CPQ Orthopaedics (2023) 6:4
Cross Sectional Reliability Study

Interrater and Intrarater Reliability for Measuring the Length of the Flexor Digitorum Superficialis

John Connor Montemayor*, David Iodice, Michael Esposito, Lucas Ferreira & Jason Grimes

Department of Physical Therapy and Human Movement Science, Stanford Health Care, USA

***Correspondence to:** Dr. John Connor Montemayor, Bachelor of Medicine and Bachelor of Surgery, Xuzhou Medical University, China.

Received: 03 May 2023

Published: 06 June 2023

Keywords: *Flexor Pronator Mass; Muscle Length; Forearm Flexor; Goniometer*

Abstract

Introduction

Measuring Flexor Digitorum Superficialis (FDS) length can provide information regarding functional gripping, forearm muscle power, and possibly risk of injury. To our knowledge, no study has previously investigated measuring FDS length.

Methods

FDS length was assessed by two examiners in 21 healthy individuals (total of 42 wrists) through measuring wrist extension range of motion (ROM) in two different positions. Both positions involved elbow extension, forearm pronation, and wrist extension. Position 1 included finger flexion while Position 2 included finger extension. Test outcome was defined as a positive test being less ROM in Position 2 compared to Position 1.

Results

Intrarater reliability for Position 1 produced an ICC3,1 of 0.85 and 0.91 for Examiner 1 and 2, respectively. For Position 2, the ICC3,1 was 0.85 and 0.94 for Examiner 1 and 2, respectively. For Position 1, the standard error of measure (SEM) was 2.2 and 2.1 degrees, and the minimal detectable change (MDC) was 6.2 and 5.8 degrees for Examiner 1 and 2, respectively. For Position 2, SEM was 2.2 and 1.7 degrees, and the MDC was 6.1 and 4.6 degrees for Examiner 1 and 2, respectively. Interrater reliability produced an ICC2,3 for Position 1 of 0.74 and Position 2 of 0.81. The percent agreement for test outcome (positive vs. negative) was 69% with a kappa value of 0.26.

Conclusion

The findings suggest good to excellent intrarater reliability and moderate to good interrater reliability for this proposed measure of FDS length.

Introduction

The flexor digitorum superficialis (FDS) muscle has been a topic of interest in recent studies [1-3]. The FDS is located in the anteromedial forearm and is the largest of the extrinsic wrist flexors [4]. The primary action of the FDS is gripping, which includes flexion of the wrist and flexion of the proximal interphalangeal (PIP) joints. As hand grip strength increases, FDS electromyographic (EMG) activity increases [5]. A secondary function of the FDS is working with the ulnar collateral ligament (UCL) to provide the greatest dynamic stabilization to valgus force at the elbow [1,6].

The FDS is a multi-joint fusiform muscle and due to the muscle architecture, the muscle fails resulting in a degree of muscle strain, with minimal elongation from its resting length [7,8]. The length of the FDS has a direct relationship with the muscle's capacity to generate force [9]. A shortened FDS muscle length of 5% induces a decrease in muscle force capacity by 50% [9]. Both shortened and lengthened muscles have implications that can contribute to muscle injury [10]. Other multi-joint muscles, such as the hamstring, have been extensively studied and demonstrate an increase in muscular strain associated with decreased muscle length [11]. Muscle strains, being one of the most common non-contact injuries, are seen across athletes and require immediate attention to prevent more serious complications [12]. Predisposing factors for strains are eccentrically loading a muscle, muscles that are composed primarily of type 2 fibers, and muscles that cross two or more joints. Askling et al. (2006) stated that there is no evidence for the exact point in the range of motion (ROM) that muscle strains are more likely to occur, but muscle strains are more related to the rate of force production and muscle length changes [13]. Therefore, the ability to reliably measure the length of the FDS may give clinicians an idea of whether the FDS is contributing to forearm muscle strains or upper extremity pain and dysfunction.

An assessment of FDS length is also important for clinicians who are treating patients with spinal cord injuries. Tenodesis grip is used to supplement hand function with certain patients within the spinal cord injury population [14,15]. Tenodesis grip requires shortening of the forearm flexor muscles, including the FDS [16-18]. Passively measuring the FDS can allow clinicians to objectify the shortening of the muscle.

