

Maciej Biały^{1,3 (A,B,D,E,F)}, Agata Dziewońska^{1,3 (A,B,C,E,F)}, Magdalena Sroka- Buszydlik^{2,3 (A,B,C,E,F)}

Wiarygodność pomiaru przekroju poprzecznego mięśnia wielodzielnego z wykorzystaniem ultrasonografii

Reliability of measurement of lumbar multifidus thickness using ultrasonography

¹ Studia doktoranckie na Wydziale Wychowania Fizycznego, Akademia Wychowania Fizycznego im. Jerzego Kukuczki, Katowice

² Centrum Leczenia Schorzeń Kręgosłupa i Limfologii, Ustroń

³ Grupa badawcza Collegium Magicum

STRESZCZENIE

Wstęp: Ocenie poddana została wiarygodność pomiaru, przekroju poprzecznego mięśnia wielodzielnego w lędźwiowym odcinku kręgosłupa z wykorzystaniem ultrasonografii.

Materiał i metody: W badaniach uczestniczyło dziesięć osób (K = 4), w przedziale wiekowym 19–26 lat. Byli to studenci Akademii Wychowania Fizycznego, prezentujący relatywnie zbliżony, aktywny styl życia. Kryterium wyłączenia z badania, stanowiły aktualne dolegliwości bólowe: lędźwiowego odcinka kręgosłupa, miednicy i kończyn dolnych oraz urazy i zabiegi chirurgiczne w tych okolicach ciała w przeszłości. Jako narzędzie pomiarowe wykorzystano aparat USG Mindray DP6600. Badania przeprowadzono trzy razy w tygodniu (poniedziałek, środa, piątek). Pomiar powtarzany był pięciokrotnie przez dwóch badających. Pozwoliło to na oszacowanie współczynników określających wiarygodność pomiaru zarówno jednego, jak i dwóch badaczy, a tym samym na identyfikację wielkości i źródeł błędów pojawiających się przy pomiarach tego typu.

Wyniki: Wykonanie co najmniej dwóch powtórzeń pomiaru wielkości przekroju poprzecznego mięśnia wielodzielne-

ABSTRACT

Background: The objective of this study was to evaluate the reliability of measurement of lumbar multifidus muscle thickness using real-time ultrasonography.

Material and methods: Ten subjects (F = 4) aged 19–26 years were selected for participation. All were students of the University of Physical Education presenting relatively similar, active lifestyle. All met the following inclusion criteria: lack of any current pain in the area of the lumbar spine, pelvis and lower limbs; lack of any previous injuries/surgeries in these areas of the body in the past. A Mindray DP6600 ultrasound device was used with a linear 75L38EA array. The measurements were carried out three times a week (Monday, Wednesday, Friday). Measurements were repeated five times by the two investigators. This allowed for estimation of reliability indices and identification of the measurement error size and its sources.

Results: The results indicate that two repetitions of the muscle thickness measurements are sufficient to obtain fair and excellent reliability defined by intraclass correlation coefficients (ICCs ranging from 0,81 to 0,99). Five repetitions of change in muscle thickness

Udział współautorów / Participation of co-authors: A – przygotowanie projektu badawczego/ preparation of a research project; B – zbieranie danych / collection of data; C – analiza statystyczna / statistical analysis; D – interpretacja danych / interpretation of data; E – przygotowanie manuskryptu / preparation of a manuscript; F – opracowanie piśmiennictwa / working out the literature; G – pozyskanie funduszy / obtaining funds

go przez jednego i dwóch badających, wykazuje dobrą i doskonałą wartość wskaźników wiarygodności ICC – *intraclass correlation coefficients* (ICC od 0,81 do 0,99). Natomiast wskaźniki wiarygodności zmiany wielkości przekroju poprzecznego, osiągały poziom dobry (ICC od 0,82 do 0,86) po wykonaniu pięciu pomiarów.

Wnioski: By ocenić wiarygodność zmian przekroju poprzecznego, rekomenduje się wykonanie pięciu powtórzeń pomiaru, natomiast badając przekrój poprzeczny – dwóch. Taki poziom wiarygodności usprawiedliwia wykorzystanie przedstawionej metodyki w warunkach klinicznych oraz na potrzeby badań naukowych.

Słowa kluczowe: mięsień wielodzielny, wiarygodność, ultrasonografia

Wstęp

Ultrasonografia (USG) jest jedną z najpopularniejszych metod nieinwazyjnego obrazowania układu mięśniowo-szkieletowego. Do ważniejszych zastosowań tej techniki diagnostycznej zalicza się możliwość dokonania pomiaru przekroju poprzecznego tkanki mięśniowej, co może być wykorzystane w stwierdzaniu jej atrofii bądź hipertrofii. W literaturze odnaleźć można wiele doniesień na temat wykorzystania USG w pomiarach mięśnia czworogłowego uda, piszczelowego przedniego, mięśnia wielodzielnego w odcinku lędźwiowym (ML) oraz poprzecznego brzucha (TrA) [1, 2, 3, 4, 5, 6, 7, 8]. Dodatkowym zastosowaniem tej metody może być kontrola pozycji elektrod igłowych w elektromiografii [9, 10], a także badania naukowe, w których ocenie poddana jest trafność i wiarygodność wykonywanych pomiarów [4, 6], czy testy określające sztywność stawów krzyżowo-biodrowych [11, 12].

W ostatnim dziesięcioleciu w dziedzinie fizjoterapii bardzo popularne stało się wykorzystanie USG w celu oceny efektywności ćwiczeń stosowanych w leczeniu dolegliwości bólowych lędźwiowego odcinka kręgosłupa. W większości tego typu doniesień ocenie poddany był przekrój poprzeczny TrA, jego zmiana obserwowana pomiędzy fazą skurczu i rozluźnienia, a także wiarygodność przeprowadzonych pomiarów [13]. Badaniom z wykorzystaniem USG był również poddawany mięsień ML. Niektóre z nich, podobnie jak w przypadku TrA, miały na celu ocenę wiarygodności pomiarów ML podczas jego aktywacji [4, 14, 15, 16, 17, 18], w innych analizie został poddany jego przekrój poprzeczny (m.in. u badanych z zespołami bólowymi okolicy lędźwiowo-krzyżowej) [18, 14, 19, 20, 21, 22, 23, 24, 25, 26, 27].

Celem niniejszej pracy, jest określenie poziomu wiarygodności (jednego oraz wielu badaczy), który uzasadni zastosowanie pomiarów wykonywanych za pomocą USG zarówno w badaniach naukowych, jak i klinicznych. Ocenie poddano: przekrój poprzeczny ML na poziomie L4 (jego warstwę głęboką, do której można zaliczyć mięśnie skręcające oraz warstwę powierzchowną tego mięśnia), zmianę przekroju poprzecznego ww. mięśnia,

measurement are necessary to obtain fair reliability (ICCs ranging from 0,82 to 0,86).

