

Submitted: 19.11.2015
Accepted: 06.04.2016

Inter-rater reliability in the classification of supraspinatus tendon tears using 3D ultrasound – a question of experience?

Zgodność oceny stopnia uszkodzenia ścięgna mięśnia nadgrzebieniowego w ultrasonografii 3D – kwestia doświadczenia?

Giorgio Tamborrini¹, Christian Marx¹, Raphael Micheroli²

¹ *Ultrasound Center and Rheumatology Department, Bethesda Hospital Basel, Switzerland*

² *University Hospital Zürich, Rheumatology Department, Switzerland*

Correspondence: Giorgio Tamborrini, MD, Ultrasound Center Rheumatology, Gellerstrasse 144, Switzerland – 4002 Basel, e-mail: sono@irheuma.com

DOI: 10.15557/JoU.2016.0025

Key words

ultrasound,
inter-rater reliability,
shoulder rotator cuff,
3D ultrasound,
reproducibility

Słowa kluczowe

ultrasonografia,
zgodność ocen,
pierścień rotatorów,
ultrasonografia 3D,
odtwarzalność

Abstract

Background: Three-dimensional (3D) ultrasound of the shoulder is characterized by a comparable accuracy to two-dimensional (2D) ultrasound. No studies investigating 2D versus 3D inter-rater reliability in the detection of supraspinatus tendon tears taking into account the level of experience of the raters have been carried out so far. **Objectives:** The aim of this study was to determine the inter-rater reliability in the analysis of 3D ultrasound image sets of the supraspinatus tendon between sonographer with different levels of experience. **Patients and methods:** Non-interventional, prospective, observational pilot study of 2309 images of 127 adult patients suffering from unilateral shoulder pain. 3D ultrasound image sets were scored by three raters independently. The intra- and inter-rater reliabilities were calculated. **Results:** There was an excellent intra-rater reliability of rater A in the overall classification of supraspinatus tendon tears (2D vs 3D $\kappa = 0.892$, pairwise reliability 93.81%, 3D scoring round 1 vs 3D scoring round 2 $\kappa = 0.875$, pairwise reliability 92.857%). The inter-rater reliability was only moderate compared to rater B on 3D ($\kappa = 0.497$, pairwise reliability 70.95%) and fair compared to rater C ($\kappa = 0.238$, pairwise reliability 42.38%). **Conclusions:** The reliability of 3D ultrasound of the supraspinatus tendon depends on the level of experience of the sonographer. Experience in 2D ultrasound does not seem to be sufficient for the analysis of 3D ultrasound imaging sets. Therefore, for a 3D ultrasound analysis new diagnostic criteria have to be established and taught even to experienced 2D sonographers to improve reproducibility.

Streszczenie

Wprowadzenie: Ultrasonografia trójwymiarowa (3D) barku jest uznawana za metodę o porównywalnej dokładności w stosunku do ultrasonografii dwuwymiarowej (2D). Jak dotąd nie przeprowadzono żadnego badania porównującego ultrasonografię 2D i 3D w diagnostyce zerwania ścięgna mięśnia nadgrzebieniowego z uwzględnieniem poziomu doświadczenia badaczy. **Cele:** Celem niniejszego badania było porównanie oceny patologii ścięgna nadgrzebieniowego za pomocą badania 3D pomiędzy operatorami o różnym poziomie doświadczenia. **Pacjenci i metody:** Pilotażowe nieinterwencyjne, prospektywne, obserwacyjne badanie 2309 obrazów uzyskanych z badań 127 dorosłych pacjentów cierpiących z powodu jednostronnego bólu barku. Zestawy obrazów 3D zostały ocenione w sposób niezależny przez trzech badaczy. Obliczono poziom zgodności pomiędzy badaczami oraz