To indirectly assess FDS length, a measure of passive wrist extension ROM can be obtained with the FDS maximally elongated with the fingers extended and repeated with the FDS on slack with the fingers flexed. The difference between the wrist extension ROM with the FDS maximally lengthened and the FDS on slack can then be calculated to provide an indication of FDS length. The measurement of wrist extension ROM, which is routinely used in clinical practice, can be obtained using a standard 8-inch goniometer. To our knowledge, no study has previously investigated interrater and intrarater reliability of a measurement of FDS length. The purpose of this study was therefore to examine the intrarater and interrater reliability of this proposed assessment of FDS length using a goniometric measure of wrist extension in a healthy population. Examining this measure in a healthy population is an important first step before investigating it in a specific population of interest such as individuals with medial elbow pain, spinal cord injury, or overhead throwers.

Methods

This study was a cross sectional reliability study where two examiners performed the measures of FDS length on all participants for multiple trials in order to examine the intrarater and interrater reliability of the measure. One examiner was a student Physical Therapist and the other was an expert Physical Therapist with 17 years of clinical experience. Each examiner was assigned a separate recorder familiar with goniometry. The job of each recorder was to determine the measurement value from the goniometer and record the value, ensuring that each examiner remained blinded to the measurement values. Each examiner assessed the passive ROM of wrist extension bilaterally in both positions (Position 1=fingers flexed, Position 2=fingers extended) using a standard 8-inch goniometer.

Participants were recruited at a single university by means of email or verbal communication. Eligibility requirements were for healthy individuals over 18 years old without current upper extremity pain. A past medical history of upper extremity injury was allowed as long as the individual was asymptomatic at the time of testing. Randomization was used for determining the order of examiner seen, which arm was to be measured first, and starting test position (Position 1 or Position 2). Each participant first pulled a number out of a hat to indicate which examiner would be seen first. A second paper was then picked again by the participant to indicate the starting arm (left or right) and test position (Position 1 or Position 2). Measurements of passive ROM for wrist extension were then obtained with the upper extremity in both positions: Position 1 with elbow extended, forearm pronated, wrist extended, and fingers relaxed in a flexed position; and Position 2 with elbow extended, forearm pronated, wrist extended, and fingers extended or hyperextended as able. Each ROM with the respected Position (1 or 2) was taken to maximal wrist extension end range.

The technique for measuring the passive ROM was very specific and was considered to be especially important due to experience obtained during pilot testing. The participant and examiner were seated at the treatment table and the participant's test arm rested comfortably on the treatment table. The technique used can be seen in Figures 1 and 2. The examiner's proximal forearm was positioned to lock the participant's elbow out into extension and pronation, and the web space of the examiner's proximal hand wrapped around the participant's second proximal phalanx. The fingers of the examiner's proximal arm were on the volar surface of the participant's hand, while the examiner's thumb was on the dorsal side of the participant's hand. For Position 1, the examiner's hand only stabilized the participant's hand along the volar surface of the metacarpals (Figure 1). For Position 2, the examiner's hand stabilized the participant's hand across the volar surface of the proximal and middle phalanges, and special attention was placed on having the examiner's thumb on the dorsal surface of the participant's second through fifth PIP joints to ensure PIP joint flexion did not occur when measuring passive wrist extension (Figure 2). To ensure that only the FDS was being put on tension and not flexor digitorum profundus, the participant's distal interphalangeal joint was relaxed in a flexed position. The distal arm of the examiner was used to obtain the goniometric measure of passive wrist extension.