Conclusions: To evaluate the reliability of muscle thickness change authors indicate to perform five measurements of mentioned variable, to evaluate muscle thickness reliability alone – two measurement. Such a number of repeated measurements justifies the use of presented methodology in clinical and scientific settings.

Key words: lumbar multifidus, reliability, ultrasonography

Introduction

Ultrasonography (US) is one of the most popular methods for a non-invasive imaging of the musculoskeletal system. The most important applications of this diagnostic technique include the ability to measure the cross-sectional area of muscle tissue, which can be used in verifying its atrophy or hypertrophy. There are many reports on the use of US in the measurement of the quadriceps, tibialis anterior muscle, lumbar multifidus (ML) and transversus abdominis muscle (TrA) in literature [1, 2, 3, 4, 5, 6, 7, 8]. Other additional application of this method may include the control of needle electrodes in fine-wire electromyography [9, 10], as well as some scientific research on validity and reliability of measurements [4, 6], or tests determining sacroiliac joint stiffness [11, 12].

In the last decade in the field of physiotherapy US become very popular tool for the evaluation of exercise effectiveness in low back pain treatment. In majority of cases the thickness of TrA was examined as well as its observed change between contraction and relaxation phase, along with reliability of these measurements [13]. The ML muscle was also subject to US examination. Some of these studies, similarly to TrA examination, focused on the reliability of measurements evaluation during ML activation phase [4, 14, 15, 16, 17, 18], while other analysed ML cross-sectional area (e.g. in subjects with lumbosacral pain syndromes) [18, 14, 19, 20, 21, 22, 23, 24, 25, 26, 27].

The objective of the presented study, regarding measurements of ML thickness, is to define the intra- and inter-rater reliability level (of one or more raters) which justifies the use of presented method for scientific research and clinical trials. The ML cross-sectional area at the level of L4 (both deep and superficial layer of ML), muscle thickness change, and the reliability of repeated analysis of ultrasound images were subject to estimation (carried out by two raters). The effect of the number of repeated measurements on the parameters associated with the reliability in question was attempted to be determined as well.

oraz wiarygodności powtórnej analizy obrazów USG (dwóch badających). Podjęto także próbę określenia wpływu liczby powtarzanych pomiarów na parametry z nią związane.

Materiał i metody

Materiał badawczy

Do udziału w badaniach przystąpiło 15 osób. Na drodze doboru celowego, w oparciu o kryteria selekcji zakwalifikowano 10 osób – 4 kobiety i 6 mężczyzn (wiek 19–26, $X = 23,2 \pm 4,95$; masa ciała 56–92, $X = 71,9 \pm 25,46$ kg; wzrost 164–186, $X = 178,4 \pm 15,56$ cm).

Badane osoby prowadziły podobny, aktywny styl życia oraz prezentowały porównywalny poziom obciążenia wysiłkiem fizycznym.

Kryteria włączenia:

- brak dolegliwości bólowych okolicy lędźwiowo-miedniczno-biodrowej (LMB) w czasie przeprowadzania badania,
- brak dolegliwości bólowych i urazów kompleksu LMB, trwających dłużej niż dwa tygodnie i/lub wymagających interwencji lekarskiej lub hospitalizacji,
- brak przebytych zabiegów chirurgicznych w okolicy LMB i kończyn dolnych,
- prawidłowy obraz ultrasonograficzny ML, umożliwiający dokładne odczytanie danych.

Wszyscy uczestnicy wyrazili zgodę na udział w eksperymencie, jak również zostali poinformowani o jego przebiegu i czynnościach, jakie będą podejmowane w trakcie jego trwania. Badanie zostało zaakceptowane przez Komisję Bioetyczną Akademii Wychowania Fizycznego w Katowicach.

Badający

Pomiary wykonane zostały przez dwoje studentów V roku fizjoterapii (określeni zostali jako badający A i B), posiadających podstawowe wiadomości dotyczące badania wielkości przekroju poprzecznego mięśni z wykorzystaniem USG. Dodatkowo przeszli oni dwutygodniowe szkolenie z zakresu obsługi USG i wykonywania pomiarów tego typu, które nadzorowane było przez doświadczonego, niezależnego specjalistę. Przed przystąpieniem do próby oceniono również, czy poziom umiejętności zdobyty przez badających jest wystarczający do prawidłowego przeprowadzenia pomiarów.

Narzędzie i przygotowanie do pomiaru

Pomiary przekroju poprzecznego ML, wykonane zostały aparatem ultrasonograficznym typu Mindray DP 6600, z płaską głowicą 75L38EA. Częstotliwość, z jaką obrazowano mięsień wielodzielny wynosiła 5 MHz, co pozwoliło na penetrację w głąb tkanek na głębokość około 6 cm.

Pozycja każdego badanego była ustandaryzowana – leżenie przodem z odsłoniętą dolną częścią tułowia

Subjects and methods

Participants

Of the 15 healthy volunteers only 10 of them, including 4 women, who passed the inclusion criteria (age 19–26, $X = 23.2 \pm 4.95$, weight 56–92, $X = 71.9 \pm 25.46$ kg, height 164–186, $X = 178.4 \pm 15.56$ cm) were qualified. All subjects led similar, active lifestyle, and presented a comparable level of daily physical strain load.

Inclusion criteria were:

- lack of low back pain during the measurements, no pain and injuries of lumbo-pelvic complex lasting more than two weeks, and/or requiring any medical intervention or hospitalization,
- lack of surgery in the lumbo-pelvic region and lower limbs,
- appropriate US image quality of ML (allowing precise data readout).

All participants agreed to participate in the experiment and were also informed about its course and procedures undertaken during its duration. The study was approved by the local bioethical committee.

Raters

All measurements were carried out by two raters with some basic experience in acquiring and analyzing US images. Both were students of physiotherapy in Year V and they were described as Rater A and Rater B respectively. Moreover, they went through a two-week training on US maintenance and measuring of such type. This training was guided by an experienced, independent professional who was not directly involved in the study. Prior to the trial, both raters were evaluated whether their level of skills acquired through the training is sufficient enough to carry out US measurements.

Equipment and preparations for measurement

A US imaging device Mindray DP 6600 equipped with the 75L38EA linear array was used for the ML measurements (frequency 5 MHz, tissue penetration depth approx. 6 cm).