zgodność ocen jednego badacza. **Wyniki:** Badacz A wykazał się bardzo wysokim poziomem zgodności swoich ocen w ogólnej klasyfikacji zerwania ścięgna mięśnia nadgrzebieniowego (2D vs 3D $\kappa = 0,892$, porównywalność parami 93,81%, runda ocen obrazów 3D nr 1 vs runda ocen obrazów 3D nr 2 $\kappa = 0,875$, porównywalność parami 92,857%). Zgodność pomiędzy badaczami była tylko umiarkowana w porównaniu z badaczem B (obrazy 3D) ($\kappa = 0,497$, porównywalność parami 70,95%) i przeciętna w porównaniu z badaczem C ($\kappa = 0,238$, porównywalność parami 42,38%). **Wnioski:** Wiarygodność interpretacji obrazów 3D ścięgna mięśnia nadgrzebieniowego jest zależna od poziomu doświadczenia operatora. Doświadczenie w ultrasonografii 2D zdaje się być niewystarczające do analizy zestawów obrazów 3D. Z tego powodu należy opracować nowe kryteria diagnostyczne dla analizy obrazów 3D i wprowadzić nauczanie interpretacji badań ultrasonografii trójwymiarowej, aby w ten sposób zwiększyć poziom odtwarzalności.

Background

Two-dimensional (2D) ultrasound (US) is an inexpensive, painless, non-invasive, and fast method of a dynamic examination of the shoulder in real-time. It is characterized by a similar accuracy to the MRI examination to evaluate superficial tendons and soft tissues with a high degree of resolution, including the rotator cuff and muscles⁽¹⁻⁵⁾. Furthermore, US is not limited by MRI contraindications, such as the patient's body habitus, claustrophobia, inability to lie flat, or implanted devices not compatible with the magnetized environment of MRI. Yet, the major disadvantage of 2D US, including the sonography of the shoulder, is a high rater dependence^(6,7).

Three-dimensional (3D) US with the advantages of multiplanar imaging facility with three perpendicular planes observed in any position, plus direction, has been an established prenatal diagnostic technique since the early 90's⁽⁸⁾. The use of this technique in musculoskeletal diagnostics has been reported in the examination of the neonatal hip and in the diagnosis of meniscus lesions^(9,10). However, the role of 3D US in the diagnosis of shoulder pathologies is not clear with only few studies, and on small populations, investigating the rater dependency in 3D US⁽¹¹⁻¹³⁾.

Objectives

The objective of this study was to determine the inter-rater reliability in the analysis of 3D US image sets of the supraspinatus tendon between sonographers with different levels of experience in musculoskeletal 2D US.

Patients and methods

We performed a non-interventional, prospective, observational pilot study involving 127 patients who were referred for the sonographic examination of a painful shoulder with clinical suspected or considered supraspinatus tendon pathology by the referring doctor. They were recruited in the period between January, 2014 and July, 2014 and were enrolled in a single outpatient clinic after signing an informed consent. The approval of the local ethics review board Basel was obtained. The study was performed in

Wprowadzenie

Ultrasonografia dwuwymiarowa (USG 2D) to niedroga, bezbolesna, nieinwazyjna i szybka metoda dynamicznego badania barku w czasie rzeczywistym. Metoda ta charakteryzuje się dokładnością zbliżoną do badania rezonansu magnetycznego (MRI) w ocenie powierzchniowych ścięgien i tkanek miękkich, w tym pierścienia rotatorów i mięśni⁽¹⁻⁵⁾. Ponadto diagnostyka USG nie jest ograniczona przez przeciwwskazania typowe dla MRI, takie jak rozmiar ciała, klaustrofobia, niezdolność do utrzymania pozycji leżącej na czas badania lub obecność implantów, które nie są kompatybilne z namagnetyzowanym otoczeniem MRI. Jednakże główną wadą ultrasonografii dwuwymiarowej, w tym również ultrasonografii barku, jest wysoki stopień zależności od operatora^(6,7).