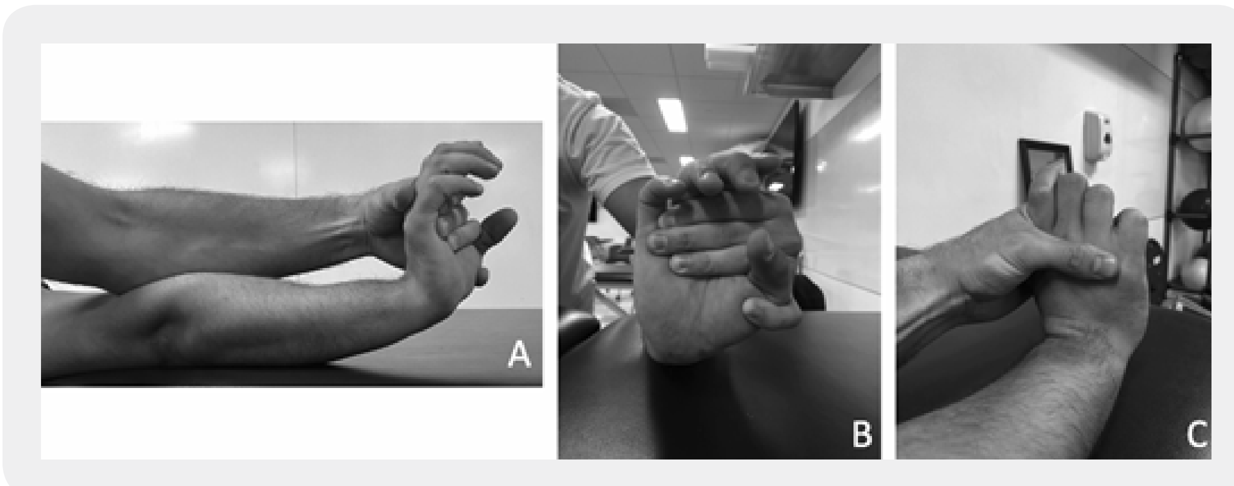

Figure 1: *Position 1 technique. With the participant and the examiner sitting, wrist extension ROM was measured with elbow extended, forearm pronated, wrist extended, fingers flexed. A) Emphasis on forearm lock between examiner and participant. B) Emphasis on contact between all metacarpal bones. C) Emphasis on allowing the proximal interphalangeal joints to relax.*

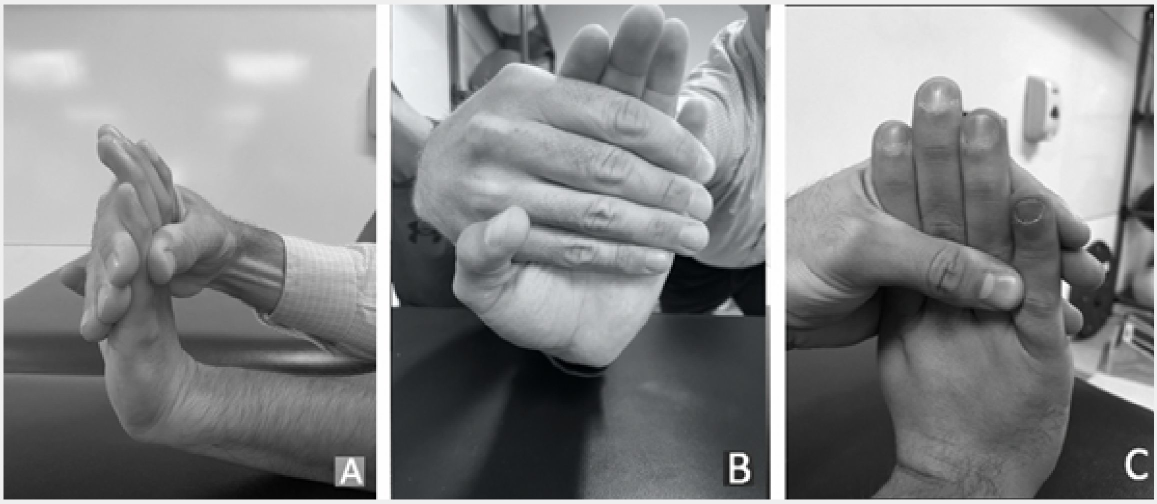

Figure 2: *Position 2 technique. With the participant and the examiner sitting, wrist extension ROM was measured with elbow extended, forearm pronated, wrist extended, fingers extended. A) Emphasis on forearm lock between examiner and participant. B) Emphasis on contact between all metacarpal bones and proximal interphalangeal joints being held into extension. C) Emphasis on examiner's thumb on the dorsal aspect of the proximal interphalangeal joints, locking fingers into extension`*