Position of each participant was standardized – lying prone with low back exposed, upper limbs placed parallel to the torso. A roll, with the diameter of 15 cm, was placed under the ankles to make the position possibly the most comfortable. Then the rater, palpated L4 spinous process and marked a perpendicular line to the spine at that height. This line became a steady reference point against which the array was applied. Among all subjects the screen magnification was adjusted to 400%. The raters, depending on the results of measurement randomization, performed different tasks. While the rater A was to start the measurement, he was responsible for the correct positioning and holding the array, at the same time, the rater's B task was to record US images. Later the raters swapped their tasks.

i kończynami górnymi ułożonymi wzdłuż ciała. Pod stawy skokowe podkładano okrągły wałek terapeutyczny o średnicy 15 cm, by przyjęta pozycja była jak najbardziej komfortowa. Następnie badający, palpacyjnie odszukiwał wyrostek kolczysty kręgu L4 i na jego wysokości, prostopadle do kręgosłupa zaznaczał linię, stanowiącą stały punkt odniesienia, względem którego przykładano głowicę aparatu USG. U wszystkich osób ustawiono powiększenie obrazu do 400%. Badający, w zależności od wyników losowania kolejności pomiarów, pełnili różne funkcje. W przypadku, gdy badający A rozpoczynał procedurę, do jego zadań należało prawidłowe ustawienie i trzymanie głowicy, w tym samym czasie badający B dokonywał rejestracji obrazu. Jeśli pomiar rozpoczynał badający B, dochodziło do zamiany pełnionych funkcji.

Struktura badania

Ze względu na to, że istotą badań były pomiary przekroju poprzecznego ML, w trakcie jego świadomej i kontrolowanej aktywacji, do procedury wprowadzono tygodniowy trening, którego celem było nauczenie tej umiejętności każdej z zakwalifikowanych osób. Prowadzone w tym okresie ćwiczenia, polegały na nauce świadomego pobudzeniu mięśnia ML. W poniedziałek osoba badana wykonywała 10 powtórzeń 5-sekundowej aktywacji ML w dwóch seriach, przeprowadzonych z wykorzystaniem USG jako biofeedback'u, wszystko to odbywało się pod kontrolą jednego z badaczy. Podobne sesje treningowe wykonywano w środę i piątek. W pozostałe dni osoba badana zobligowana była do przeprowadzania treningu z tą samą ilością serii i powtórzeń w warunkach domowych.

Badania przeprowadzone zostały tydzień po sesji treningowej. Kolejność wykonywania pomiarów przez badających została poddana randomizacji. Dodatkowo, przed przystąpieniem do nich, osoba badana w ramach rozgrzewki wykonywała pięć 5-sekundowych aktywacji ML, by odtworzyć odczucia proprioceptywne, które towarzyszą pobudzeniu tego mięśnia. Każdy badający zapisywał obraz ML w czasie rozluźnienia i aktywacji. Cykl 'aktywacja-rozluźnienie' powtarzany był 5 razy, w wyniku czego jeden badacz otrzymywał 10 obrazów ML w ciągu jednego dnia pomiarowego. Powyżej opisana procedura powtórzona została kolejno w poniedziałek, środę i piątek. Wszystkie badane osoby poinstruowano by w tygodniu, w którym dokonywano pomiarów, utrzymały stały poziom aktywności ruchowej. Dodatkowo zostały one zobowiązane do poinformowania badających o wystąpieniu jakichkolwiek dolegliwości bólowych w okolicy lędźwiowo-miedniczno-biodrowej i/lub kończyn dolnych. Strukturę badania przedstawiono na ryc. 1.

Przetwarzanie danych

Wszystkie obrazy zarejestrowane przez aparat USG zostały zapisane w formacie JPG, a następnie przeniesione do komputera w celu dalszego przetwarzania. Do

Structure

Due to the fact that the measurements encompassed ML conscious controlled activation while the measurement process, the procedure included a weekly training, whose aim was to teach this ability to each of the enrolled subjects. The training consisted of several sessions to train the subjects conscious ML muscle activation. On Monday, the subject performed 10 repetitions of 5-second ML activation in two series, with US as a biofeedback, under the supervision of one of the raters. Similar training sessions were performed on Wednesday and Friday. On other days, the tested person was obliged to train at home following the same number of sets and repetitions.

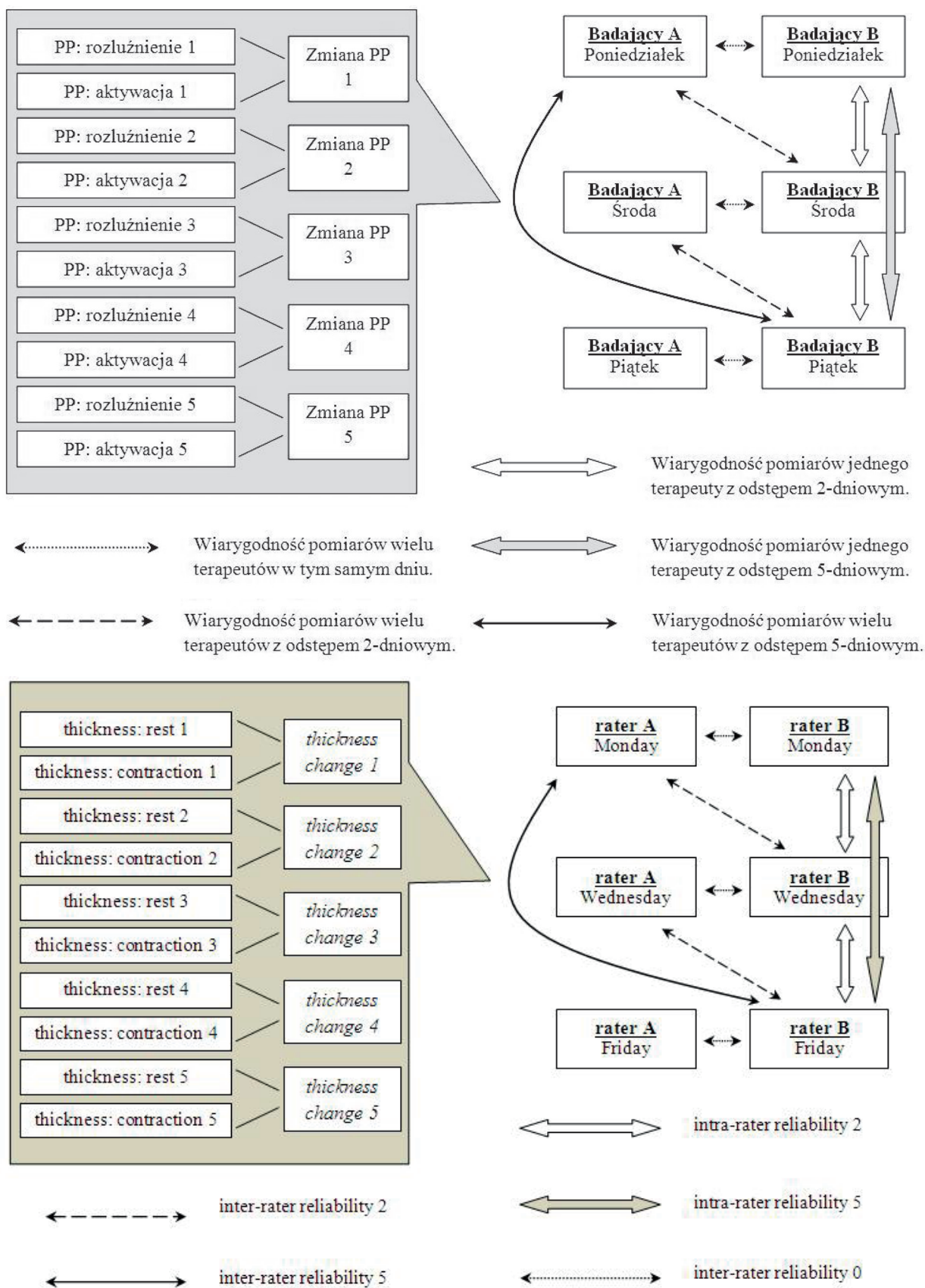
The measurements were carried out a week after the training session. The sequence of measurements performed by raters has been randomized. In addition, prior to them, the examined person, as a part of the warm-up performed five 5-second ML activation to reproduce proprioceptive sensations which accompany each ML contraction. Each rater recorded the ML image during its relaxation and activation phase. The 'activation-relaxation' cycle was repeated 5 times so that one rater was given 10 ML images a day. Such a procedure was repeated successively on Monday, Wednesday and Friday. All subjects were instructed to maintain a constant level of physical activity during the measurement week. In addition, they were obliged to inform the raters in case any pain in the lower back or extremities. Figure 1 presents an outline of the study structure.

