Reliability for the measurements of the active neck and little finger extension, and trunk flexion ranges using a tape measure in healthy young adults

Marta Grychtoł[1], Wiktoria Nowak[1], Bartosz Głowacki[1], Weronika Nalewajka[1], Wiktoria Gural[1], Agnieszka Opala-Berdzik[2]

[1] Department of Physiotherapy in Internal Diseases, Students Scientific Association on the Analysis of the Influence of Pregnancy on the Musculoskeletal System, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland

[2] Institute of Physiotherapy and Health Sciences, Department of Physiotherapy in Internal Diseases, The Jerzy Kukuczka Academy of Physical Education, Katowice, Poland

Correspondence to: Agnieszka Opala-Berdzik, email: a.opala-berdzik@awf.katowice.pl

DOI: https://doi.org/10.5114/phr.2024.140785

Received: 08.12.2023 **Reviewed:** 21.01.2024 **Accepted:** 22.01.2024

Abstract

Background: Measurements of range of motion (ROM) of the spine and extremities are essential in physiotherapy research and practice. These measurements help to identify hypo- or hypermobility in joints associated with various medical conditions and to assess the effects of therapies. Simple tools such as goniometers and tape measures are commonly used, but their accuracy can vary. It is therefore important to determine the reliability of these tools for different joints and movements.

Aims: The aim of this study was to determine the repeatability of tape measurements of neck and little finger extension and trunk flexion.

Material and methods: Thirty young healthy adults (15 females and 15 males) participated in two test sessions 7-14 days apart. During each session, two assessors took tape measurements of their active neck and little finger extension ranges and trunk flexion range using the finger floor distance (FFD) test. They also took goniometric measurements of the little finger extension range. Intra- and inter-rater measurement reliability was assessed using intraclass correlation coefficient (ICC) analyses.

Key words

tape measurement, range of motion, neck, trunk, finger, measurement reliability.

Results: The ICCs for intra- and interrater reliability of the neck extension tape measurements were 0.88-0.98, the active little finger extension 0.95-0.97, and the FFD test 0.90-0.98. For the little finger extension goniometer measurements, the ICCs were 0.95-0.98.

Conclusions: The repeatability of the tape measurements of the active neck extension range was good or excellent, that of the little finger extension range was excellent, and that of the FFD test ranged from good to excellent. The repeatability of the goniometer measurements of the little finger extension range was excellent. The results suggest that the measurements evaluated can be used in scientific research and clinical practice.

Introduction

Measurements of range of motion (ROM) of the spine and extremities are a very important part of physiotherapy research and practice. Pathological ROM, which is associated with various medical conditions, can manifest as either hypo- or hypermobility of the joints. The results of ROM measurements are used in observational studies of healthy subjects and people with various medical conditions, as well as to assess the effects of therapy [1–4].

The basic, simple and inexpensive tools for measuring active joint ROM are the goniometer and the tape measure. However, these measurements are subject to varying degrees of error, especially in relation to specific joints or directions of movement [5–7]. For this reason, it is very important to know which ROM measurements can be used in goal-directed research because of their acceptable reliability, and which are less reliable and should not be considered.

With this in mind, we reviewed the literature for studies that evaluated the reliability of ROM measurements of the spine and extremities using a tape measure, goniometer or other simple tools. The majority of these were on the cervical spine [8–12], thoracolumbar spine [5, 6, 13–15] and hand [16–18]. Some of these studies differed in the interpretation of the intraclass correlation coefficient (ICC) results. For consistency, this paper follows a guideline for reporting ICCs in reliability research [19].

According to two studies, of all active neck ROM measurements (in all planes) using cervical and universal goniometers, neck extension measurements had the highest intra- and interrater reliability [10, 11]. In one of these studies, the interrater reliability of this measurement was excellent when using a cervical goniometer and good when using a universal goniometer. The authors suggested that neck extension represents the full range of active cervical motion and that its measurement should be used to screen for neck motion [10]. However, in the other study, the intra- and interrater reliability of this measurement using a cervical goniometer was lower, ranging from fair to good [11]. Another study reported that the intrarater reliability of tape measurements of neck flexion and extension ranges was moderate [8].

For measurements of thoracolumbar spine ROM/trunk mobility, the intra- and interrater reliability of the forward flexion finger-to-floor distance (FFD) measurement using a tape measure has been reported to be good to excellent [5, 6, 13, 14]. Because of its excellent validity, reliability and responsiveness, this measurement has been suggested for use in clinical practice and therapeutic trials [15]. In turn, the intrarater reliability for lateral flexion FFD measurements was moderate to good in one study [5], and the interrater reliability was excellent in another [6]. It has also been reported that the intrarater reliability of the Schober test was poor to moderate [5], whereas

the interrater reliability of the modified Schober test was good [14] and excellent [6].

