵%) که با استفاده از مدل‌های ۱، ۲ برای اندازه‌گیری‌های تکی، ۲، ۲ برای ۲ تکرار و ۲، ۳ برای ۳ اندازه‌گیری مکرر محاسبه شده‌اند، گزارش شده‌اند. علاوه‌براین، خطای استاندارد اندازه‌گیری (SEM)

1. ANOVA

جدول ۱. مقادیر ثبت‌شده در آزمون‌های طول خم‌کننده‌ها و بازکننده‌های زانو براساس سه ارزیابی مکرر انجام‌شده توسط هر ارزیاب در ۳ روز اندازه‌گیری متوالی

روزهای هفته	ارزیاب	میانگین±انحراف‌معیار/حداقل–حداکثر			
		خم‌کننده‌های زانو		بازکننده‌های زانو	
		چپ (درجه)	راست (درجه)	چپ (درجه)	راست (درجه)
دوشنبه	A	۱۸/۸±۱۲/۴۸ (۵۳/۶۷–۹/۰۰)	۱۰/۸۳۴۰±۲۷/۸۳ (۴۹/۳۳–۸/۰۰)	۱۱۷/۸۰±۴/۶۶ (۱۳۰/۶۷–۱۱۰/۰۰)	۱۱۸/۰۷±۵/۴۸ (۱۳۵/۳۳–۱۱۰/۰۰)
	B	۲۹/۰۵±۱۲/۸۳ (۵۴/۳۳–۷/۰۰)	۲۷/۱۲±۹/۶۵ (۴۷/۶۷–۹/۳۳)	۱۱۵/۹۵±۵/۱۲ (۱۳۲/۳۳–۱۰۸/۳۳)	۱۱۷/۳۵±۵/۶۹ (۱۳۷/۳۳–۱۰۸/۰۰)
چهارشنبه	A	۹۴/۲۷±۱۳/۵۳ (۵۳/۳۳–۶/۰۰)	۲۸/۴۸±۱۲/۴۸ (۵۳/۶۷–۹/۰۰)	۱۱۷/۹۰±۵/۴۵ (۱۳۴/۳۳–۱۱۰/۶۷)	۱۱۸/۱۹±۵/۴۵ (۱۳۵/۳۳–۱۱۰/۶۷)
	B	۲۷/۱۵±۱۳/۸۹ (۵۴/۶۷–۵/۳۳)	۲۹/۲۱±۱۳/۰۵ (۵۴/۶۷–۷/۶۷)	۱۱۶/۹۸±۵/۷۸ (۱۳۵/۰۰–۱۰۸/۶۷)	۱۱۹/۰۱±۵/۸۱ (۱۳۶/۶۷–۱۰۸/۳۳)
جمعه	A	۲۷/۵۶±۱۲/۵۸ (۵۴/۶۷–۸/۳۳)	۲۸/۱۳±۱۰/۸۱ (۵۰/۶۷–۹/۳۳)	۱۱۸/۱۲±۴/۹۵ (۱۳۱/۳۳–۱۰۹/۳۳)	۱۱۷/۳۵±۵/۷۸ (۱۳۵/۰۰–۱۰۹/۰۰)
	B	۲۸/۰۶±۱۳/۲۶ (۵۳/۶۷–۵/۶۷)	۲۸/۳۴±۱۲/۳۳ (۵۳/۳۳–۸/۰۰)	۱۱۷/۳۲±۶/۰۵ (۱۳۴/۶۷–۱۰۹/۳۳)	۱۱۸/۵۹±۵/۱۱ (۱۳۵/۳۳–۱۰۹/۶۷)

توانبخشی

و کوچک‌ترین تفاوت قابل‌تشخیص (SDD) ارائه شده‌اند. پایایی درون ارزیاب ۲ به ثبات اندازه‌گیری‌های انجام‌شده توسط یک ارزیاب با فاصله ۲ روزه (مثلاً دوشنبه تا چهارشنبه و چهارشنبه تا جمعه) اشاره دارد، درحالی‌که پایایی درون ارزیاب ۵ مربوط به اندازه‌گیری‌های انجام‌شده با فاصله ۵ روزه (مثلاً دوشنبه تا جمعه) است **(تصویر شماره ۱)**.

برای پایایی بین ارزیاب **(جدول شماره ۳)**، مقادیر ICC پایین‌تری به‌خصوص با فاصله زمانی طولانی‌تر بین اندازه‌گیری‌ها به دست آمد. اغلب این اتفاق می‌افتاد که یک اندازه‌گیری مکرر، ICCهایی کمتر از ۰/۹۰ را نشان می‌داد (کمتر از ۰/۸۰ در یک مورد که قبلاً در این بخش ذکر شد). افزایش تعداد اندازه‌گیری‌های مکرر باعث افزایش قابل‌توجه مقادیر ICC شد.