Data processing

All images taken by US imaging device were recorded in JPG format and later stored in a computer for further processing. Photoshop 8.0 (Adobe Systems Incorporated, San Jose, USA) program was used. After the brightness and contrast of the image were adjusted, the layers of ML fascia and L4 articular process were selected using Photoshop tools and shaded in white (Fig. 2). This procedure minimised visual distortions typical for the 'frame over the screen' method used previously. The single measurement as well as the average value of two, three, four and five ML measurements, in both activation and relaxation phases, were subjected to the statistical analysis. All images were analysed separately by the two raters. The names of image files were coded by a third person so that none of the raters was able to figure out what phase (activation or relaxation) was the ML muscle in on each analyzed image.

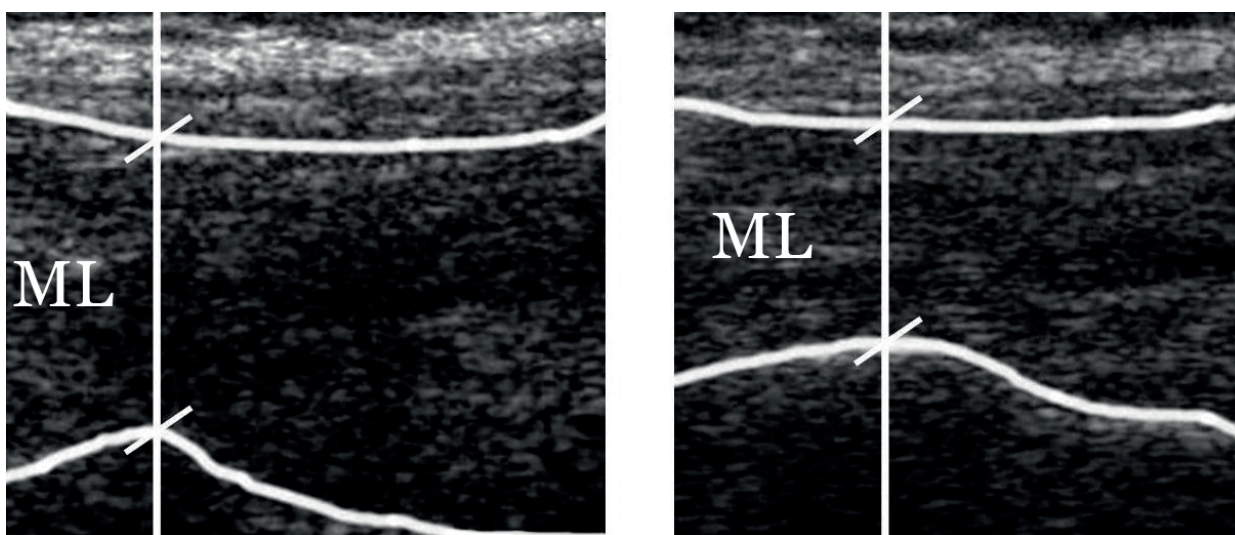
Statistical analysis

In the statistical analysis, we used ANOVA model with repeated measurements. ICC coefficients values were calculated on the basis of this analysis, which enabled further definition of reliability levels (Table 1) [28].



Ryc. 1. Procedura pomiaru (lewa strona ryciny) i struktura obliczania współczynników ICC (prawa strona ryciny), strzałki wskazują różne typy badanej wiarygodności, PP – przekrój poprzeczny mięśnia ML

Fig. 1. Measurement procedure (left-hand section) and reliability calculating structure (right-hand section). All arrows point out different types of investigated reliability, PP – ML cross-section area



Ryc. 2. Łędźwiowa część mięśnia wielodzielnego (ML) w fazie rozluźnienia (prawy obraz) i aktywacji (lewy obraz).

Zaprezentowano obrazy po przetworzeniu w programie Photoshop 8.0. Zastosowane zostały funkcje: rozjaśnienia oraz kontrastu. Na biało zaznaczono powięź ML (górną linię) i wyrostek stawowy kręgu L4 (dolną linię). Pomiaru przekroju poprzecznego ML dokonano wzdłuż białej pionowej linii oznaczonej na każdym z obrazów

Fig. 2. Multifidus lumborum (ML) during relaxation (right picture) and contraction phase (left picture). Presented images are after editing in Photoshop 8.0 software, using contrast and highlighting tools. Upper white line indicates ML fascia, and lower white line – L4 articular process. ML muscle thickness was measured along the vertical white line depicted on the edited image

analizy przetransportowanych obrazów użyto programu Photoshop 8.0 (Adobe Systems Incorporated, San José, USA), w którym zastosowano opcje rozjaśniania i kontrastowania obrazu. Następnie, wykorzystując narzędzia tego samego oprogramowania, oznaczono warstwę powięzi otaczającą mięsień ML oraz wyrostek stawowy kręgu L4 (ryc. 2). Kompletowanie danych w sposób opisany powyżej, minimalizowało błędy, które mogły pojawić się w przypadku zbierania ich bezpośrednio z monitora aparatu USG. Analizie statystycznej, został poddany pojedynczy pomiar jak również średnie dwóch, trzech, czterech oraz pięciu pomiarów ML w fazie aktywacji i rozluźnienia. Każdy obraz był analizowany przez dwóch badaczy. Nazwy wszystkich uzyskanych w badaniu plików JPG zostały zakodowane przez osobę trzecią, w taki sposób, aby żaden z oceniających nie wiedział, w jakiej fazie jest mięsień ML (aktywacji lub rozluźnienia) na analizowanym przez niego obrazie.