Trójwymiarowe (3D) badanie USG, polegające na wielopłaszczyznowym obrazowaniu w trzech prostopadłych płaszczyznach, uzyskiwanych z dowolnej pozycji, to dobrze ugruntowana technika diagnostyki prenatalnej, stosowana od wczesnych lat 90.⁽⁸⁾ Metoda ta została wykorzystana w diagnostyce układu mięśniowo-szkieletowego w badaniu biodra noworodka oraz w diagnozie patologii łokotki^(9,10). Mimo to rola ultrasonografii trójwymiarowej w diagnostyce patologii barku nie została jasno określona. Powstało na ten temat tylko kilka badań, przeprowadzonych na niewielkich populacjach, dotyczących zależności oceny od operatora⁽¹¹⁻¹³⁾.

Cele

Celem niniejszego badania było porównanie oceny patologii ścięgna nadgrzebieniowego w badaniu trójwymiarowym pomiędzy operatorami o różnym poziomie doświadczenia w wykonywaniu badań ultrasonograficznych układu mięśniowo-szkieletowego 2D.

Pacjenci i metody

Autorzy niniejszej pracy wykonali pilotażowe, nieinterwencyjne, prospektywne, obserwacyjne badanie, obejmujące 127 pacjentów, którzy zostali skierowani przez lekarza zlecającego na badanie ultradźwiękowe bolesnego barku z podejrzeniem lub rozpoznaniem patologii ścięgna mięśnia nadgrzebieniowego. Pacjenci byli rekrutowani w okresie od stycznia do czerwca 2014 roku, a następnie, po podpisaniu formularza świadomości

2D US findings	Diagnostic criteria
Tendinosis	Thick swollen tendon with hypoechoic echotexture. Partial interruption may occur inside the tendon. Irregularity of fibrillar pattern, fragmentation, and focal hypoechoic or hyperechoic areas
Bursal sided partial-thickness tear	Focal hypoechoic or anechoic defect in the tendon, involving the bursal surface
Articular sided partial-thickness tear	Focal hypoechoic or anechoic defect in the tendon, involving the articular surface
Intra-substance sided partial-thickness tear	Focal hypoechoic or anechoic defect inside the tendon
Full thickness tear	Hypoechoic zone extending through the entire substance of the cuff or non-visualization of the rotator cuff with "approximation" of the deltoid muscle to the surface of the humeral head

Tab. 1. Diagnostic criteria for the supraspinatus tendon in 2D US

accordance with good clinical practice and carried out in compliance with the Helsinki Declaration.

Dynamic 2D US examination of both shoulders was conducted in accordance with the guidelines set by the Swiss Society of US Medicine (Schweizerische Gesellschaft für Ultraschall in der Medizin, SGUM), musculoskeletal section⁽¹⁴⁾ which are compatible with the Musculoskeletal US Technical Guidelines from the European Society of Musculoskeletal Radiology (ESSR)⁽¹⁵⁾ and the guidelines of the European League Against Rheumatism (EULAR).

2D US examination was performed by an advanced sonographer (EULAR Teacher/EFSUMB Level 3 = rater A). GE Logic E9 with ML6-15 probe was used and the scanning time for dynamic examination was 10 minutes.

Diagnostic criteria for 2D US findings of the supraspinatus tendon are listed in Tab. 1.

A volumetric 3D image set (6–16D probe on GE Logic E9 and 6–16D probe on GE Voluson E6, scanning time 2 minutes) with the visualization of the long biceps tendon at the level of the rotator cuff interval and the footprint of the supraspinatus tendon was acquired during the same examination. The arm was held in a modified Crass position. Afterwards, a set of transverse, longitudinal, axial and rendered volumetric 3D images of the supraspinatus tendon insertion was created automatically.

Based on the assumption that the accuracy of ultrasound is comparable to MRI in the evaluation of the rotator cuff lesions^(1–5), we decided not to perform additional MRI examination. We selected 3 raters with different experience level in musculoskeletal ultrasound.