با ۳ اندازه‌گیری مکرر، تمام ICCها، حتی آن‌هایی که برای اندازه‌گیری‌های انجام‌شده با وقفه ۵ روزه محاسبه شده‌اند، بالاتر از ۰/۹۰ بودند. مجدداً، تمام SEMها و SDDها به‌تدریج کاهش یافتند.

برای هر نوع قابلیت اطمینان تحت بررسی، افزایش تعداد اندازه‌گیری‌های مکرر باعث باریک شدن تدریجی فواصل اطمینان ICCها شد، حتی زمانی که مقدار ICC دیگر به‌طور قابل‌توجهی افزایش نمی‌یافت. مقادیر زیر ارائه شده‌اند:

ضرایب همبستگی درون‌گروهی (فاصله اطمینان ICC±۹۵%)، براساس مدل ۱، ۲ برای اندازه‌گیری‌های تکی، ۲، ۲ برای ۲ اندازه‌گیری مکرر و ۲، ۳ برای ۳ اندازه‌گیری مکرر، خطای استاندارد اندازه‌گیری (SEM) و کوچک‌ترین تفاوت قابل‌تشخیص (SDD). پایایی بین ارزیاب (۰) به اندازه‌گیری‌های انجام‌شده توسط ارزیابان مختلف در همان روز اشاره دارد، درحالی‌که پایایی بین ارزیاب ۲ نشان‌دهنده اندازه‌گیری‌های انجام‌شده با فاصله ۲ روزه (مثلاً دوشنبه–چهارشنبه وچهارشنبه–جمعه) و پایایی بین ارزیاب ۵ مربوط به فاصله ۵ روزه (مثلاً دوشنبه–جمعه) است **(تصویر شماره ۱)**.

بحث

نتایج مطالعه حاضر نشان می‌دهد دو آزمون اصلاح‌شده طول عضله برای آزمون‌های خم‌کننده‌های زانو و بازکننده‌های زانو، سطح بسیار رضایت‌بخشی از پایایی را نشان می‌دهند (برای ۳ اندازه‌گیری مکرر، تمام ICCهای محاسبه‌شده بالاتر از ۰/۹۰ بودند). این یافته نشان‌دهنده کارایی این آزمون‌ها در محیط‌های تحقیقات علمی پشتیبانی است و جایگزینی قابل‌اعتماد برای ابزارهای تکنولوژیکی بسیار پیچیده می‌باشند. این ابزارها ممکن است گاهی اوقات نتوانند جنبه‌های عملکردی گسترده‌تر حرکت انسان را که آزمایش‌های بالینی می‌توانند به آن‌ها بپردازند، به تصویر بکشند. این یافته‌ها به‌ویژه برای محققانی که قصد دارند

جدول ۲. تخمین‌های پایایی درون ارزیاب

متغیر	پایایی	تعداد اندازه‌گیری‌ها	چپ				راست			
			ICC (±۹۵%CI)	خطای استاندارد اندازه‌گیری (درجه)	کوچک‌ترین تفاوت‌های قابل‌شناسایی (درجه)	ICC (±۹۵%CI)	خطای استاندارد اندازه‌گیری (درجه)	کوچک‌ترین تفاوت‌های قابل‌شناسایی (درجه)
خم‌کننده‌های زانو	درون ارزیاب ۲	۱	۰/۹۸ (۰/۹۴–۰/۸۵)	۴/۰۹	۱۱/۳۵	۰/۸۷ (۰/۹۳–۰/۸۲)	۴/۴۵	۱۲/۳۳
		۲ (میانگین)	۰/۹۴ (۰/۹۷–۰/۹۱)	۳/۰۲	۸/۳۸	۰/۹۱ (۰/۹۴–۰/۸۸)	۳/۷۰	۱۰/۲۶
		۳ (میانگین)	۰/۹۶ (۰/۹۸–۰/۹۴)	۲/۴۷	۶/۸۴	۰/۹۵ (۰/۹۸–۰/۹۲)	۲/۷۶	۷/۶۵
	درون ارزیاب ۵	۱	۰/۸۷ (۰/۹۱–۰/۸۳)	۴/۳۳	۱۲/۰۱	۰/۸۹ (۰/۹۳–۰/۸۵)	۳/۹۹	۱۱/۰۵
		۲ (میانگین)	۰/۹۳ (۰/۹۶–۰/۹۱)	۳/۱۸	۸/۸۲	۰/۹۲ (۰/۹۵–۰/۹۰)	۳/۴۰	۹/۴۲
		۳ (میانگین)	۰/۹۵ (۰/۹۷–۰/۹۳)	۲/۶۹	۷/۴۵	۰/۹۵ (۰/۹۶–۰/۹۳)	۲/۶۸	۷/۴۲
عضلات بازکننده‌های زانو	درون ارزیاب ۲	۱	۰/۸۹ (۰/۹۴–۰/۸۵)	۱/۷۳	۴/۸۰	۰/۹۰ (۰/۹۴–۰/۸۷)	۱/۷۹	۴/۹۵
		۲ (میانگین)	۰/۹۵ (۰/۹۸–۰/۹۲)	۱/۰۹	۳/۰۳	۰/۹۶ (۰/۹۸–۰/۹۳)	۱/۱۴	۳/۱۶
		۳ (میانگین)	۰/۹۶ (۰/۹۷–۰/۹۵)	۱/۱۸	۳/۲۸	۰/۹۶ (۰/۹۸–۰/۹۵)	۱/۰۴	۲/۸۸
	درون ارزیاب ۵	۱	۰/۸۸ (۰/۹۳–۰/۸۴)	۱/۸۰	۵/۰۰	۰/۷۸ (۰/۹۲–۰/۸۴)	۱/۸۳	۵/۰۹
		۲ (میانگین)	۰/۹۲ (۰/۹۶–۰/۸۹)	۱/۴۷	۴/۰۸	۰/۹۲ (۰/۹۵–۰/۹۰)	۱/۴۴	۳/۹۹
		۳ (میانگین)	۰/۹۵ (۰/۹۷–۰/۹۳)	۱/۱۶	۳/۲۳	۰/۹۴ (۰/۹۶–۰/۹۲)	۱/۲۴	۳/۴۶