Analiza statystyczna

W obliczeniach statystycznych, wykorzystano model analizy wariancji z powtarzającymi pomiarami. W oparciu o wyniki tej analizy, obliczono współczynniki ICC, na podstawie których, określano poziom wiarygodności (tabela 1) [28]. Ocenie poddany został pomiar wielkość przekroju poprzecznego ML w fazie aktywacji i rozluźnienia oraz wielkość różnicy pomiędzy tymi fazami.

By obliczyć współczynniki ICC jednego badającego (intra-rater) dla pojedynczego pomiaru zastosowano model (2,1), natomiast w celu określenia wiarygodności

Differences in ML thickness between activation and relaxation phase were calculated for all subjects.

To calculate the coefficients of a single rater ICC (intra-rater) a single measurement model was used (2,1), and to determine the reliability of the average of two, three, four, and five measurements the following models were used – ICC^{(2,2),(2,3),(2,4),(2,5)}. The reliability of measurements performed with a two-day break, i.e. Monday-Wednesday, Wednesday-Friday (intra-rater 2) and five-day break i.e. Monday to Friday (intra-rater 5) was thus determined. To define ICC values for many raters (inter-rater), analogous calculations were performed (based on the models: (2,1) – a single measurement, (2,2) – the average of two measurements, (2,3) – the average of three measurements, (2,4) – the average of four measurements and (2,5) – the average of five measurements). The reliability of measurements made: on the same day (inter-rater 0), with a two-day break Monday-Wednesday, Wednesday-Friday (inter-rater 2) and a five-day break Monday to Friday (inter-rater 5) was calculated, too- see Fig. 1.

In addition, the standard error of measurement (SEM) and the smallest detectable difference (SDD) were calculated. The calculations were based on the formula: SEM = standard deviation $\times \sqrt{1 - ICC}$, SDD = 1.96 \times SEM $\times \sqrt{2}$.

Results

The measurement of ML thickness showed good and excellent reliability for intra - (Table 2) and inter-rater (Table 3). High ICC values were observed in both

Tabela 1. W tabeli przedstawiono możliwe wartości współczynnika ICC oraz przyporządkowane im poziomy wiarygodności

Table 1. Values of intraclass correlation coefficients (ICCs) and corresponding reliability levels

Wartość współczynnika ICC / ICC values	Poziom wiarygodności / Reliability level
0.00–0.50	słaby / poor
0.50–0.75	umiarkowany / moderate
0.75–0.90	dobry / good
powyżej / above 0.90	doskonały / excellent

średniej dwóch, trzech, czterech i pięciu prób, wykorzystano odpowiednio modele ICC^{(2,2),(2,3),(2,4),(2,5)}. Określona została wiarygodność pomiarów wykonywanych: z dwudniową przerwą, tj. poniedziałek – środa, środa – piątek (intra-rater 2) oraz pięciodniową przerwą, tj. poniedziałek – piątek (intra-rater 5). Dla wyznaczenia ICC wielu badaczy (inter-rater), przeprowadzono analogiczne kalkulacje w oparciu o modele: (2,1) – pojedynczy pomiar, (2,2) – średnia dwóch pomiarów, (2,3) – średnia trzech pomiarów, (2,4) – średnia czterech pomiarów i (2,5) – średnia pięciu pomiarów. Obliczona została wiarygodność pomiarów wykonanych: tego samego dnia (inter-rater 0), z dwudniową przerwą, tj. poniedziałek – środa, środa – piątek (inter-rater 2) oraz pięciodniową przerwą tj. poniedziałek – piątek (inter-rater 5). Zob. ryc. 1.

Ponadto, obliczony został standardowy błąd pomiaru (SEM – *standard error of measurement*) oraz najmniejsza wykrywalna różnica (SDD – *smallest detectable differences*). Obliczeń dokonano w oparciu o formuły: SEM = odchylenie standardowe $\times \sqrt{1 - ICC}$, SDD = 1.96 \times SEM $\times \sqrt{2}$.

Wyniki

Pomiar wielkości przekroju poprzecznego mięśnia ML wykazał dobrą i doskonałą wiarygodność intra (tabela 2) i inter-rater (tabela 3). Wysokie współczynniki ICC zaobserwowano zarówno w pomiarach przeprowadzonych z dwu- bądź pięciodniową przerwą. Wszystkie uzyskane współczynniki ICC osiągnęły wartości przekraczające 0.81. Najwyższy SEM wynosi 2.54 mm, natomiast SDD 7.03 mm. Obie wartości zostały uzyskane w przypadku wykonania pojedynczego pomiaru, w odstępie 5 dni (poniedziałek – piątek). Niezależnie od tego czy brano pod uwagę średnią z dwóch, czy z pięciu pomiarów, wiarygodność intra-rater i inter-rater odnotowana została na poziomie doskonałym. Wyjątkiem są wartości uzyskane ze średniej czterech pomiarów wykonywanych z pięciodniowym odstępem czasu, które prezentują poziom dobry. Uzyskane wyniki wskazują, że ilość pomiarów przekroju poprzecznego mięśnia ML, nie wpływa znacząco na poziom wiarygodności.

Inaczej przedstawiają się dane dotyczące zmiany wielkości przekroju poprzecznego ML, które są zależne od ilości wykonanych pomiarów. Dla jednego badającego, współczynnik ICC wzrasta od 0.22 (intra-rater 2 dla

measurements (with two or five-day breaks respectively). The observed ICCs were 0.81 or higher. The highest SEM was 2.54 mm, and the highest SDD was 7.03 mm. Both values were obtained for the implementation of a single measurement, with a five-day break (Monday – Friday). Regardless whether the average two or five measurements was taken into account, the reliability of intra-rater and inter-rater reached the excellent level. The exceptions are the values obtained from the average of four measurements carried out with a five-day break, which presented a good level. The results indicated that the number of thickness measurements of the ML muscle, does not significantly affect the level of reliability.

Different results are observed concerning ML thickness change, in which the reliability level is dependent on the number of taken measurements. For one rater, the ICC coefficient increased from 0.22 (intra-rater 2) to 0.82 (intra-rater 5) (Table 2). A similar trend can be observed for measurements made by two raters, ICCs from 0.30 (inter-rater 2) to 0.86 (inter-rater 0) (Table 3). The lowest ICC values were observed when a single measurement was taken, in such cases, SEM ranged from 2.07 to 4.77%, and the SDD from 5.73 to 13.22%.

Discussion

Our results indicated that the most decisive factor that had a significant impact on the measuring reliability was the number of repeated measurements. Considering the reliability results of ML thickness measurements, regardless of the time lag between measurement day and the number of raters, it can be concluded that even a single measurement was enough to get at least a satisfactory outcome. Interestingly, adding another repetition (measurement), increases ICC coefficients to the excellent level. It was also noticed that increasing the number of measurements to five had no significant effect on improving the reliability level. Consequently, it was recommended to carry out at least two repetitions to achieve the excellent reliability of measurement of ML cross-sectional area. Similar recommendations have been proposed by Koppenhaver [16], in studies with congruent methodology. Comparing measurements from a 3-day time interval, he obtained results comparable to those presented in this paper.