Ustalenia na podstawie badania USG 2D	Kryteria diagnostyczne
Przewlekłe uszkodzenie ścięgna (<i>tendinosis</i>)	Ścięgno pogrubiałe, obrzęknięte, hipoechogeniczne; możliwe śródścięgniste uszkodzenia; echostruktura niejednorodna albo zachowana budowa włókienkowa, fragmentacja, ogniska hipoechogeniczne lub hiperechogeniczne
Zerwanie częściowe po stronie kaletki	Ognisko hipoechogeniczne lub bezechowe w obrębie ścięgna, obejmujące włókna obwodowe od strony kaletki
Zerwanie częściowe po stronie stawowej	Ognisko hipoechogeniczne lub bezechowe w obrębie ścięgna, obejmujące włókna dośrodkowe od strony powierzchni stawowej
Zerwanie częściowe odcinka środkowego	Ognisko hipoechogeniczne lub bezechowe w obrębie ścięgna
Zerwanie na pełnej grubości	Obszar hipoechogeniczny przechodzący przez całą masę pierścienia rotatorów lub brak wizualizacji pierścienia rotatorów („przybliżenie” mięśnia naramiennego do powierzchni głowy kości ramiennej)

Tab. 1. Kryteria diagnostyczne dla ścięgna mięśnia nadgrzebieniowego w badaniu USG 2D

mej zgody, zostali zarejestrowani w jednym zakładzie opieki zdrowotnej. Uzyskano zgodę lokalnej rady etycznej w Bazylei. Badanie zostało przeprowadzone zgodnie z dobrymi praktykami klinicznymi oraz Deklaracją Helsińską.

Przeprowadzono dynamiczne badanie USG 2D obu barków zgodnie z wytycznymi Szwajcarskiego Stowarzyszenia Ultrasonograficznego (Schweizerische Gesellschaft für Ultraschall in der Medizin, SGUM)⁽¹⁴⁾, które są kompatybilne z wytycznymi Europejskiego Towarzystwa Radiologii Mięśniowo-Szkieletowej (European Society of Musculoskeletal Radiology, ESSR)⁽¹⁵⁾ oraz wytycznymi Europejskiej Ligi do Walki z Reumatyzmem (European League Against Rheumatism, EULAR).

Badanie USG 2D zostało przeprowadzone przez doświadczonego operatora USG (nauczyciel EULAR/EFSUMB, poziom 3 – badacz A). Wykorzystano urządzenie GE Logic E9 z sondą ML6-15. Czas skanowania w badaniu dynamicznym wynosił 10 minut.

Kryteria diagnostyczne dla ścięgna mięśnia nadgrzebieniowego w badaniu USG 2D przedstawiono w tab. 1.

Podczas jednego badania pozyskano zestaw wolumetrycznych obrazów 3D (urządzenie GE Logic E9 z sondą 6–16D oraz urządzenie GE Voluson E6 z sondą 6–16D; czas skanowania – 2 minuty) z wizualizacją ścięgna głowy długiej mięśnia dwugłowego ramienia na poziomie pomiędzy ścięgnami mięśni podłopatkowego i nadgrzebieniowego oraz miejsca przyczepu ścięgna mięśnia nadgrzebieniowego (*footprint*). Ramię było trzymane w pozycji zmodyfikowanej Crassa. Następnie automatycznie wykonano zestaw wolumetrycznych renderowanych obrazów 3D przyczepu ścięgna mięśnia nadgrzebieniowego w płaszczyźnie poprzecznej, wzdłużnej i osiowej.