The anatomical landmarks used for the goniometric measure of wrist extension were as follows: the fulcrum was placed on the ulnar aspect of the wrist over the triquetrum, the proximal arm was placed on the midline of the ulna using the olecranon and ulnar styloid process as reference points, and the distal arm was placed in line with the 5th metacarpal being careful to disregard the soft tissue mass of the hypothenar eminence. Once the alignment of the goniometer was completed by the examiner, the goniometer was read, and the value was recorded by the recorder to maintain blinding of the examiner. The measurements for the two testing positions were taken for both the right and left wrists for three trials. There was a rest period of 30 seconds for measuring each wrist, during which time the contralateral wrist was measured. The participant was given a 3-minute rest and the procedure was then repeated by the other examiner. The order of testing was again randomized by pulling a paper from a hat. The study protocol was approved by an Institutional Review Board to meet ethical standards.

A sample size estimation was run where a minimal accepted reliability (ICC) of 0.6, an expected reliability (ICC) of 0.8, a statistical power of 0.85, repetitions per subject of 3, and a significance level of 0.05 would require a sample size of 38 wrists [19]. Wrists rather than participants were the unit of analysis, presuming that each wrist is independent to account for discrepancies with dominant arm and non-dominant arm. SPSS version 26 was used for statistical analysis. Independent samples t-test was used to investigate for differences in motion between Position 1 and Position 2. A two-way mixed, absolute agreement, single measure intraclass correlation coefficient ($ICC_{3,1}$) was used to assess intrarater reliability for Examiner 1 and Examiner 2 for both Position 1 and Position 2. The standard error of the measure (SEM) and minimal detectable change (MDC) were also calculated at the 90% confidence level for each examiner using the following formulas: SEM_{90}=1.64 x SD x √(1-ICC) and MDC_{90}=SEM_{90} x √2. A two-way random, absolute agreement, average measures $ICC_{(2,3)}$ was used to determine interrater reliability between the two examiners for both Position 1 and Position 2. Cohen's kappa and positive test outcome were calculated to investigate agreement between the two examiners on whether Position 2 had less ROM than Position 1. Positive test outcome was defined as less ROM in Position 2 vs. Position 1. The 95% limits of agreement (LOA) were calculated between examiners for Position 1 and Position 2 measurements and Bland-Altman plots were produced (Figure 3).

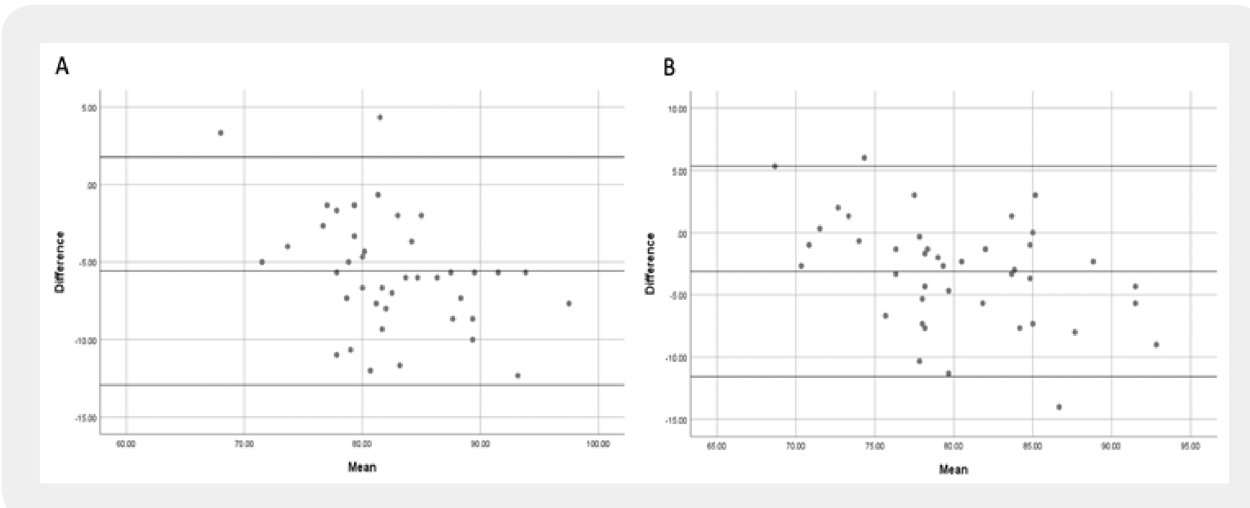

Figure 3: *Bland-Altman plot for Position 1 and Position 2. Difference (degrees) and mean (degrees) between Examiner 1 and Examiner 2. A) Position 1. B) Position 2.*