Another dependence can be observed while assessing the reliability of measuring ML thickness

Tabela 2. Pomiar wiarygodności intra-rater (jeden badający). Zaprezentowano wyniki: współczynnik korelacji wewnątrz klasowej (ICC ($\pm 95\%$ przedziały ufności)); modele: (2,1) dla pojedynczego, (2,2) dla podwójnego oraz (2,3) dla potrójnego, (2,4) dla czterokrotnego i (2,5) dla pięciokrotnego pomiaru. PP (przekrój poprzeczny), SEM (standardowy błąd pomiaru) oraz SDD (najmniejsza wykrywalna różnica). Wiarygodność intra-rater 2 określa wiarygodność pomiaru wykonywanego z dwudniowym odstępem czasu (t. j. poniedziałek – środa, środa – piątek); wiarygodność intra-rater 5 – wiarygodność pomiaru wykonywanego z pięciodniowym odstępem czasu (tj. poniedziałek–piątek) (zob. ryc. 1)

Table 2. Measures of intra-rater reliability. Presented are interclass correlation coefficient (ICC ($\pm 95\%$ Confidence Interval)); models: (2,1) for single measurements (2,2) for two repeated measurements, (2,3) for three repeated measurements, (2,4) for four repeated measurements and (2,5) for five repeated measurements. Standard error of the measurement (SEM) and the smallest detectable difference (SDD). Intra-rater reliability 2 describes reliability of measurements performed with 2-day break (i.e. Monday – Wednesday, Wednesday – Friday); intra-rater reliability 5 – reliability of measurements performed with 5-day break (i.e. Monday – Friday) (see Fig. 1)

	Liczba pomiarów / N measurements	PP / Thickness			Zmiana w PP / Thickness change		
		ICC ($\pm 95\%$ PU)	SEM (mm)	SDD (mm)	ICC ($\pm 95\%$ PU)	SEM (%)	SDD (%)
intra-rater 2	1	0.89(0.83-0.93)	2,02	5,61	0.22(-0.10-0.49)	4,77	13,22
	2(średnia / mean)	0.95(0.92-0.97)	1,36	3,77	0.55(0.16-0.76)	3,33	9,24
	3(średnia / mean)	0.95(0.93-0.97)	1,36	3,78	0.66(0.35-0.82)	2,80	7,76
	4(średnia / mean)	0.93(0.89-0.96)	1,63	4,53	0.67(0.37-0.82)	2,76	7,66
	5(średnia / mean)	0.96(0.93-0.97)	1,20	3,34	0.69(0.42-0.84)	2,64	7,33
intra-rater 5	1	0.82(0.81-0.95)	2,47	6,84	0.48(0.07-0.76)	4,36	12,09
	2(średnia / mean)	0.93(0.86-0.96)	1,54	4,26	0.74(0.34-0.89)	2,93	8,11
	3(średnia / mean)	0.92(0.86-0.96)	1,63	4,53	0.63(0.08-0.85)	3,37	9,33
	4(średnia / mean)	0.86(0.75-0.93)	2,25	6,23	0.76(0.41-0.91)	2,73	7,55
	5(średnia / mean)	0.93(0.86-0.96)	1,54	4,26	0.82(0.54-0.93)	2,38	6,59

pojedynczego pomiaru) do 0,82 (intra-rater 5 – średnia 5 pomiarów przeprowadzonych z pięciodniowym odstępem czasu) (tabela 2). Podobną tendencję można zaobserwować dla pomiarów wykonanych przez dwie osoby, gdzie odnotowano wartości od 0.30 (inter-rater 2 dla pojedynczego pomiaru) do 0.86 (inter-rater 0 – średnia 5 pomiarów wykonywanych tego samego dnia) (tabela 3). Najniższe ICC zaobserwowano, gdy wykonany został pojedynczy pomiar, w tych przypadkach wielkości SEM wynosiły od 2.07 do 4.77%, a SDD od 5.73 do 13.22%.

Dyskusja

Zdaniem autorów niniejszej pracy, decydującym czynnikiem, który wywiera znaczący wpływ na wiarygodność pomiarów o takim charakterze, jest liczba powtarzanych pomiarów. Rozpatrując wyniki wiarygodności pomiarów przekroju poprzecznego ML, niezależnie od odstępu czasu pomiędzy wykonywanym badaniem oraz od ilości badaczy, można dojść do wniosku, że już pojedynczy pomiar wystarczy, by uzyskać wynik na poziomie przynajmniej dobrym. Co ciekawe, wykonanie drugiego powtórzenia, powoduje wzrost wartości współczynników ICC do poziomu doskonałego. Zauważono również, że zwiększenie ilości pomiarów do pięciu, nie wywiera znaczącego wpły-

change which varies depending on the time interval between measurements, the number of repetitions and the number of raters. During the assessment of the obtained results, the tendency to increase the level of reliability with the increased number of repetitions could be observed. After 5 measurements, (regardless whether done by one or two raters), it possible to achieve a level ranging from moderate to good (depending on the time interval between tests). For this reason such a number of trials is recommended in research aiming at evaluating ML thickness change.

It should also be pointed that according to the authors, there are factors that may influence the reliability of measurement of these tested variables. One of them is the degree of experience of the individual rater. Based on the obtained results, which were similar to the results of other research reports [18], it can be stated that the high reliability of measurement of ML thickness can be achieved by raters who despite their little experience have went through appropriate training. In the framework of the disputed topic, we should put forward the issue of the lumbar spine palpation level, where the measurements were made. Better results perhaps, could be achieved, in the case of measurements carried out by some individuals

Tabela 3. Pomiar wiarygodności inter-rater (dwóch badających). Zaprezentowano wyniki: współczynnik korelacji wewnątrzklasowej (ICC ($\pm 95\%$ przedziały ufności)); modele: (2,1) dla pojedynczego, (2,2) dla podwójnego oraz (2,3) dla potrójnego, (2,4) dla czterokrotnego i (2,5) dla pięciokrotnego pomiaru. PP (przekrój poprzeczny), SEM (standardowy błąd pomiaru) oraz SDD (najmniejsza wykrywalna różnica). Inter-rater 0 opisuje wiarygodność pomiaru wykonywanego tego samego dnia. Wiarygodność inter-rater 2 – pomiar wykonany z dwudniowym odstępem czasu (tj. poniedziałek – środa, środa – piątek); wiarygodność inter-rater 5 – pomiar wykonany z pięciodniowym odstępem czasu (tj. poniedziałek – piątek) (zob. ryc. 1)