توانبخشی

اندازه‌گیری‌های بالینی را در مطالعات خود پیاده‌سازی کنند، مفید است، جایی که روش‌های ساده‌تر و در دسترس‌تر ترجیح داده می‌شوند.

ازنظر پایایی مطلق، SDD ما برای ۳ اندازه‌گیری مکرر، حداکثر ۹/۹۳ درجه برای خم‌کننده‌های زانو و ۴/۴۸ درجه برای بازکننده‌های زانو بود، درحالی‌که پیامدهای بالینی یا علمی این مقادیر به شرایط بستگی دارد، مشاهده شد تفاوت بین ارزیاب‌ها یا در روزهای مختلف کمتر از SDDها بود که نشان می‌دهد اندازه‌گیری‌ها در طول زمان پایدار و ثابت مانده‌اند. این امر، استحکام ابزارهای اصلاح‌شده ما را برجسته می‌کند و پشتیبانی‌کننده استفاده از آن‌ها برای ارزیابی‌های بالینی و علمی قابل‌اعتماد است.

مطالعه حاضر تحت شرایط چالش‌برانگیزی انجام شد که ممکن است بر قابلیت اطمینان تأثیر منفی بگذارد. افراد موردمطالعه در اواسط دوران بلوغ بودند، دوره‌ای که با تغییرات سریع و غیرقابل‌پیش‌بینی در مورفولوژی و عملکرد مشخص می‌شود. علاوه‌براین، فواصل طولانی بین آزمون‌ها (تا ۵ روز) در نظر گرفته شد و از دو ارزیاب نسبتاً بی‌تجربه که تحت یک دوره آموزشی کامل قرار گرفته بودند، استفاده شد. علی‌رغم این چالش‌های بالقوه، یافته‌های ما قابلیت اطمینان بسیار خوبی را نشان داد و استفاده از این آزمون‌های اصلاح‌شده را هم در عمل بالینی و هم در کاربردهای تحقیقاتی تأیید کرد.

آزمون طول خم‌کننده زانو اغلب موردتوجه محققان در سراسر جهان است. به‌طورکلی، نتایج ما با نتایج مطالعات قبلی همسو است؛ بااین‌حال، توانستیم جنبه‌هایی از پایایی آزمون را که قبلاً بررسی نشده‌اند، بررسی کنیم، درحالی‌که بسیاری از مطالعات بر پایایی درون ارزیاب و اندازه‌گیری‌های انجام‌شده در همان روز تمرکز دارند، داده‌ها در مورد پایایی بین ارزیاب و فواصل طولانی‌تر آزمون بازآزمون محدود هستند. تاآنجاکه می‌دانیم، حمید و همکاران [۳۸] تنها نویسندگانی هستند که پایایی بین ارزیاب را برای این آزمون با فاصله یک هفته گزارش کرده‌اند. آن‌ها ICC را ۰/۸۷-۰/۰۸۱ [۱، ۲] ثبت کردند که با نتایج ما (بین ارزیابان در ۵ روز) قابل‌مقایسه است، اگرچه فواصل اطمینان آن‌ها بسیار گسترده‌تر بود و از ۰/۳۲ تا ۰/۹۲ متغیر بود. این اختلاف احتمالاً به‌دلیل حجم نمونه کوچک آن‌ها (۱۴ نفر) و عدم اندازه‌گیری‌های مکرر است (آن‌ها فقط اندازه‌گیری‌های تکی انجام دادند). درنتیجه، می‌توانیم ادعا کنیم که روش ما تخمین‌های دقیق‌تری از ICCهای واقعی برای این اندازه‌گیری ارائه می‌دهد که به‌ویژه در محیط‌های تحقیقات علمی اهمیت دارد. علاوه‌براین، تیم ما قبلاً پایایی بین ارزیاب آزمون خم‌کننده زانو را گزارش