With regard to hand joints, the assessment of intra- and interrater reliability of finger goniometer measurements covered the active and passive ranges of middle finger flexion in the metacarpophalangeal, proximal interphalangeal and distal interphalangeal joints. The reliability of the measurements ranged from poor to excellent and was lower for passive than for active ROM [16]. Another study evaluated the reliability of different active ROM measurement methods for the thumb carpometacarpal joint. For example, the interrater reliability of radial abduction measurement was excellent using the pollexograph-thumb angle method and the thumb-distal-interphalangeal distance tape measure method. Good interrater reliability was achieved using the pollexograph metacarpal angle method and the Moeltgen goniometer intermetacarpal angle method [17]. Test-retest reliability was also reported for the goniometer measurements of active wrist and finger ROM. Reliability was good for wrist flexion/extension range and moderate to good for the I-V metacarpophalangeal joints, with the highest results for the little finger [18].

The above-mentioned literature suggests that some measurement methods vary in reliability between research centres and that not all of them achieve at least good reliability, which is considered to be clinically acceptable [5]. Therefore, we wanted to verify or determine whether selected ROM tape measurements in three representative parts of the body, such as the neck, trunk and fingers, could be performed reliably. As the tape measurement of the little finger extension area proposed in this study has not yet been presented and evaluated, we also evaluated the measurement of this ROM with a goniometer.

If these methods of ROM measurement prove to be reliable, they could be used, for example, in longitudinal studies of the influence of hormonal changes associated with pregnancy or the menstrual cycle on joint mobility. In such studies, the mobility of at least a few body parts should be examined. The measurements should also be considered when determining the effect of specific treatments on joint mobility in conditions such as "smartphone neck" with shortening of the superficial neck flexors, ulnar nerve palsy, reduced ROM associated with non-specific pain in the spinal region, or reduced hip joint mobility due to hamstring tightness, etc.

Aims

The aim of the study was to assess the intra- and interrater reliability of tape measures of neck and little finger extension and trunk flexion (using the finger floor distance test) in healthy young adults. In addition, the reliability of the goniometric measurement of the little finger extension range was to be tested. In this way, we wanted to determine whether these non time-consuming, easy to perform and interpret measurements could be used in scientific research. We assumed that all measurements would have at least good reliability.

Material and methods

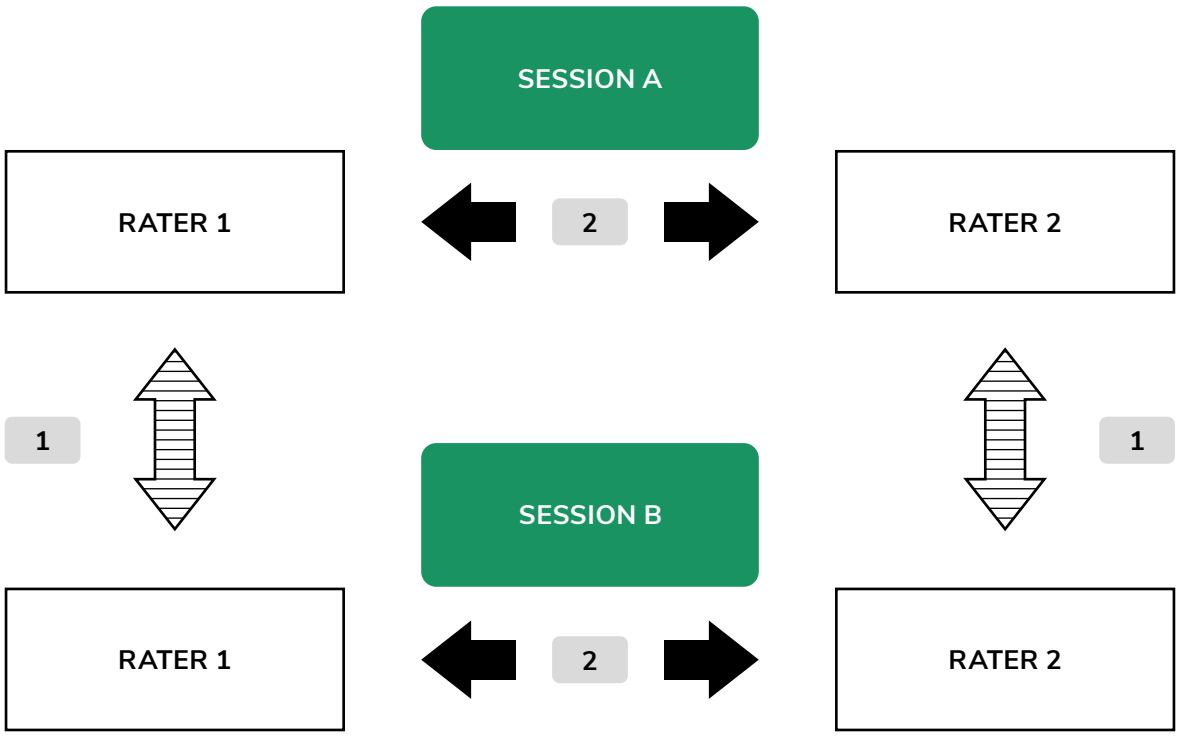
Participants

Thirty young healthy adults (15 females and 15 males) took part in the study. Participants were recruited from students at the Academy of Physical Education in Katowice, Poland. Recruitment was carried out by members of the research team using social media and word of mouth. Inclusion criteria for participation in the study were age between 18 and 30 years and BMI ≤ 30. As joint mobility may depend on changes in estrogen levels during the menstrual cycle [20], the inclusion criteria for women were also being in the luteal phase of the menstrual cycle or using oral contraception/hormonal therapy. Exclusion criteria for women were being in the follicular and ovulatory phases or being pregnant. The phases of the menstrual cycle were monitored by the participants using a special telephone application. Exclusion criteria for the whole group were any