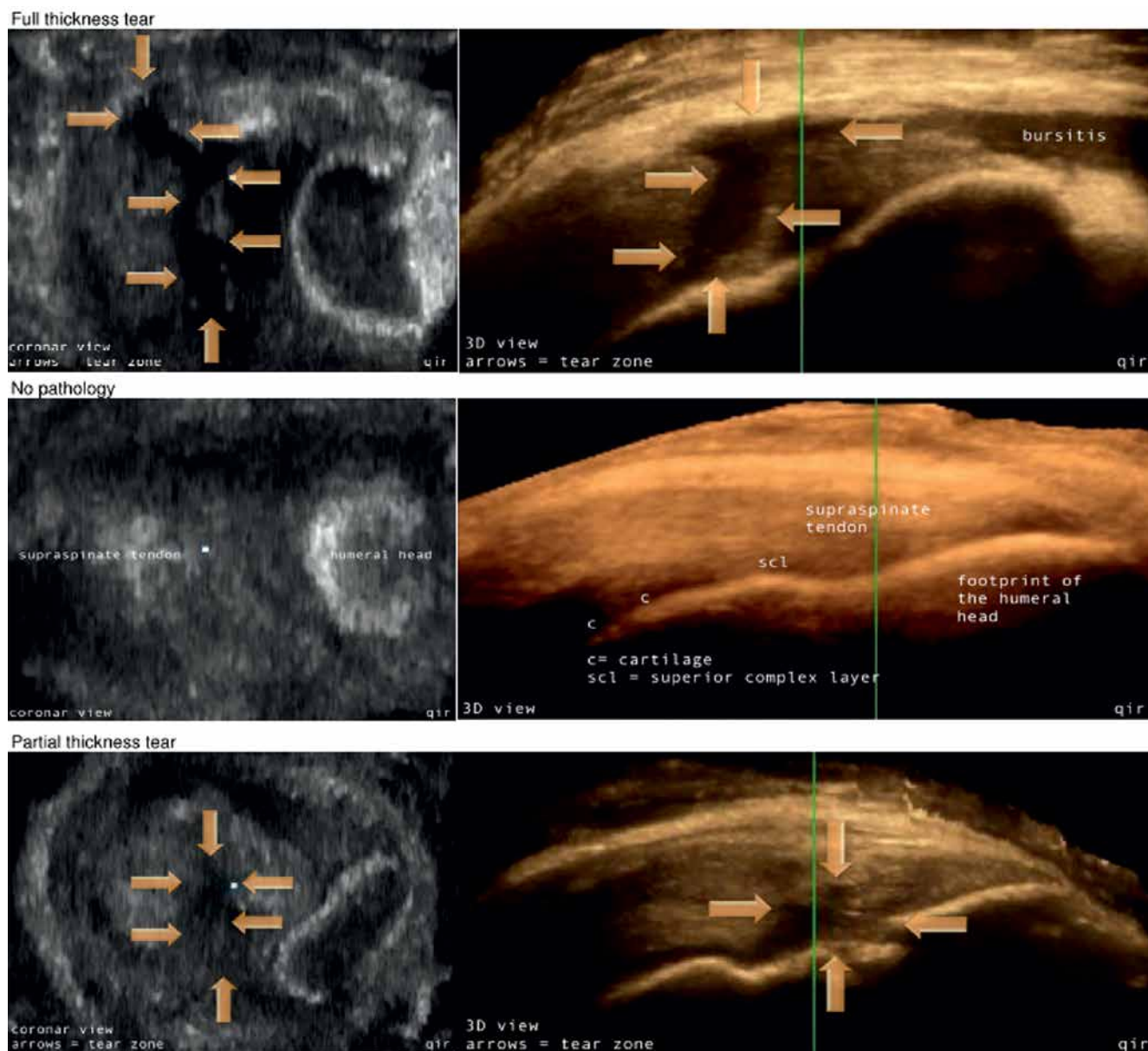


Fig. 1. Examples of 3D US images (coronar and 3D rendered views)

Ryc. 1. Przykłady obrazów USG 3D (przekrój czołowy i renderowany obraz 3D)

Rater A initially evaluated all 2D images and the 3D image sets. Afterwards, the blinded 3D image sets were scored independently by two other raters (one advanced sonographer EULAR Teacher/EFSUMB Level 2 = rater B and one fellow sonographer EULAR intermediate level = rater C) and with an interval of 3 months again by rater A. The most advanced rater A with 13 years of experience in ultrasound imaging set as “gold standard”. The findings of rater A were then compared to those of rater B and C.