جدول ۳. تخمین‌های پایایی بین ارزیابان

متغیر	پایایی	تعداد اندازه‌گیری‌ها	چپ				راست			
			ضریب همبستگی درون گروهی ICC (±95%CI)	خطای استاندارد اندازه گیری (درجه)	کوچک‌ترین تفاوت‌های قابل شناسایی (درجه)	ضریب همبستگی درون گروه ICC (±95%CI)	خطای استاندارد اندازه گیری (درجه)	کوچک‌ترین تفاوت‌های قابل شناسایی (درجه)
خم‌کننده‌های زانو	درون ارزیاب ۰	۱	۰/۸۸ (۰/۹۳–۰/۸۴)	۴/۱۴	۱۱/۴۸	۰/۸۹ (۰/۹۵–۰/۸۳)	۴/۱۰	۱۱/۳۵
		۲ (میانگین)	۰/۹۳ (۰/۹۶–۰/۹۰)	۳/۱۶	۸/۷۶	۰/۹۲ (۰/۹۵–۰/۸۹)	۳/۴۹	۹/۶۸
		۳ (میانگین)	۰/۹۶ (۰/۹۹–۰/۹۳)	۲/۳۹	۶/۶۲	۰/۹۶ (۰/۹۹–۰/۹۲)	۲/۴۷	۶/۸۵
	درون ارزیاب ۲	۱	۰/۸۵ (۰/۹۰–۰/۷۹)	۴/۶۹	۱۲/۹۸	۰/۸۸ (۰/۹۳–۰/۸۳)	۴/۱۷	۱۱/۵۴
		۲ (میانگین)	۰/۹۱ (۰/۹۶–۰/۸۷)	۳/۶۳	۱۰/۰۵	۰/۹۳ (۰/۹۶–۰/۸۹)	۳/۱۸	۸/۸۱
		۳ (میانگین)	۰/۹۴ (۰/۹۶–۰/۹۱)	۲/۹۶	۸/۲۱	۰/۹۵ (۰/۹۸–۰/۹۲)	۲/۶۹	۷/۴۵
	درون ارزیاب ۵	۱	۰/۷۹ (۰/۸۴–۰/۷۴)	۵/۴۷	۱۵/۱۷	۰/۸۲ (۰/۸۶–۰/۷۷)	۵/۰۴	۱۳/۹۷
		۲ (میانگین)	۰/۸۸ (۰/۹۲–۰/۸۴)	۴/۱۴	۱۱/۴۷	۰/۸۹ (۰/۹۲–۰/۸۶)	۳/۹۴	۱۰/۹۲
		۳ (میانگین)	۰/۹۱ (۰/۹۳–۰/۸۸)	۳/۵۸	۹/۹۳	۰/۹۳ (۰/۹۵–۰/۹۱)	۳/۱۵	۸/۷۱
عضلات بازکننده‌های زانو	درون ارزیاب ۰	۱	۰/۸۷ (۰/۹۱–۰/۸۳)	۱/۹۶	۵/۴۴	۰/۹۰ (۰/۹۴–۰/۸۶)	۱/۶۹	۴/۶۹
		۲ (میانگین)	۰/۹۲ (۰/۹۵–۰/۸۹)	۱/۵۴	۴/۲۷	۰/۹۳ (۰/۹۵–۰/۹۰)	۱/۴۲	۳/۹۳
		۳ (میانگین)	۰/۹۵ (۰/۹۸–۰/۹۳)	۱/۲۲	۳/۳۸	۰/۹۴ (۰/۹۶–۰/۹۲)	۱/۳۱	۳/۶۴
	درون ارزیاب ۲	۱	۰/۸۶ (۰/۹۰–۰/۸۲)	۱/۹۶	۵/۴۲	۰/۸۵ (۰/۸۹–۰/۸۱)	۲/۰۴	۵/۶۶
		۲ (میانگین)	۰/۹۱ (۰/۹۵–۰/۸۸)	۱/۵۹	۴/۴۲	۰/۹۱ (۰/۹۴–۰/۸۹)	۱/۵۸	۴/۳۹
		۳ (میانگین)	۰/۹۳ (۰/۹۵–۰/۹۱)	۱/۳۸	۳/۸۳	۰/۹۴ (۰/۹۶–۰/۹۱)	۱/۲۹	۳/۵۸
	درون ارزیاب ۵	۱	۰/۸۱ (۰/۸۵–۰/۷۷)	۲/۲۳	۶/۱۷	۰/۸۴ (۰/۸۸–۰/۸۱)	۲/۰۶	۵/۷۲
		۲ (میانگین)	۰/۸۵ (۰/۸۸–۰/۸۲)	۱/۹۸	۵/۴۸	۰/۸۹ (۰/۹۱–۰/۸۶)	۱/۷۱	۴/۷۴
		۳ (میانگین)	۰/۹۰ (۰/۹۳–۰/۸۸)	۱/۶۲	۴/۴۸	۰/۹۲ (۰/۹۵–۰/۹۰)	۱/۴۶	۴/۰۴

توانبخشی

کرده است [۲۹]، اگرچه آن اندازه‌گیری‌ها توسط دو ارزیاب در همان روز انجام شده است، اما از آن زمان، ما متدولوژی خود را با معرفی یک گیج تنش و تکنیک اعمال نیرو که برای کاربر "دوست‌دارتر" باشد و همچنین یک روش متفاوت برای تعیین نقطه پایانی حرکت، بهبود بخشیده‌ایم.