limitations in the range of motion of the musculoskeletal system, especially pain and/or stiffness in the neck, right hand, back and/or hamstrings. Informed consent was obtained from each participant after full explanation of the purpose and protocol of the study. The study was approved by the institutional bioethics committee (approval number 2/2023).

Procedure

The study consisted of two examination sessions, one to two weeks apart (sessions A and B). Both sessions took place in the same room, at the same temperature and at the same time of day. Participants were asked not to change their dietary, mealtime, or exercise habits between sessions. Each session included measurements of 1) neck extension range using a tape measure, 2) little (fifth) finger extension range using a) a tape measure and b) a goniometer, 3) trunk flexion range using the finger-floor distance test (using a tape measure). Each measurement was performed by two raters in random order. Both raters were blinded to each other's results during the measurements. **Figure 1** shows a scheme for assessing the intra- and inter-rater reliability of the measurements. The subjects' height and mass were measured at the first session to calculate BMI.



Notes: *Striped arrows (1) indicate intrarater reliability (Session A vs. Session B); solid arrows (2) indicate interrater reliability, separately in Session A and Session B.*

Figure 1. Scheme for assessing the reliability of measurements of the active neck and little finger extension, and trunk flexion range in 30 healthy young adults (15 females and 15 males).

Raters

All measurements were carried out by the first two authors, who were fourth year physiotherapy students. They were trained by the senior author – a physiotherapist with 15 years' experience in physiotherapy research – to be proficient in each of the measurement methods before the study began. The senior author observed each rater's technique and gave additional instructions when needed to achieve the best accuracy.

Measurements

To measure the range of active neck extension, the subject was seated in a straight-back chair with the lower limbs resting on the floor. The upper limbs rested freely on the thighs and the head was in a neutral vertical position. The subject was asked to perform a maximal extension of the neck (as she/he wanted to look up and back as far as possible), keeping the trunk and shoulders still and moving the head in the sagittal plane. She/he was told that the movement should be natural, not forced, and that it should be performed only in the cervical region. A tape measure was used to take a measurement at the extreme of active neck extension. The zero point of the tape measure was placed on the jugular notch of the sternum (its lowest point) and the distance from this point to the tip of the chin (the lowest point of the mental prominence) was measured [8, 9] (**Figure 2a**).

To measure the range of active extension of the little finger with a tape measure, the subject was seated with the right forearm and palm resting on the tabletop (at the edge) with the thumb and all fingers flat on the tabletop, parallel to each other. He/she was asked to lift the little finger off the tabletop as far as possible. During the test, the uninspected fingers of the right hand remained flat on the tabletop and did not move. The tape measure was used to measure the shortest distance between the most distal part of the little fingertip and the tabletop (**Figure 2b**).

To measure the range of active extension of the little finger with a goniometer, the procedure (position of the subject, their right upper extremity and the task performed) was the same as described above. The measurement was made by applying the goniometer to the fifth metacarpophalangeal joint (**Figure 2c**). The hand position for measuring the active range of little finger extension was chosen according to the indications and protocols used in previous studies [21, 22].

The finger-to-floor distance test was used to measure the range of trunk flexion. For the test, the subject assumed a standing position (barefoot) on a 15 cm platform with feet hip width apart. The subject was asked to raise the upper limbs with the elbow joints extended (to an angle of approximately 90° flexion at the shoulder joints). The trunk flexion movement began with a slow neck flexion and continued with successive flexions of the lower spinal segments. The knee joints were straight throughout the test. The movement should be natural and not forced. The subject was instructed not to deepen the flexion of the trunk, so that the movement took place to its first limit. A tape measure was used to measure from the tip of the longest finger of the right hand to the top of the platform. The top of the platform had a value of zero. If the subject reached below the height of the platform, the distance was recorded in negative values, and if he did not reach the platform – in positive values [15] (**Figure 2d**).

Each of the four measurements was taken three times in order to use the averaged results in the statistical analysis. There were approximately 60 seconds between measurements.