Table 3. Measures of inter-rater reliability. Presented are interclass correlation coefficient (ICC ($\pm 95\%$ Confidence Interval)); models: (2,1) for single measurements, (2,2) for two repeated measurements, (2,3) for three repeated measurements, (2,4) for four repeated measurements and (2,5) for five repeated measurements. Standard error of the measurement (SEM) and the smallest detectable difference (SDD). Inter-rater reliability 0 describes reliability of measurements performed on the same day; inter-rater reliability 2 – reliability of measurements performed with 2-day break (i.e. Monday – Wednesday, Wednesday – Friday); intra-rater reliability 5 – reliability of measurements performed with 5-day break (i.e. Monday – Friday) (see Fig.1)

Wiarygodność Reliability	Liczba pomiarów / N measurements	PP / Thickness			Zmiana w PP / Thickness change		
		ICC ($\pm 95\%$ CI)	SEM (mm)	SDD (mm)	ICC ($\pm 95\%$ PU)	SEM (%)	SDD (%)
inter-rater 0	1	0.96 (0.94-0.98)	1.20	3.33	0.50 (0.17-0.72)	3.98	11.02
	2(średnia / mean)	0.99(0.98-0.99)	0.60	1.66	0.80(0.57-0.90)	2.34	6.49
	3(średnia / mean)	0.99(0.98-0.99)	0.60	1.66	0.82(0.62-0.91)	2.15	5.95
	4(średnia / mean)	0.92(0.87-0.95)	1.73	4.80	0.86(0.70-0.93)	1.90	5.26
	5(średnia / mean)	0.99(0.98-0.99)	0.60	1.65	0.86(0.71-0.93)	1.89	5.24
inter-rater 2	1	0.88(0.83-0.92)	2.11	5.86	0.30(-0.01-0.55)	2.07	5.73
	2(średnia / mean)	0.95(0.93-0.97)	1.36	3.77	0.59(0.24-0.79)	3.18	8.82
	3(średnia / mean)	0.96(0.93-0.97)	1.22	3.38	0.68(0.40-0.83)	2.72	7.53
	4(średnia / mean)	0.92(0.87-0.95)	1.75	4.84	0.69(0.41-0.84)	2.68	7.43
	5(średnia / mean)	0.96(0.93-0.97)	1.20	3.34	0.69(0.42-0.84)	2.64	7.33
inter-rater 5	1	0.81(0.67-0.90)	2.54	7.03	0.39(-0.79-0.68)	4.72	13.09
	2 (średnia)	0.92(0.85-0.96)	1.64	4.55	0.61(0.02-0.84)	3.58	9.93
	3 (średnia)	0.92(0.85-0.96)	1.63	4.53	0.58(-0.48-0.83)	3.59	9.94
	4 (średnia)	0.85(0.72-0.92)	2.33	6.45	0.72(0.32-0.89)	2.94	8.16
	5 (średnia)	0.93(0.86-0.96)	1.54	4.26	0.79(0.49-0.92)	2.57	7.12

wu na poprawę wiarygodności. W związku z tym, autorzy rekomendują wykonać co najmniej dwóch powtórzeń, by osiągnąć doskonałą wiarygodność pomiaru przekroju poprzecznego ML. Podobne zalecenia, zostały zaproponowane przez Koppenhaver'a [16], który w badaniach wykorzystujących zbliżoną metodologię, porównywał pomiary przeprowadzane z 3-dniowym odstępem czasu. Uzyskał on wyniki, mieszczące się w zakresach porównywalnych do przedstawionych w prezentowanej pracy.

Inną zależność można zaobserwować oceniając wiarygodność pomiaru zmiany przekroju poprzecznego. Jest ona zmienna zależnie od odstępu czasu pomiędzy pomiarami,

with previous clinical experience. According to the authors, the training time extension, may also influence the reliability of measurement of ML thickness. Another controversy arose when the shape of the US array used in the study was concerned. Recent studies provided evidence that the reliability of the measurements was not affected by the shape of the array[29]. However, it is highly probable that the array width that clings to the body of the person examined at the measuring point may have the impact on the results.

liczby powtórzeń oraz ilości badaczy przeprowadzających badanie. Dokonując jednak oceny przedstawionych wyników, uchwytana staje się tendencja wzrostu poziomu wiarygodności wraz ze zwiększeniem liczby powtórzeń. Po wykonaniu 5 pomiarów, niezależnie czy przez jednego czy dwóch badających, możliwe jest osiągnięcie poziomu umiarkowanego i dobrego (w zależności od odstępu czasu pomiędzy badaniami). Z tego powodu, taką liczbę prób rekomenduje się w badaniach mających na celu ocenę zmiany wielkości przekroju poprzecznego ML.

Należy również zaznaczyć, iż w opinii autorów istnieją czynniki, które mogą mieć wpływ na zwiększenie wiarygodności pomiaru badanych zmiennych. Pierwszym z nich jest stopień doświadczenia poszczególnych osób dokonujących pomiarów. Na podstawie przedstawionych wyników, które są zbliżone do rezultatów innych doniesień naukowych [18], stwierdzić można, że wysoką wiarygodność pomiarów przekroju poprzecznego ML mogą osiągnąć osoby, które mimo niewielkiego doświadczenia, zostały poddane odpowiedniemu szkoleniu. W ramy obszaru spornego należy umieścić zagadnienie palpacji poziomu kręgosłupa lędźwiowego, na którym dokonuje się pomiarów. Być może lepsze wyniki, uzyskałoby się, w przypadku pomiarów przeprowadzonych przez osoby posiadające pewne doświadczenie kliniczne. Według autorów, wydłużenie czasu przeszkolenia badających, również może mieć wpływ na wiarygodność pomiaru przekroju poprzecznego ML. Kolejną kwestią, która przybiera charakter kontrowersyjny, jest kształt głowic aparatów ultrasonograficznych używanych w badaniach. Najnowsze badania dostarczają dowodów, iż na wiarygodność pomiarów, nie wpływa kształt używanej głowicy [29]. Jednakże, można spekulować, że wpływ na wyniki może mieć szerokość głowicy, która przylega do ciała osoby badanej w miejscu pomiaru.

Wnioski

1. By osiągnąć akceptowalny poziom wiarygodności pomiaru mięśnia ML na poziomie L4 (zarówno dla jednego jak i wielu badaczy), należy wykonać od dwóch (dla przekroju poprzecznego) do pięciu (dla badania zmiany przekroju poprzecznego) powtórzeń omawianych pomiarów.
2. Ocena wiarygodności pomiarów, przekroju poprzecznego ML, u osób nieposiadających dolegliwości bólowych w rejonie kompleksu biodrowo-miedniczno-lędźwiowego, może być egzekwowana przez niedoświadczone osoby, które odbyły odpowiednie przeszkolenie.