The possible findings with corresponding diagnostic criteria were the same as for 2D US (cf. Tab. 1). Figure 1 shows the examples of the 3D US findings.

Bazując na założeniu, że w ocenie patologii pierścienia rotatorów dokładność ultradźwięków jest porównywalna do dokładności MRI⁽¹⁻³⁾, autorzy zdecydowali, aby nie wykonywać dodatkowego badania MRI. Wybrano trzech badaczy o różnym doświadczeniu w badaniach USG układu ruchu.

Badacz A ocenił wszystkie obrazy 2D oraz zestawy obrazów 3D. Następnie obrazy 3D zostały niezależnie ocenione przez dwóch pozostałych badaczy (doświadczony operator USG, nauczyciel EULAR/EFSUMB, poziom 2 – badacz B; operator USG, poziom średnio zaawansowany EULAR – badacz C), a po 3 miesiącach ponownie przez badacza A. Oceny najbardziej doświadczonego badacza A, mającego 13 lat doświadczenia w obrazowaniu ultradźwiękowym, zostały uznane za „złoty standard” i porównane z ocenami badaczy B i C.

Characteristic	Patients (n = 127)
Male patients – no. (%)	51 (40)
Total shoulders scanned – no.	210
Left shoulders scanned – no.	103
Right shoulders scanned – no.	107
Mean age – yrs	59.67 (range 18–86 yrs)

Tab. 2. Baseline characteristics of the patients enrolled in the study (n=127)

To assess the inter-rater reliability between different raters with regard to the findings in the 3D US image sets we used κ statistic. For rater A intra-rater reliability analyses were conducted between the findings in 2D US and 3D image sets. Overall interpretations of the κ statistic were based on the criteria described by Landis and Koch⁽¹⁶⁾. The level of reliability was defined as follows: κ values of 0.81 to 0.99 were considered to represent almost a perfect reliability; 0.61 to 0.80 – substantial reliability; 0.41 to 0.60 – moderate reliability; 0.21 to 0.40 – fair reliability; and 0.01 to 0.20 – slight reliability.

Results

We obtained 2309 images (210 3D image sets) of 127 adult patients suffering from unilateral shoulder pain. In our study, the 3D image set acquisition took about 2 minutes, compared to 10 minutes in the 2D US examination.

Baseline characteristics are shown in Tab. 2 and the US findings for the 2D US examinations are shown in Tab. 3.

There was almost a perfect intra-rater reliability of rater A in the overall classification of supraspinatus tendon tears (2D vs 3D κ = 0.892, pairwise reliability 93.81%, 3D scoring round 1 vs 3D scoring round 2 κ = 0.875, pairwise reliability 92.857%). There was almost a perfect overall reliability in the classification of full thickness tears between 2D and 3D (κ = 0.810), and a substantial reliability in the classification of partial thickness tears (κ = 0.667).

The inter-rater reliability between rater A and B was only moderate (κ = 0.497, pairwise reliability 70.95%) and fair between rater A and C (κ = 0.238, pairwise reliability 42.38%) comparing findings in the 3D US imaging sets.

Finding	Shoulders (n = 210)
No pathology	73 (34.8%)
Tendinosis	11 (5.2%)
Bursal sided partial-thickness tear	11 (5.2%)
Articular sided partial-thickness tear	69 (32.9%)
Intrasubstance sided partial-thickness tear	3 (1.4%)
Full thickness tear	44 (20.9%)

Tab. 3. 2D US findings (n=210)

Charakterystyka	Pacjenci (n = 127)
Liczba pacjentów płci męskiej (%)	51 (40)
Liczba skanowanych barków	210
Liczba skanowanych lewych barków	103
Liczba skanowanych prawych barków	107
Średni wiek (lata)	59,67 (zakres 18–86 lat)

Tab. 2. Podstawowa charakterystyka osób biorących udział w badaniu (n = 127)

Możliwe wyniki obserwacji wraz z kryteriami diagnostycznymi były takie same jak w przypadku USG 2D (tab. 1). Na ryc. 1 przedstawiono przykładowe wyniki obserwacji obrazów USG 3D.