Statistical analysis

Sample size was determined based on an intraclass correlation coefficient (ICC) test with a significance level (α) of 0.05 (type I error), target power 1 - β of 0.8 (80% power) and number of raters / repetitions per subject (k) = 2. The minimum

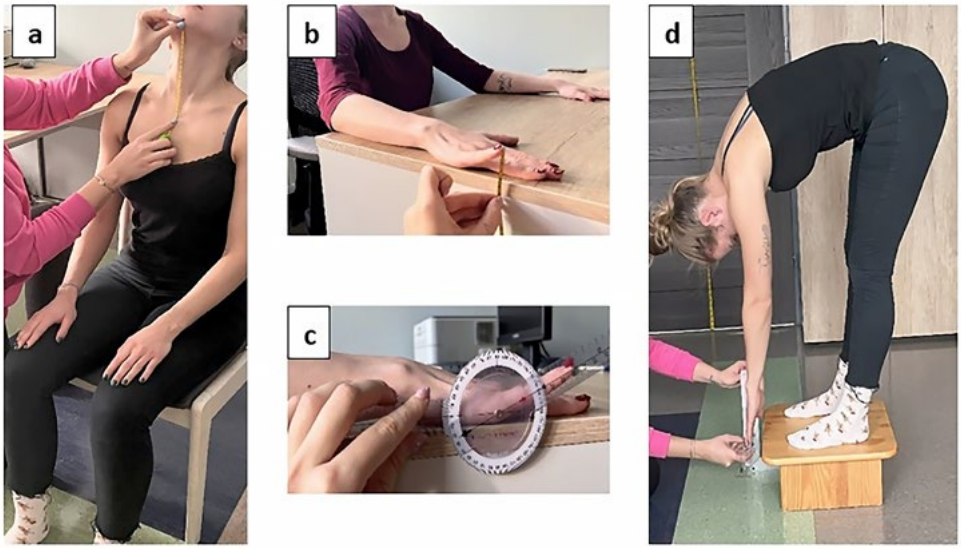


Figure 2. Tape measurements of active neck extension range (a), little finger extension range (b), trunk flexion/finger floor distance test (d), and goniometer measurement of little finger extension range (c).

acceptable reliability was 0.75 and the expected reliability was 0.90 [19, 23]. The results of the calculation were as follows: the sample size was n = 26 and the sample size with a 10% dropout rate was ndrop = 29. Thirty subjects were recruited to form equal groups of females and males.

The Shapiro-Wilk test was used to determine whether the data had a normal distribution, and Levene's test was used to check for homogeneity of variance in the whole group and in the female and male subgroups (for each measurement: 1) neck range of active extension with a tape measure, 2) little finger range of active extension a) with a tape measure and b) with a goniometer, and 3) trunk flexion with the finger floor distance test). Skewness, kurtosis and modality of the data were checked. As there were no reasons to reject the null hypotheses, ICC analyses were performed to assess the intra- and interrater reliability of the measurements. Means and standard deviations were calculated for each parameter. Repeated measures ANOVA (repeated factorial – two levels) was used to calculate 2k ICCs (two-way random effects, absolute agreement, multiple raters/measurements) and 95% confidence intervals were calculated. As each measurement was performed three times, k refers to the number of measurements averaged before entering the ICC analysis [19]. The following ICC intervals were chosen: values < 0.5 indicate poor reliability, 0.5-0.75 indicate moderate reliability, 0.75-0.90 indicate good reliability, and > 0.90 indicate excellent reliability [19].

The standard error of measurement (SEM) was also calculated ($SEM=SD\sqrt{(1-ICC)}$) and the minimal detectable change (MDC,95%) was calculated as $MDC = SEM*1.96*\sqrt{2}$ [24]. Statistical significance was set at $p < 0.05$. The data was analyzed using RStudio 2022.02.0+443 Prairie Trillium (Posit PBC, Boston, MA, USA).

Results

Of the 15 male participants, nine were examined seven days apart and six were examined 14 days apart. Of the 15 female participants, nine were examined seven days apart and six were examined 11-14 days apart. Six women were using oral contraceptives/hormone therapy. It was not possible to repeat the seven-day measurements in all subjects as originally planned because some subjects had to postpone their second visit. The characteristics of the subjects are presented in **Table 1**.

The results of two raters' measurements of active neck and little finger extension and trunk flexion range in the whole group and separately for men and women from two examination sessions are shown in the table below **Table 2**.

Each rater took a total of 720 measurements on 30 participants during sessions A and B (i.e., 180 measurements with a tape measure of 1) neck extension range, 2) finger extension range, 3) trunk flexion range, and 4) finger extension range with a goniometer).

Intrarater measurement reliability

Tape measurements of active neck extension showed good intra-rater reliability when taken by each rater in the total group of 30 adults and in a subgroup of 15 males. They showed excellent intra-rater reliability when taken in the subgroup of 15 females.

Tape and goniometer measurements of active little finger extension showed excellent intrarater reliability when taken by each rater in the total group and in the subgroups of males and females.

The intrarater reliability of the active trunk flexion range measurement (using the FFD test) was excellent when taken by Rater 1 (excellent in the male and good in the female subgroups). The intrarater reliability of this test was good when performed by rater 2 (excellent in male and moderate in female subgroups). The ICC, 95% CI, SEM and MDC95 values are shown in **Table 3**.

Table 1. Characteristics of 30 healthy young adults (15 females and 15 males) who underwent measurements of active neck and little finger extension and trunk flexion range during two examination sessions that were 7-14 days apart*.