Results

Twenty-one (6 men and 15 women) healthy participants with a mean age of 23.5 years (SD=1.6 years) had their right and left wrist extension ROM measured in both Position 1 and Position 2. From the 21 participants, 81% were right hand dominant and 57.1% were current or previous overhead athletes. Each wrist was assessed independently, resulting in 42 total wrists that were measured to examine the length of the FDS. For Examiner 1, there was no significant difference between Position 1 and Position 2 for both the right (t=1.48, p=0.15) and left (t=-0.02, p=0.98) wrists. For Examiner 2, there was no significant difference between Position 1 and Position 2 for both the right (t=1.29, p=0.20), and left (t=2.0, p=0.052) wrists. (Table 1 & 2)

Intrarater reliability for Position 1 and Position 2 for both examiners was good to excellent. For Position 1, Examiner 1 had an $ICC_{(3,1)}$ value of 0.85 (95% confidence interval (CI): 0.76-0.91) and Examiner 2 had an ICC(3,1) value of 0.91 (95% CI: 0.85-0.95). For position 2, Examiner 1 had an $ICC_{(3,1)}$ value of 0.85 (95% CI: 0.75-0.91) and Examiner 2 had an $ICC_{(3,1)}$ value of 0.94 (95% CI: 0.91-0.97). For Position 1, the SEM was 2.2 and 2.1 degrees, and the MDC was 6.2 and 5.8 degrees for Examiner 1 and 2, respectively. For Position 2, the SEM was 2.2 and 1.7 degrees, and the MDC was 6.1 and 4.6 degrees for Examiner 1 and 2, respectively.

Interrater reliability for Position 1 was moderate with an $ICC_{(2,3)}$ value of 0.74 (95% CI: -0.19-0.92). The interrater reliability for Position 2 was good with an $ICC_{(2,3)}$ value of 0.81 (95% CI: 0.46-0.92). The 95% LOA was -12.93-1.77 for Position 1 and -11.58-5.34 for Position 2. Percent agreement was 69% for the two examiners on the final muscle length outcome (positive or negative result for FDS tightness) and the kappa value was 0.26 (95% CI: -0.01-0.54), representing a fair level of agreement.

Table 1: *Participant Descriptive Information.*

	Minimum	Maximum	Mean	Std Deviation
Age (years)	21	27	23.5	1.7
Height (inches)	59.0	71.0	65.4	3.6
Weight (pounds)	100.0	220.0	144.5	32.3
BMI (kg/m2)	18.9	31.6	23.5	3.1
Hand Dominance				
	Frequency		Percentage	
Right	17		81.0	
Left	4		19.0	
Previous or Current Overhead Athlete				
	Frequency		Percentage	
Yes	12		57.1	
No	9		42.9	

Table 2: *Measures of Wrist Extension ROM for Position 1 and 2 for each Examiner*

	N	Minimum	Maximum	Mean	Std Deviation
Examiner 1 Position 1	42	69°	94°	79.7°	5.4°
Examiner 2 Position 1	42	66°	101°	85.3°	6.9°
Examiner 1 Position 2	42	69°	89°	78.6°	5.3°
Examiner 2 Position 2	42	66°	97°	81.7°	7.1°

***Table 3:** Summary of Reliability Values*

Intrarater Reliability				
		SEM	MDC	ICC (95% CI)
Examiner 1	Position 1	2.2°	6.2°	0.85 (0.76-0.91)
	Position 2	2.1°	5.8°	0.91 (0.85-0.95)
Examiner 2	Position 1	2.2°	6.1°	0.85 (0.75-0.91)
	Position 2	1.7°	4.6°	0.94 (0.91-0.97)
Interrater Reliability				
		Position 1		Position 2
ICC (95% CI)		0.74 (-0.19-0.92)		0.81 (0.46-0.92)
95% Limits of Agreement		-12.93-1.77		-11.58-5.34
Percent Agreement		69%		
Cohen's Kappa (95% CI)		0.26 (-0.01-0.54)		