Aby określić zgodność ocen poszczególnych badaczy, wykorzystano współczynnik zgodności κ . W przypadku badacza A przeprowadzono analizę zgodności jego ocen pomiędzy obrazami USG 2D i zestawami obrazów USG 3D. Ogólne interpretacje współczynnika κ zostały oparte na kryteriach Landisa i Kocha⁽¹⁶⁾. Poziom zgodności określano w następujący sposób: wartości κ od 0,81 do 0,99 – bardzo wysoka zgodność; wartości od 0,61 do 0,80 – wysoka zgodność; od 0,41 do 0,60 – umiarkowana zgodność; wartości od 0,21 do 0,40 – przeciętna zgodność; wartości od 0,01 do 0,20 – niska zgodność.

Wyniki

Uzyskano 2309 obrazów (210 zestawów obrazów 3D) z badań 127 dorosłych pacjentów cierpiących z powodu jednostronnego bólu barku. Czas pozyskiwania zestawu obrazów 3D wyniósł około 2 minut, w porównaniu z czasem 10 min w przypadku badania USG 2D.

Podstawowa charakterystyka badania została zawarta w tab. 2; w tab. 3 przedstawiono rozpoznania uzyskane na podstawie analizy obrazów USG 2D.

Badacz A osiągnął bardzo wysoki poziom zgodności swoich ocen w ogólnej klasyfikacji zerwania ścięgna mięśnia

Ustalenia	Liczba barków (n = 210)
Brak patologii	73 (34,8%)
Przewlekłe uszkodzenie ścięgna (tendinosis)	11 (5,2%)
Zerwanie częściowe po stronie kaletki	11 (5,2%)
Zerwanie częściowe po stronie stawowej	69 (32,9%)
Zerwanie częściowe odcinka środkowego	3 (1,4%)
Zerwanie na pełnej grubości	44 (20,9%)

Tab. 3. Rozpoznane patologie ścięgna nadgrzebieniowego na podstawie analizy obrazów USG 2D (n = 210)

Discussion

3D US of the shoulder is as accurate as 2D US when compared to MRI for the diagnosis of full- and partial-thickness supraspinatus tendon tears, and the 3D US examination significantly reduced the time between the initial scan and interpretation by a radiologist, which ultimately improved the efficiency of the workplace⁽¹¹⁾. However, the time spent on the interpretation of data obtained from 3D US was not measured in our study, therefore one could assume that the overall time will remain the same, except that the 3D US data are interpreted from a third party.

The inter-rater reliability of 2D US of the shoulder had been investigated in the past with experienced sonographer^(17,18). The corresponding κ values were only fair-to-moderate. The conclusion was that further consensus regarding the standardisation of scanning technique and diagnostic criteria is necessary to improve the reproducibility of musculoskeletal ultrasonography even for an experienced sonographer.

Given different levels of experience in 2D US, sonography has a high level of inter-rater reliability for full-thickness rotator cuff tears and a lower inter-rater reliability for partial-thickness and intratendinous rotator cuff tears⁽¹⁹⁾.

Studies have even shown that with 3D US, rotator cuff tears were more often correctly diagnosed than with a conventional 2D sonography^(12,20). However, 3D US of the shoulder seems to have a poorer inter- and intra-rater reliability compared to 2D sonography as investigated by Hayter et al.⁽¹³⁾ Using MRI for the classification of rotator cuff tears, fellowship-trained, experienced orthopaedic surgeons had a good reliability for predicting full-thickness rotator cuff tears and the number of tendons involved, and a moderate reliability in predicting the involved side of the partial-thickness rotator cuff tear, but a poor reliability in predicting the grade of the partial-thickness tear⁽²¹⁾.