Chatracter	Females	Males
Age (years)	23.6 ± 2.5 (21 - 30)	22.1 ± 1.2 (19 - 24)
Body Height (cm)	169.7 ± 5.0 (163 - 183)	180.7 ± 4.7 (172 - 190)
Body mass (kg)	69.1 ± 10.2 (53.1 – 85.6)	77.25 ± 10.7 (59.2 - 92.1)
BMI	24.0 ± 3.6 (18.6 – 28.6)	23.7 ± 3.2 (18.3 - 30)

Notes: * *Data are shown as means ± standard deviations (ranges).*

Table 2. Measurements of active neck and little finger extension and trunk flexion range by two raters in 30 young adults (15 males and 15 females) from two examination sessions 7-14 days apart.

Measurements	Session A			Session B		
	Subjects	Males	Females	Subjects	Males	Females
Rater 1						
Neck extension (cm)	21.9±1.3	22.1±1.3	21.8±1.4	22.1±1.5	22.6±1.5	21.5±1.3
Little finger extension (cm)	5.0±1.0	4.8±1.0	5.1±1.1	5.1±1.0	5.0±0.8	5.3±1.2
Little finger extension (°)	27.1±8.9	25.4±6.4	28.8±10.9	27.2±9.5	26.2±7.3	28.3±11.4
FFD test (cm)	-0.7±9.4	1.4±11.1	-2.8±7.2	-2.3±8.3	-0.4±9.2	-4.2±7.0
Rater 2						
Neck extension (cm)	22.0±1.5	22.0±1.5	21.9±1.5	22.2±1.5	22.8±1.4	21.6±1.4
Little finger extension (cm)	5.0±1.0	4.9±0.9	5.0±1.2	5.0±0.9	4.9±0.6	5.1±1.1
Little finger extension (°)	29.1±9.1	27.2±7.6	30.9±10.3	28.8±8.6	27.7±6.3	29.9±10.6
FFD test (cm)	-3.4±8.3	-1.1±9.8	-5.8±5.9	-4.0±7.8	-2.1±8.7	-6.0±6.4

Notes: * *Data are shown as means ± standard deviations.*

Abbreviations: *(cm) – measurement with a tape measure; (°) – measurement with a goniometer; FFD – Finger floor distance test.*

Table 3. Intrarater reliability for the measurements of active neck and little finger extension, and trunk flexion range performed by two raters in 30 young adults (15 males and 15 females) during two examination sessions (A and B) 7-14 days apart.

Measurements	Intrarater reliability (session A vs Session B)			
	ICC 2,k	95% CI	SEM	MDC_{95}
YOUNG ADULTS (n=30)				
	Rater 1			
Neck extension (cm)	0.90	0.80-0.95	0.4	1.2
Little finger extension (cm)	0.97	0.90-0.99	0.2	0.5
Little finger extension (°)	0.98	0.95-0.99	1.3	3.6
FFD test (cm)	0.94	0.87-0.97	2.0	5.5

	Rater 2			
Neck extension (cm)	0.88	0.75-0.94	0.5	1.4
Little finger extension (cm)	0.95	0.90-0.98	0.2	0.6
Little finger extension (°)	0.95	0.90-0.98	2.0	5.4
FFD test (cm)	0.90	0.79-0.95	2.5	7.0
MALES (n=15)				
	Rater 1			
Neck extension (cm)	0.87	0.53-0.96	0.5	1.4
Little finger extension (cm)	0.94	0.83-0.98	0.2	0.6
Little finger extension (°)	0.95	0.85-0.98	1.5	4.2
FFD test (cm)	0.96	0.87-0.99	2.0	5.6
	Rater 2			
Neck extension (cm)	0.83	0.29-0.95	0.6	1.7
Little finger extension (cm)	0.92	0.76-0.97	0.2	0.6
Little finger extension (°)	0.95	0.86-0.98	1.5	4.3
FFD test (cm)	0.97	0.92-0.99	1.6	4.4
FEMALES (n=15)				
	Rater 1			
Neck extension (cm)	0.94	0.81-0.98	0.3	0.9
Little finger extension (cm)	0.99	0.94-1.00	0.1	0.3
Little finger extension (°)	0.99	0.96-1.00	1.1	3.0
FFD test (cm)	0.90	0.71-0.97	2.2	6.2
	Rater 2			
Neck extension (cm)	0.93	0.78-0.98	0.4	1.0
Little finger extension (cm)	0.97	0.92-0.99	0.2	0.5
Little finger extension (°)	0.95	0.86-0.98	2.3	6.4
FFD test (cm)	0.66	-0.05-0.89	3.5	9.8

Abbreviations: *ICC 2,k – intraclass correlation coefficient; CI – confidence interval; SEM – standard error of measurement; MDC – minimal detectable change; (cm) – measurement with a tape measure, refers to SEM and MDC; (°) – measurement with a goniometer, refers to SEM and MDC; FFD test – finger floor distance test.*

Interrater measurement reliability

Tape measures of active neck extension range showed excellent inter-rater agreement when considering the results for the whole group and the male and female subgroups in sessions A and B.