SEM=Standard Error of the Measure; MDC=Minimal Detectable Change; ICC=Intraclass Correlation Coefficient; CI=confidence interval

Discussion

To our knowledge, there is no evidence of the reliability of a measurement assessing the length of the FDS. However, it has been shown that length of the FDS correlates to force production [9]. To further investigate how the length of the FDS impacts force production and energy dissipation, a reliable clinical measure is needed. The proposed measure for assessing the length of the FDS involves measuring passive ROM of wrist extension with the FDS maximally elongated and then slackened. As described earlier, previous pilot work on this measurement indicated that discrepancies in hand placement by the examiners during the measurement of Position 2 influenced the results. The measurement was improved by making sure that the examiner's thumb was on the dorsal surface of the participant's second through fifth proximal interphalangeal joints to ensure no flexion occurred at that joint when taking the measurement. This level of detail and attention to proper hand placement is important when performing the measurement for Position 2.

The intrarater reliability for Position 1 and Position 2 was good for Examiner 1 and excellent for Examiner 2. The reliability for Position 2 was stronger and had a narrower confidence interval than Position 1, possibly due to the associated end feel in that test position. Position 2 puts the FDS in a maximally lengthened position, which promotes a soft tissue stretch end feel. Differences in end feel between the two test positions likely contributed to the difference in the SEM and the MDC between Position 1 and Position 2. The reliability of Position 1 and Position 2 obtained from this study is similar to reliability values commonly reported for other goniometric measurements [20]. Goniometric measurements have been found to have good to excellent intratester reliability at the wrist joint, with most ICC values ranging between 0.80 and 0.95 [20]. The SEM ranges from 3 to 5 degrees for most goniometric measurements [20]. While this is the first study to investigate the reliability of a measure for FDS length, a comparison can be made to a previous study on the knee extension angle measure for hamstring length which reported an intra-tester ICC of 0.94 [21].

The interrater reliability was moderate for both Position 1 and Position 2. For Position 1, we believe the large confidence interval obtained was most likely due to variability in the amount of force that was applied when putting the wrist at its end range of extension. Examiner 2 measured more ROM for both Position 1 and Position 2 compared to Examiner 1 (Table 2), likely as a result of utilizing a greater force at end range. Attempting to standardize the force used would likely help to provide a more consistent measurement between examiners. Additionally, the sample used for this study might have affected the reliability of the measurement. The measurement may have been less reliable and more variable in this sample due to the likelihood that the healthy population tested did not possess significant FDS stiffness. Performing the test on a population that is more prone to FDS stiffness may result in improved measurement accuracy.

Differences in measurements between Position 1 and Position 2 can be interpreted as an indication of the length of the FDS. For both examiners, there was no significant difference on either the right or left sides between Position 1 and Position 2. With a population likely presenting with FDS stiffness or shortness, significant differences between Position 1 and 2 would be expected and will be a future area of investigation. The proposed method to measure the length of the FDS is simple, inexpensive, and feasible with no extra training and no extra personnel to assist in the measurement. It should be considered relevant because of its importance in terms of force production, muscle injury, and tenodesis grip.

This study has limitations. Due to the ease of access and various restrictions in place due to the COVID-19 pandemic, we utilized a convenience sample of graduate students at a single university. The exclusion of different age ranges across the life span limits the generalizability of the findings and should be investigated in future studies. Another limitation was an inability to perform test-retest reliability of this measurement due to restrictions encountered with trying to bring participants back for a second assessment period. A final limitation of our study was not accounting for the amount of force used when passively moving the participant into maximal end-range wrist extension. Future studies should include standardization and training to ensure both examiners are consistently using the same amount of force on each individual.

Bibliography

1. Udall, J., Fitzpatrick, M., McGarry, M., *et al.* (2009). Effects of flexor-pronator muscle loading on valgus stability of the elbow with an intact, stretched, and resected medial ulnar collateral ligament. *Journal of Shoulder and Elbow Surgery, 18*(5), 773-778.

2. Buffi, J., Werner, K., Kepple, T., *et al.* (2015). Computing muscle, ligament, and osseous contributions to the elbow varus moment during baseball pitching. *Ann Biomed Eng., 43*(2), 404-415.