علاوه‌براین، از ابزاری ساده برای پشتیبانی از ساق پای آزمودنی‌ها استفاده کردیم که در مقایسه با دستگاه‌های مورداستفاده در سایر مطالعات [۳۳، ۳۸] حتی ساده‌تر به نظر می‌رسد. علاوه‌براین، داده‌های موجود در مقالات اغلب از نمونه‌های کوچک جمع‌آوری می‌شوند [۲۶، ۲۷، ۲۹، ۳۰، ۳۸]. گاهی نویسندگان از روش‌های نامناسبی برای تعیین نقطه پایانی حرکت [۳۳] یا از ابزارهای پیچیده برای کنترل حرکت لگن استفاده می‌کنند [۲۶، ۳۱].

همان‌طور که نتایج ما نشان می‌دهد، به نظر نمی‌رسد چنین اقداماتی برای دستیابی به اندازه‌گیری زاویه‌ای قابل‌اعتماد در آزمون طول خم‌کننده‌های زانو ضروری باشند.

داده‌های موجود در مورد آزمون طول عضله بازکننده زانو تا حدودی محدود است. گاجدوسیک [۲۸] با استفاده از اندازه‌گیری‌های گونیامتری، پایایی بسیار خوبی را نشان داد، اما مطالعه او شامل یک ارزیاب واحد بود که اندازه‌گیری‌ها را در همان روز با تنها ۱۵ آزمودنی انجام می‌داد. تیم ما [۲۹] قبلاً داده‌هایی در مورد پایایی درون و بین ارزیاب برای این آزمون ارائه کرده بود، اما باز هم، همه اندازه‌گیری‌ها در ۱ روز با حجم نمونه کوچک ۱۴ آزمودنی انجام شد. در مطالعه حاضر، این محدودیت‌ها را با استفاده از یک کرنش‌سنج متفاوت و دوباره

با یک روش اصلاح‌شده برای تعیین نقطه پایانی حرکت، برطرف کردیم. همچنین تصمیم گرفتیم از تثبیت خارجی لگن استفاده نکنیم، زیرا دریافتیم که تثبیت ایجادشده توسط چرخش متقابل دو استخوان لگن کافی است **(تصویر شماره ۳)**. تمام داده‌های دیگر مربوط به آزمون طول بازکننده‌های زانو که در این مطالعه ارائه شده است، جدید است.

برای ارائه توصیه‌هایی درمورد استفاده از دو آزمون طول عضله، پیشنهاد می‌کنیم در محیط‌های علمی، انجام ۳ اندازه‌گیری تکراری از آزمون مفید خواهد بود، به‌ویژه زمانی که فاصله زمانی طولانی‌تری بین آزمون‌ها و بازآزمون‌ها وجود دارد (تا ۵ روز در این مطالعه) یا زمانی که آزمون‌ها و بازآزمون‌ها توسط ارزیابان مختلف انجام می‌شود. در شرایطی که فاصله زمانی کوتاه‌تر است و/یا یک ارزیاب واحد وجود دارد، ۲ اندازه‌گیری مکرر ممکن است کافی باشد. کاهش تعداد اندازه‌گیری‌های مکرر به کمتر از ۲ را توصیه نمی‌کنیم، درحالی‌که این رویکرد ممکن است در یک محیط بالینی قابل‌قبول باشد (با فرض اینکه پزشک آموزش مناسب دیده و به‌طور دقیق به پروسه پایبند باشد)، جایی که الزامات قابلیت اطمینان کمی انعطاف‌پذیرتر است، باتوجه‌به اینکه برای یک اندازه‌گیری مکرر، پایین‌ترین ICC که ثبت کردیم ۰/۷۹ بود (برای آزمون طول خم‌کننده‌های زانو، قابلیت اطمینان بین ارزیاب با فاصله ۵ روز)، همچنان نتیجه‌ای رضایت‌بخش برای استفاده بالینی است. همچنین فراتر از یک نقطه خاص، افزایش تعداد اندازه‌گیری‌های مکرر به افزایش بیشتر مقدار ICC منجر نمی‌شود، اما به محدود کردن فاصله اطمینان ICC کمک می‌کند. این پدیده از دیدگاه علمی سودمند است، زیرا دقت تخمین‌های قابلیت اطمینان را افزایش می‌دهد.