Tape measurements of active little finger extension also showed excellent inter-rater agreement for the whole group and male and female subgroups in both sessions.

In sessions A and B, goniometer measurements of active little finger extension showed excellent inter-rater agreement when taken in the whole group and in the female subgroup, and good agreement in the male subgroup.

In both sessions, the FFD test showed excellent inter-rater agreement when taken in the whole group and in the male subgroup. It also showed excellent agreement when administered in the female subgroup in session B, but good agreement in session A. The ICC, 95% CI, SEM and MDC95 values are shown in **Table 4**.

Table 4. Interrater reliability for the measurements of active neck and little finger extension, and trunk flexion range performed in 30 young adults (15 males and 15 females) during two examination sessions (A and B) 7-14 days apart.

Measurements	Interrater reliability (Rater 1 vs Rater 2)			
	ICC 2,k	95% CI	SEM	MDC_{95}
YOUNG ADULTS (n=30)				
	Session A			
Neck extension (cm)	0.97	0.94-0.99	0.2	0.7
Little finger extension (cm)	0.97	0.93-0.98	0.2	0.5
Little finger extension (°)	0.95	0.86-0.98	2.0	5.6
FFD test (cm)	0.92	0.74-0.97	2.5	7.0
	Session B			
Neck extension (cm)	0.98	0.96-0.99	0.2	0.6
Little finger extension (cm)	0.97	0.93-0.99	0.2	0.5
Little finger extension (°)	0.96	0.88-0.98	1.7	4.6
FFD test (cm)	0.98	0.80-0.99	1.1	3.1
MALES (n=15)				
	Session A			
Neck extension (cm)	0.97	0.90-0.99	0.2	0.7
Little finger extension (cm)	0.95	0.84-0.99	0.2	0.4
Little finger extension (°)	0.77	0.27-0.93	1.8	4.9
FFD test (cm)	0.97	0.75-0.99	1.8	5.0

	Session B			
Neck extension (cm)	0.97	0.92-0.99	0.3	0.7
Little finger extension (cm)	0.95	0.85-0.98	0.2	0.4
Little finger extension (°)	0.74	0.26-0.92	1.8	5.0
FFD test (cm)	0.98	0.85-1.00	1.3	3.5
FEMALES (n=15)				
	Session A			
Neck extension (cm)	0.98	0.93-0.99	0.2	0.6
Little finger extension (cm)	0.96	0.88-0.99	0.2	0.6
Little finger extension (°)	0.96	0.85-0.99	2.1	5.8
FFD test (cm)	0.78	0.33-0.93	3.1	8.6
	Session B			
Neck extension (cm)	0.99	0.96-1.00	0.1	0.4
Little finger extension (cm)	0.98	0.92-0.99	0.2	0.5
Little finger extension (°)	0.98	0.92-0.99	1.5	4.2
FFD test (cm)	0.97	0.68-0.99	1.2	3.2

Abbreviations: *ICC 2,k – intraclass correlation coefficient; CI – confidence interval; SEM – standard error of measurement; MDC – minimal detectable change; (cm) – measurement with a tape measure, refers to SEM and MDC; (°) – measurement with a goniometer, refers to SEM and MDC; FFD test – finger floor distance test.*

Discussion

The aim of this study was to assess the intra- and interrater reliability of tape measures of active neck and little finger extension range and active trunk flexion range (using the finger floor distance test) in 30 healthy young adults. In addition, the reliability of goniometric measurement of the little finger extension range was tested. Our practical aim was to determine whether these quick and simple measurements could be used in scientific research.

Our results indicate that the between-session and between-rater agreement of tape measurements of active neck extension range was good to excellent, the little finger extension range was excellent, and the trunk flexion range (FFD test) was good to excellent. The agreement of the goniometer measurements of the little finger extension range was excellent.

In the present study, the tape measurements of active neck extension made by raters 1 and 2 showed good intrarater reliability with ICCs of 0.90 and 0.88, 95% CIs of 0.80-0.95 and 0.75-0.94, and SEMs of 0.4 and 0.5 cm respectively. These results may suggest that the measurements made by Assessor 1 were marginally more repeatable than those made by Assessor 2 (the measurement

error for Assessor 2 was only 0.1 cm greater). At the same time, the inter-rater agreement for the tape measurements of active neck extension was excellent in each of the two examination sessions. In sessions A and B, the ICCs were 0.97 and 0.98, and the 95% CIs were 0.94-0.99 and 0.96-0.99, respectively, while the SEM values were 0.2 cm in round numbers. Slightly higher ICC and confidence interval values may indicate that the ability to make accurate measurements improved with the number of measurements taken. In our study, the reliability of this measurement was higher than in another study, which showed moderate intra- and interrater reliability with an ICC of 0.69 and 0.58, respectively [8]. It cannot be excluded that this is related to the greater homogeneity of age of the participants in the present study. The mean age of the subjects in our study was 22.9±1.9 years, compared to 30.5±9.1 years in that study [8].