3. Barco, R. & Antuña, S. (2017). Medial elbow pain. *EFORT Open Reviews., 2*(8), 362-371.

4. Okafor, L. & Varacallo, M. (2020). Anatomy, shoulder and upper limb, hand flexor digitorum superficialis muscle. *StatPearls*.

5. Mohideen, A. & Sidek, S. (2011). Development of EMG circuit to study the relationship between flexor digitorum superficialis muscle activity and hand grip strength. In: 2011 4th International Conference on Mechatronics (ICOM)., 1-7.

6. Hoshika, S., Nimura, A., Takahashi, N., *et al.* (2020). Valgus stability is enhanced by flexor digitorum superficialis muscle contraction of the index and middle fingers. *J Orthop Surg Res., 15*(1), 121.

7. Shoja, M., Tubbs, R., Loukas, M., *et al.* (2008). The split flexor digitorum superficialis. *Ital J Anat Embryol., 113*(2), 103-107.
6. Chitgopkar, S. D. (2008). Flexible nailing of fractures in children using stainless steel Kirschner wires. *J Pediatr Orthop B., 17*(5), 251-255.

7. Garg, S., Dobbs, M. B., Schoenecker, P. L., Luhmann, S. J. & Gordon, J. E. (2009). Surgical treatment of traumatic pediatric humeral diaphyseal fractures with titanium elastic nails. *J Child Orthop., 3*(2), 121-127.

8. Garrett, W., Nikolaou, P., Ribbeck, B., *et al.* (1988). The effect of muscle architecture on the biomechanical failure properties of skeletal muscle under passive extension. *Am J Sports Med., 16*(1), 7-12.

9. Hauraix, H., De Monsabert, B., Herbaut, A., et al. (2018). Force-length relationship modeling of wrist and finger flexor muscles. *Medicine and Science in Sports and Exercise, 50*(11), 2311-2321.

10. Butterfield, T. & Herzog, W. (2006). Effect of altering starting length and activation timing of muscle on fiber strain and muscle damage. *Journal of Applied Physiology, 100*(5), 1489-1498.

11. Liu, H., Garrett, W., Moorman, C., *et al.* (2012). Injury rate, mechanism, and risk factors of hamstring strain injuries in sports: A review of the literature. *Journal of Sport and Health Science, 1*(2), 92-101.

12. Iwata, M., Yamamoto, A., Matsuo, S., *et al.* (2019). Dynamic stretching has sustained effects on range of motion and passive stiffness of the hamstring muscles. *J Sports Sci Med., 18*(1), 13-20.

13. Askling, C. (2006). Type of acute hamstring strain affects flexibility, strength, and time to return to pre-injury level. *British Journal of Sports Medicine., 40*(1), 40-44.

14. Harvey, L. (1996). Principles of conservative management for a non-orthotic tenodesis grip in tetraplegics. *Journal of Hand Therapy, 9*(3), 238-242.

15. Johanson, M. & Murray, W. (2002). The unoperated hand: the role of passive forces in hand function after tetraplegia. *Hand Clinics., 18*(3), 391-398.

16. Curtin, M. (1994). Development of a tetraplegic hand assessment and splinting protocol. *Spinal Cord.*, *32*(3), 159-169.

17. DiPasquale-Lehnerz, P. (1994). Orthotic intervention for development of hand function with c-6 quadriplegia. *American Journal of Occupational Therapy, 48*(2), 138-144.

18. Jung, H. Y., Lee, J. & Shin, H. I. (2018). The natural course of passive tenodesis grip in individuals with spinal cord injury with preserved wrist extension power but paralyzed fingers and thumbs. *Spinal Cord.*, *56*(9), 900-906.

19. Walter, S., Eliasziw, M. & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Stat Med., 17*(1), 101-110.

20. Norkin, C., White, D., Torres, J., *et al.* (2016). Measurement of Joint Motion: A Guide to Goniometry, 140-142.

21. Davis, D., Quinn, R., Whiteman, C., *et al.* (2008). Concurrent validity of four clinical tests used to measure hamstring flexibility. *Journal of Strength and Conditioning Research, 22*(2), 583-588.