نتیجه‌گیری

پس از اصلاحات خاص مناسب، امکان دستیابی به سطح بسیار خوب و عالی از پایایی آزمون‌های طول خم‌کننده‌ها و بازکننده‌های زانو وجود دارد. با استفاده از ۳ اندازه‌گیری مکرر، تمام ICCهای ثبت‌شده بالاتر از ۰/۹۰ بودند. چنین سطح پایایی بالایی، استفاده از نسخه‌های اصلاح‌شده آزمون‌ها را در محیط‌های علمی توجیه می‌کند. برای دستیابی به این هدف، نیازی به تجهیزات پیشرفته یا هزینه‌های زمانی، شخصی و اقتصادی بالا نیست.

محدودیت‌ها

یکی از محدودیت‌های مطالعه ما، اعتبار خارجی محدود آن است، زیرا نمونه فقط شامل نوجوانان بود. آزمون‌های خم‌کننده‌ها و بازکننده‌های زانو اغلب در ارزیابی‌های وضعیتی در این مرحله رشدی استفاده می‌شوند، اما هنگام تلاش برای تعمیم این یافته‌ها به سایر گروه‌های سنی، احتیاط لازم است. باوجوداین، اصلاحاتی که به کار بردیم مزایایی ازجمله هزینه کم، حداقل تجهیزات (گونیومتر دیجیتال و نیروسنج) و کاهش نیاز به پرسنل داشت که قابل‌دسترس در محیط‌های مختلف است. علاوه‌براین، نتایج ما نشان می‌دهد که آزمایش‌های اصلاح‌شده را می‌توان بدون نیاز به ابزارهای پیچیده یا گران‌قیمت مانند اشعه ایکس یا سیستم‌های تحلیل حرکت اجرا کرد. درنهایت، تصمیم گرفتیم ICCها را برای مدل k,۲ ارائه دهیم که امکان تعمیم وسیع‌تری را در بین ارزیابان مختلف فراهم می‌کند، هرچند که در مدل k,۳ احتمالاً قابلیت اطمینان بالاتری را نشان می‌دهد. قابلیت اطمینان درون ارزیاب را برای اندازه‌گیری‌های همان روز ارائه ندادیم، زیرا این موضوع به‌طور ذاتی در قابلیت اطمینان‌های بالاتر منعکس شده است، جایی که به‌دلیل کاهش واریانس، مقادیر ICC بالاتری را مشاهده می‌کنیم.

نتایج امیدوارکننده‌ی این مطالعه، راه‌های متعددی را برای تحقیقات آینده در مورد قابلیت اطمینان و کاربرد آزمون‌های اصلاح‌شده طول عضله بالینی، به‌ویژه برای استفاده در عمل بالینی و تحقیقات علمی، باز می‌کند. نقاط کلیدی تحقیقات آینده ممکن است شامل موارد زیر باشد: دامنه سنی وسیع‌تری از شرکت‌کنندگان و جمعیت‌های مختلف (مانند ورزشکاران جوان)، شرایط بالینی متنوع (مانند فلج مغزی، دیستروفی‌های عضلانی و غیره)، دوره‌های پیگیری طولانی‌تر، بررسی‌های مربوط به اعتبار آزمون‌ها و درنهایت، گنجاندن آن‌ها در آزمایش‌ها به عنوان ابزارهای اندازه‌گیری.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

این مطالعه در کمیته اخلاق تحقیقات زیست پزشکی نهادی، با شماره تصویب ۲۰۲۰/۱۸، تأیید شده است.

حامی مالی

این پژوهش هیچ‌گونه کمک مالی از سازمانی‌های دولتی، خصوصی و غیرانتفاعی دریافت نکرده است.

مشارکت نویسندگان

مفهوم‌سازی، منابع، نگارش-تهیه و نگارش پیش‌نویس اصلی- بررسی و ویرایش: همه نویسندگان؛ روش‌شناسی و نظارت: رافال گنات، آنا گوگولا و توماش ولنی؛ نرم‌افزار: رافال گنات؛ اعتبارسنجی، تحلیل رسمی و مدیریت پروژه: رافال گنات و آنا گوگولا؛ تحقیق: آنا گوگولا؛ سرپرستی داده‌ها: آگنیشکا پولاچک و پیوتر ووزنیاک؛ تجسم: رافال گنات، آگنیشکا پولاچک و پیوتر وونیاک.

تعارض منافع

بنابر اظهار نویسندگان، این مقاله تعارض منافع ندارد.