With regard to the active extension range of the little finger, both tape and goniometer measurements in this study showed excellent intra- and interrater reliability. Tape measurements by Raters 1 and 2 had ICCs of 0.97 and 0.95, 95% CIs of 0.90-0.99 and 0.90-0.98, respectively, and SEM values of 2 mm in round numbers. The goniometer measurements had ICCs of 0.98 and 0.95, 95% CIs of 0.95-0.99 and 0.90-0.98, and SEMs of 1.3 and 2.0 degrees, respectively. These results may indicate a slightly higher repeatability of the tape and goniometer measurements taken by Rater 1. Tape measurements in sessions A and B again had ICCs of 0.97, 95% CIs of 0.93-0.98 and 0.93-0.99 respectively, and SEMs of 0.2 cm. For the goniometer measurements, the ICCs were 0.95 and 0.96, 95% CIs were 0.86-0.98 and 0.88-0.98, and SEMs were 2 and 1.7 degrees. These results may suggest that the ability to make accurate measurements of finger extension range using the goniometer improves slightly with more measurements.

We have not found a reliability study of active finger ROM measurement using a similar methodology to ours. The authors of one study showed that the test-retest reliability of goniometer measurements of active ROM at the fifth metacarpophalangeal joint from maximum flexion to maximum extension was good with an ICC of 0.86 [18], but we cannot compare the present results with this study and draw any conclusions because the measurement procedures of these two experiments were very different. In our study, the starting position for ROM measurement was the neutral position of the little finger, and in that study, it was maximal flexion.

The present study also shows that the intrarater reliability of the active trunk flexion range measurement (using the FFD test) was excellent when taken by Rater 1 and good when taken by Rater 2. The measurements taken by Raters 1 and 2 had ICCs of 0.94 and 0.90, 95% CIs of 0.87-0.97 and 0.79-0.95, and SEMs of 2 and 2.5 cm respectively. This suggests that evaluator 1 was able to make measurements with greater repeatability than evaluator 2. The FFD test also showed excellent inter-rater agreement. In sessions A and B, the ICCs were 0.92 and 0.98, the 95% CIs were 0.74-0.97 and 0.80-0.99, and the SEMs were 2.5 and 1.1 cm. The better results obtained in session B may indicate that the raters' ability to make accurate measurements during the FFD test has improved over time. Our results for measurement agreement on the FFD test in session B (ICC of 0.98) are slightly lower than those obtained in another study (ICC of 0.999), but the design of that study was different. It concerned the assessment of the agreement between seven raters who performed measurements on three subjects [6]. It should be remembered that the result of the ROM measurement is influenced by both the subject's performance of the task and the measurement, so an adequate sample size is important in reliability studies.

Study limitations

Our results are limited to young healthy individuals (aged between 18 and 30 years) and should not be extrapolated to younger or older populations or to populations of individuals with musculoskeletal dysfunction or obesity. In addition, this

study population consisted of cooperative volunteers who were students at the Academy of Physical Education, so the results suggesting highly reliable measurements may also be related to good compliance of the subjects in following the instructions on how to perform a given task. The ICC analyses were performed on an a priori sample of 30 subjects, but additional analyses were performed for the subgroups of 15 females and 15 males. As these analyses may have had less statistical power, the results concerning the subgroups should be considered with caution. Therefore, the discussion and conclusions are based only on the sample of 30 healthy participants.

Conclusions

In this study, the intra- and interrater reliability of tape measures of active neck extension and trunk flexion (using the FFD test) was good to excellent. The reliability of the tape and goniometer measurements of active little finger extension range was excellent. This level of agreement between measurements can be considered clinically acceptable [5]. This study suggests that proper standardisation of the measurement procedure at all stages and consistent training of the raters in the procedures by an instructor resulted in good to excellent repeatability of ROM measurements by the same rater over time and between raters. Slightly higher inter-rater agreement in session B than in session A for most measures may be related to further improvement in measurement accuracy with more practice. The results of this study suggest that tape measures of active neck and little finger extension range and trunk flexion range (using the FFD test) can be used in research and clinical practice.

Declarations

Ethical Considerations: Permission for the study was received from the institutional bioethics committee (permission nr 2/2023). Informed consent was obtained from each participant after providing comprehensive information on the study's aim and protocol. The study was designed and conducted in accordance with the Declaration of Helsinki (1964) and Good Clinical Practice (GCP) guidelines.

Clinical Trials: This study was not registered as a clinical trial as it did not involve investigational products or interventions that would classify it under clinical trial regulations.

Conflict of Interest: The authors declare no conflict of interest. The study was conducted independently and without any influence from external organizations or entities.

Funding Sources: This research received no external funding and did not receive any grants or financial support from external sources, including non-profit organizations. The study was conducted using the internal resources of the institutions involved.

References

1. Medeiros DM, Martini TF. Chronic effect of different types of stretching on ankle dorsiflexion range of motion: Systematic review and meta-analysis. Foot (Edinb). 2018; 34: 28–35.
2. Reddy RS, Tedla JS, Dixit S, Abohashrh M. Cervical proprioception and its relationship with neck pain intensity in subjects with cervical spondylosis. BMC Musculoskelet Disord. 2019; 20 (1): 44.
3. Soucie JM, Wang C, Forsyth A, Funk S, Denny M, Roach KE, Boone D. Hemophilia Treatment Center Network. Range of motion measurements: reference values and a database for comparison studies. Haemophilia. 2011; 17 (3): 500–507.