Conclusions

1. To achieve an acceptable level of the measurement reliability of ML cross-sectional area at L4 level (both for single and multiple raters), it is necessary to perform from two to five repetitions of these measurements (for the thickness and for the thickness change respectively).
2. Evaluation of the reliability of ML measurements, in people who do not suffer from pain in the lumbo-pelvic complex, can be carried out by inexperienced people who have been properly trained.

Piśmiennictwo / References

1. Stokes M, Young A. *Measurement of quadriceps cross-sectional area by ultrasonography: a description of the technique and its application in physiotherapy*. Physiother Theor Pract, 1986; 2:31-36.
2. Loo A, Stokes J. *Diagnostic ultrasound scanning for clinical estimation of quadriceps size and estimation of strength*. In: *Proceedings of the III-rd International Physiotherapy Congress*. Hong Kong 1990, pp: 655-660.

3. Martinson H, Stokes M. *Measurement of anterior tibial muscle size using real-time ultrasound imaging*. Eur J Physiol. 1991; 63: 250-254.
4. Hides J, Cooper D, Stokes J. *Diagnostic ultrasound imaging for measurement of the lumbar multifidus muscle in normal young adults*. Physiother Theor Pract, 1992; 8: 19-26.
5. Hides J, Stokes M, Saide M, Jull G, Cooper D. *Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain*. Spine, 1994; 19: 165-172.
6. Hides J, Richardson C, Jull G. *Magnetic resonance imaging and ultrasonography of the lumbar multifidus muscle: Comparison of two different modalities*. Spine, 1995; 20: 54-58.
7. Hides J, Richardson C, Jull G. *Multifidus muscle recovery is not automatic following resolution of acute first episode low back pain*. Spine, 1996; 21: 2763-2769.
8. Sipilä S, Suominen H. *Muscle ultrasonography and computer tomography in elderly trained and untrained women*. Muscle Nerve, 1993; 16: 294-300.
9. Cresswell A, Grundstrom A, Thorstensson A. *Observations on intra-abdominal pressure and patterns of abdominal intra-muscular activity in man*. Acta Physiol Scand. 1992; 144: 409-418.
10. Hodges P, Richardson C. *Relationship between limb movement speed and associated contraction of the trunk muscles*. Ergonomics, 1997; 40: 1220-1230.
11. Damen L, Stijnen T, Roebroeck M, Snijders C, Stam H. *Reliability of sacroiliac joint laxity measurement with Doppler imaging of vibrations*. Ultrasound Med Biol. 2002; 28: 407-414.
12. Richardson C, Snijders C, Hides J, Damen L, Pas M, Storm J. *The relation between the transverses abdominis muscles, sacroiliac joint mechanics and low back pain*. Spine, 2002; 27: 399-405.
13. Gnat R, Saulicz E, Miądowicz B. *Reliability of real-time ultrasound measurement of transversus abdominis thickness in healthy trained subjects*. Eur Spine J, Epub. Feb. 2012.
14. Brenner A, Gill N, Buscema C, Kiesel K. *Improved activation of lumbar multifidus following spinal manipulation: a case report applying rehabilitative ultrasound imaging*. J Orthop Sports Phys Ther. 2007; 37(10): 613-619.
15. Wallwork T, Hides J, Stanton W. *Intrarater and interrater reliability of assessment of lumbar multifidus muscle thickness using rehabilitative ultrasound imaging*. J Orthop Sports Phys Ther. 2007; 37(10): 608-612.
16. Koppenhaver S, Hebert J, Fritz J, Parent E, Teyhen D, Magel J. *Reliability of rehabilitative ultrasound imaging of the transversus abdominis and lumbar multifidus muscles*. Arch Phys Med Rehabil. 2009; 90(1): 87-94.
17. Herbert W, Heiss D, Basso D. *Influence of feedback schedule in motor performance and learning of a lumbar multifidus muscle task using rehabilitative ultrasound imaging: a randomized clinical trial*. Phys Ther. 2008; 88(2): 261-269.
18. Koppenhaver S, Parent E, Teyhen D, Hebert J, Fritz J. *The effect of averaging multiple trials on measurement error during ultrasound imaging of transversus abdominis and lumbar multifidus muscles in individuals with low back pain*. J Orthop Sports Phys Ther. 2009; 39(8): 604-611.
19. Kiesel K, Uhl T, Underwood F, Nitz A. *Rehabilitative ultrasound measurement of select trunk muscle activation during induced pain*. Man Ther. 2008; 13(2): 132-138.
20. Kiesel K, Underwood F, Mattacola C, Nitz A, Malone T. *A comparison of select trunk muscle thickness change between subjects with low back pain classified in the treatment-based classification system and asymptomatic controls*. J Orthop Sports Phys Ther. 2007; 37(10): 596-607.
21. Pressler J, Heiss D, Buford J, Chidley J. *Between-day repeatability and symmetry of multifidus cross-sectional area measured using ultrasound imaging*. J Orthop Sports Phys Ther. 2006; 36(1): 10-18.
22. Herbert W, Heiss D, Basso D. *Influence of feedback schedule in motor performance and learning of a lumbar multifidus muscle task using rehabilitative ultrasound imaging: a randomized clinical trial*. Phys Ther. 2008; 88(2): 261-269.
23. Van K, Hides J, Richardson C. *The use of real-time ultrasound imaging for biofeedback of lumbar multifidus muscle contraction in healthy subjects*. J Orthop Sports Phys Ther. 2006; 36(12): 920-925.
24. Hodges P, Holm A, Hansson T, Holm S. *Rapid atrophy of the lumbar multifidus follows experimental disc or nerve root injury*. Spine, 2006; 31(25): 2926-2933.
25. Vasseljen O, Dahl H, Mork P, Torp H. *Muscle activity onset in the lumbar multifidus muscle recorded simultaneously by ultrasound imaging and intramuscular electromyography*. Clin Biomech. 2006; 21(9): 905-913.
26. Stokes M, Rankin G, Newham D. *Ultrasound imaging of lumbar multifidus muscle: normal reference ranges for measurements and practical guidance on the technique*. Man Ther. 2005; 10(2): 116-126.
27. Coldron Y, Stokes M, Cook K. *Lumbar multifidus muscle size does not differ whether ultrasound imaging is performed in prone or side lying*. Man Ther. 2003; 8(3): 161-165.
28. Portney L, Watkins M. *Statistical measures of reliability*. W: Portney L, Watkins M. (red.). Foundations of clinical research applications to practice. Prentice-Hall, New Jersey 2000; 557-584.
29. Worsley P, Smith N, Warner M, Stokes M. *Ultrasound transducer shape has no effect on measurements of lumbar multifidus muscle size*. Man Ther. 2012; 17(2): 187-191.

Adres do korespondencji / Mailing address:

Maciej Biały
 Centrum Rehabilitacji
 ul. Kilińskiego 56/VIII, 41-200 Sosnowiec
 Tel: 782 668 617
 e-mail: mbfizjoterapia@gmail.com