References

[1] Chmielewska D, Cebula M, Gnat R, Rudek-Zeprzałka M, Gruszczyńska K, Baron J, et al. Reliability of inter-recti distance measurement on ultrasound images captured by novice examiners. Physiotherapy Theory and Practice. 2024; 40(11):2652-60. [DOI:10.1080/09593985.2023.2255897] [PMID]

[2] Gnat R, Saulicz E, Miądowicz B. Reliability of real-time ultrasound measurement of transversus abdominis thickness in healthy trained subjects. European Spine Journal. 2012; 21(8):1508-15. [DOI:10.1007/s00586-012-2184-4] [PMID]

[3] Biały M, Adamczyk W, Gnat R, Stranc T. Tissue deformation index as a reliable measure of lateral abdominal muscle activation on M-mode sonography. Journal of Ultrasound in Medicine. 2017; 36(7):1461-7. [DOI:10.7863/ultra.16.07045] [PMID]

[4] Biały M, Adamczyk WM, Marczykowski P, Majchrzak R, Gnat R. Deformations of abdominal muscles under experimentally induced low back pain. European Spine Journal . 2019; 28(11):2444-51. [DOI:10.1007/s00586-019-06016-y] [PMID]

[5] Gogola A, Gnat R, Dziub D, Gwóźdź M, Zaborowska M. The impact of the neurodevelopmental traction technique on activation of lateral abdominal muscles in children aged 11-13 years. NeuroRehabilitation. 2016; 39(2):183-90. [DOI:10.3233/NRE-161350] [PMID]

[6] Gogola A, Gnat R, Zaborowska M, Dziub D, Gwóźdź M. Asymmetry of activation of lateral abdominal muscles during the neurodevelopmental traction technique. Journal of Bodywork and Movement Therapies. 2018; 22(1):46-51. [DOI:10.1016/j.jbmt.2017.03.019] [PMID]

[7] Park SD. Reliability of ultrasound imaging of the transversus deep abdominial, internal oblique and external oblique muscles of patients with low back pain performing the drawing-in Maneuver. Journal of Physical Therapy Science. 2013; 25(7):845-7. [DOI:10.1589/jpts.25.845] [PMID]

[8] Guruhan S, Kafa N, Ecemis ZB, Guzel NA. Muscle activation differences during eccentric hamstring exercises. Sports Health. 2021; 13(2):181-6. [DOI:10.1177/1941738120938649] [PMID]

[9] Oliva-Lozano JM, Muyor JM. Core muscle activity during physical fitness exercises: A systematic review. International Journal of Environmental Research and Public Health. 2020; 17(12):4306. [DOI:10.3390/ijerph17124306] [PMID]

[10] Gnat R, Biały M, Dziewońska A. Experimentally induced low back pain influences brain networks activity. Journal of Motor Behavior. 2021; 53(6):680-92. [DOI:10.1080/00222895.2020.1839376] [PMID]

[11] Gnat R, Dziewońska A, Biały M, Wieczorek M. Differences in activity of the brain networks during voluntary motor tasks engaging the local and global muscular systems of the lower trunk. Motor Control. 2020; 24(4):624-43. [DOI:10.1123/mc.2019-0109] [PMID]

[12] Biały M, Kłaptocz P, Gnat R. Functional asymmetry of the spine in standing and sitting positions. Acta Universitatis Palackianae Olomucensis. 2010; 40(1):53-60. [Link]

[13] Gnat R, Biały M. A new approach to the measurement of pelvic asymmetry: Proposed methods and reliability. Journal of Manipulative and Physiological Therapeutics. 2015; 38(4):295-301. [DOI:10.1016/j.jmpt.2015.02.002] [PMID]

[14] Gnat R, Spoor K, Pool-Goudzwaard A. The influence of simulated transversus abdominis muscle force on sacroiliac joint flexibility during asymmetric moment application to the pelvis. Clinical Biomechanics. 2015; 30(8):827-31. [DOI:10.1016/j.clinbiomech.2015.06.006] [PMID]

[15] Hecht GG, Van Rysselberghe NL, Young JL, Gardner MJ. Gait analysis in orthopaedic surgery: History, limitations, and future directions. Journal of the American Academy of Orthopaedic 2022; 30(21):e1366-73. [DOI:10.5435/JAAOS-D-21-00785]

[16] Qiu Y, Guan Y, Liu S. The analysis of infrared high-speed motion capture system on motion aesthetics of aerobics athletes under biomechanics analysis. Plos One. 2023; 18(5):e0286313. [DOI:10.1371/journal.pone.0286313] [PMID]

[17] Kurashina W, Iijima Y, Sasanuma H, Saito T, Takeshita K. Evaluation of muscle stiffness in adhesive capsulitis with Myoton PRO. JSES International. 2022; 7(1):25-9. [DOI:10.1016/j.jseint.2022.08.017] [PMID]

[18] Kurashina W, Takahashi T, Sasanuma H, Saitsu A, Takeshita K. Relationship between achilles tendon stiffness using myoton PRO and translation using a tensile testing machine: A biomechanical study of a porcine model. Cureus. 2023; 15(11):e49359. [DOI:10.7759/cureus.49359] [PMID]

[19] Linek P, Pałac M, Wolny T. Shear wave elastography of the lateral abdominal muscles in C-shaped idiopathic scoliosis: A case-control study. Scientific Reports. 2021; 11(1):6026. [DOI:10.1038/s41598-021-85552-4] [PMID]

[20] Ormachea J, Parker KJ. Elastography imaging: The 30 year perspective. Physics in Medicine and Biology. 2020; 65(24):24TR06. [DOI:10.1088/1361-6560/abca00] [PMID]