4. Stevens-Lapsley JE, Balter JE, Wolfe P, Eckhoff DG, Kohrt WM. Early neuromuscular electrical stimulation to improve quadriceps muscle strength after total knee arthroplasty: a randomized controlled trial. Phys Ther. 2012; 92 (2): 210–226.
5. Aartun E, Degerfalk A, Kentsdotter L, Hestbaek L. Screening of the spine in adolescents: inter- and intra-rater reliability and measurement error of commonly used clinical tests. BMC Musculoskelet Disord. 2014; 15: 37.
6. Johnson M, Mulcahey MJ. Interrater Reliability of Spine Range of Motion Measurement Using a Tape Measure and Goniometer. J Chiropr Med. 2021; 20 (3): 138–147.
7. Pua YH, Wrigley TV, Cowan SM, Bennell KL. Intrarater test-retest reliability of hip range of motion and hip muscle strength measurements in persons with hip osteoarthritis. Arch Phys Med Rehabil. 2008; 89 (6): 1146–1154. Erratum in: Arch Phys Med Rehabil. 2008; 89 (8):1628. Wrigley Tim W [corrected to Wrigley, Tim V].
8. Asha SE, Pryor R. Validation of a method to assess range of motion of the cervical spine using a tape measure. J Manipulative Physiol Ther. 2013; 36 (8): 538-545.
9. Hsieh CY, Yeung BW. Active neck motion measurements with a tape measure. J Orthop Sports Phys Ther. 1986; 8 (2): 88–90.
10. Whitcroft K, Massouh L, Amirfeyz R, Bannister G. Comparison of methods of measuring active cervical range of motion. Spine (Phila Pa 1976). 2010; 35 (19): 976–980.
11. Wolan-Nieroda A, Guzik A, Mocur P, Drużbicki M, Maciejczak A. Assessment of Interrater and Intrarater Reliability of Cervical Range of Motion (CROM) Goniometer. Biomed Res Int. 2020; 2020: 8908035.
12. Wolfenberger VA, Bui Q, Batenchuk GB. A comparison of methods of evaluating a cervical range of motion. J Manipulative Physiol Ther. 2002; 25 (3): 154–160.
13. Ayala F, Sainz de Baranda P, De Ste Croix M, Santonja F. Absolute reliability of five clinical tests for assessing hamstring flexibility in professional futsal players. J Sci Med Sport. 2012; 15 (2): 142–147.
14. Robinson HS, Mengshoel AM. Assessments of lumbar flexion range of motion: intertester reliability and concurrent validity of 2 commonly used clinical tests. Spine (Phila Pa 1976). 2014; 39 (4): E270–275.
15. Perret C, Poiraudeau S, Fermanian J, Colau MM, Benhamou MA, Revel M. Validity, reliability, and responsiveness of the fingertip-to-floor test. Arch Phys Med Rehabil. 2001; 82 (11): 1566–1570.
16. Lewis E, Fors L, Tharion WJ. Interrater and intrarater reliability of finger goniometric measurements. Am J Occupational Ther. 2010; 64, 555–561.
17. Holzbauer M, Hopfner M, Haslhofer D, Kwasny O, Duscher D, Froschauer SM. Radial and palmar active range of motion measurement: reliability of six methods in healthy adults. J Plast Surg Hand Surg. 2021; 55 (1): 41–47.
18. Reissner L, Fischer G, List R, Taylor WR, Giovanoli P, Calcagni M. Minimal detectable difference of the finger and wrist range of motion: comparison of goniometry and 3D motion analysis. J Orthop Surg Res. 2019; 14 (1): 173.
19. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016; 15 (2): 155–163. Epub 2016. Erratum in: J Chiropr Med. 2017; 16 (4): 346.
20. Martínez-Fortuny N, Alonso-Calvete A, Da Cuña-Carrera I, Abalo-Núñez R. Menstrual Cycle and Sport Injuries: A Systematic Review. Int J Environ Res Public Health. 2023; 20 (4): 3264.
21. Latz D, Koukos C, Boeckers P, Jungbluth P, Schiffner E, Kaufmann R, Gehrmann SV. Influence of Wrist Position on the Metacarpophalangeal Joint Motion of the Index Through Small Finger. Hand (N Y). 2019; 14 (2): 259–263.
22. Weis CA, Grondin D, Vernon H. The Effect of Phase of Menstrual Cycle on Joint Mobility in the Cervical Spine and Extremities in Nulliparous Women: A Cross-Sectional Study. J Manipulative Physiol Ther. 2016; 39 (6): 393–400.
23. Arifin, WN. A web-based sample size calculator for reliability studies. Education in Medicine Journal. 2018; 10 (3): 67–76.
24. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res. 2005; 19 (1): 231–240.