[21] Dębski P, Białas E, Gnat R. The parameters of foam rolling, self-myofascial release treatment: A review of the literature. Biomedical Human Kinetics. 2019; 11(1):36-46. [DOI:10.2478/bhk-2019-0005]

[22] Neto T, Jacobsohn L, Carita AI, Oliveira R. Reliability of the active-knee-extension and straight-leg-raise tests in subjects with flexibility deficits. Journal of Sport Rehabilitation. 2015; 24(4):2014-0220. [DOI:10.1123/jsr.2014-0220] [PMID]

[23] Cejudo A, Sainz de Baranda P, Ayala F, Santonja F. Test-retest reliability of seven common clinical tests for assessing lower extremity muscle flexibility in futsal and handball players. Physical Therapy in Sport. 2015; 16(2):107-13. [DOI:10.1016/j.ptsp.2014.05.004] [PMID]

[24] Peeler JD, Anderson JE. Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. Journal of Athletic Training. 2008; 43(5):470-6. [DOI:10.4085/1062-6050-43.5.470] [PMID]

[25] Silvers-Granelli HJ, Cohen M, Espregueira-Mendes J, Mandelbaum B. Hamstring muscle injury in the athlete: State of the art. Journal of ISAKOS. 2021; 6(3):170-81. [DOI:10.1136/jisakos-2017-000145] [PMID]

[26] Fredriksen H, Dagfinrud H, Jacobsen V, Maehlum S. Passive knee extension test to measure hamstring muscle tightness. Scandinavian Journal of Medicine & Science in Sports. 1997; 7(5):279-82. [DOI:10.1111/j.1600-0838.1997.tb00153.x] [PMID]

[27] Gajdosik R, Lusin G. Hamstring muscle tightness. Reliability of an active-knee-extension test. Physical Therapy. 1983; 63(7):1085-90. [DOI:10.1093/ptj/63.7.1085] [PMID]

[28] Gajdosik RL. Rectus Femoris Muscle Tightness: Lntratester Reliability of an Active Knee Flexion Test. The Journal of Orthopaedic and Sports Physical Therapy. 1985; 6(5):289-92. [DOI:10.2519/jospt.1985.6.5.289] [PMID]

[29] Gnat R, Kuszewski M, Koczar R, Dziewońska A. Reliability of the passive knee flexion and extension tests in healthy subjects. Journal of Manipulative and Physiological Therapeutics. 2010; 33(9):659-65. [DOI:10.1016/j.jmpt.2010.09.001] [PMID]

[30] Guex K, Fourchet F, Loepelt H, Millet GP. Passive knee-extension test to measure hamstring tightness: Influence of gravity correction. Journal of Sport Rehabilitation. 2012; 21(3):231-4. [DOI:10.1123/jsr.21.3.231] [PMID]

[31] Hamberg J, Björklund M, Nordgren B, Sahlstedt B. Stretchability of the rectus femoris muscle: investigation of validity and intratester reliability of two methods including X-ray analysis of pelvic tilt. Archives of Physical Medicine and Rehabilitation. 1993; 74(3):263-70. [Link]

[32] Norris CM, Matthews M. Inter-tester reliability of a self-monitored active knee extension test. Journal of Bodywork and Movement Therapies. 2005; 9(4):256-9. [DOI:10.1016/j.jbmt.2005.06.002]

[33] Rakos DM, Shaw KA, Fedor RL, Lamanna M, Yocum CC, Lawrence KJ. Interrater reliability of the active-knee-extension test for hamstring length in school-aged children. Pediatric Physical Therapy. 2001; 13(1):37-41. [DOI:10.1097/00001577-200104000-00006] [PMID]

[34] Reurink G, Goudswaard GJ, Oomen HG, Moen MH, Tol JL, Verhaar JA, et al. Reliability of the active and passive knee extension test in acute hamstring injuries. The American Journal of Sports Medicine. 2013; 41(8):1757-61. [DOI:10.1177/0363546513490650] [PMID]

[35] Shimon JM, Darden GF, Martinez R, Clouse-Snell J. Initial reliability and validity of the lift-and-raise hamstring test. Journal of Strength and Conditioning Research. 2010; 24(2):517-21. [DOI:10.1519/JSC.0b013e3181ca32ae] [PMID]

[36] Arfin WN. Sample size calculator (Version 2.0) [Internet]. 2024 [Updated 2024 November 1]. Available from: [Link]

[37] Portney L, Watkins M. Foundations of clinical research: Applications to practice. 3rd ed. Harlow, Essex: Pearson/Prentice Hall; 2009. [Link]

[38] Hamid MS, Ali MR, Yusof A. Interrater and intrarater reliability of the active knee extension (AKE) test among healthy adults. Journal of Physical Therapy Science. 2013; 25(8):957-61. [DOI:10.1589/jpts.25.957] [PMID]

This Page Intentionally Left Blank