In our study the reliability for predicting full thickness tears was almost perfect ($\kappa = 0.810$) and substantial ($\kappa = 0.667$) for predicting the partial-thickness tear. According to Hayter et al.⁽¹³⁾, the reason behind is the difficulty in the interpretation of small partial-thickness tears at the footprint on 3D images, a common area affected by anisotropy. Therefore, a standardized 3D US image acquisition with minimal areas of anisotropy should be defined. Standardized cross-sectional images (like those made automatically with the 3D probe) in multiple section planes would be necessary with repositioning of the transducer.

On the other hand, the 2D US examination compared to 3D US still has the additional advantage to collect information (e.g. sonopalpation, dynamic examination) in the differentiation of articular sided partial thickness tears versus tendinosis or anisotropy. The 2D US examination is known to be an examiner-dependent method. However, the superior advantage of ultrasound compared to all other imaging modalities (including 3D US) is the possibility to create static images as well as dynamic images, which

nadgrzebieniowego (2D vs 3D: $\kappa = 0,892$, porównywalność parami 93,81%; runda ocen obrazów 3D nr 1 vs runda ocen obrazów 3D nr 2: $\kappa = 0,875$, porównywalność parami 92,857%). Osiągnięto bardzo wysoki poziom zgodności w klasyfikacji zerwania pełnej grubości pomiędzy obrazami 2D i 3D ($\kappa = 0,810$) oraz wysoki poziom zgodności w rozpoznawaniu częściowych uszkodzeń ($\kappa = 0,667$).

Zgodność ocen na podstawie obrazów 3D pomiędzy badaczami A i B była tylko umiarkowana ($\kappa = 0,497$, porównywalność parami 70,95%) oraz przeciętna pomiędzy badaczami A i C ($\kappa = 0,238$, porównywalność parami 42,38%).

Dyskusja

W rozpoznawaniu częściowego i pełnego zerwania ścięgna mięśnia nadgrzebieniowego obrazowanie USG 3D barku jest tak samo dokładne jak obrazowanie USG 2D w odniesieniu do ostatecznej weryfikacji w MRI. Przy czym czas badania USG 3D jest znacząco krótszy, co pozytywnie wpływa na wydajność pracy w placówce opieki zdrowotnej⁽¹¹⁾. Jednakże w prezentowanym badaniu nie zmierzono czasu potrzebnego na interpretację danych uzyskanych z obrazowania USG 3D, dlatego można przyjąć, że całkowity czas pozostaje taki sam, z tą różnicą, że obrazy ultrasonograficzne są interpretowane przez osoby trzecie.

W przeszłości przeprowadzono badania dotyczące zgodności ocen patologii narządu ruchu pomiędzy doświadczonymi operatorami badania USG 2D^(17,18). Wyznaczone wartości współczynnika κ mieściły się w zakresie od zgodności przeciętnej do umiarkowanej. Wysłano wniosek, że konieczna jest dalsza standaryzacja techniki skanowania i kryteriów diagnostycznych w celu poprawy powtarzalności oceny.

Biorąc pod uwagę różny poziom doświadczenia operatorów w USG 2D, badanie cechuje się wysokim poziomem porównywalności pomiędzy różnymi badaczami w przypadku zerwania całkowitego stożka rotatorów oraz niższym w przypadku zerwania częściowego i uszkodzeń śródścięgniastych⁽¹⁹⁾.

Badania wykazały, że w przypadku korzystania z USG 3D dokładniej oceniano zerwania stożka rotatorów w porównaniu z klasyczną ultrasonografią 2D^(12,20). Jednocześnie stwierdzono jednak, że badanie USG 3D barku charakteryzuje się niższą porównywalnością pomiędzy badaczami i niższą zgodnością ocen jednego badacza w porównaniu z USG 2D, co zostało ujawnione w badaniu Haytera i wsp.⁽¹³⁾ W przypadku badania MRI wykazano, że doświadczeni ortopedzi po odbyciu stażu kierunkowego uzyskali wysoki poziom rozpoznawalności pełnego uszkodzenia stożka rotatorów wraz z określeniem rodzaju zerwanych ścięgien oraz umiarkowaną do niskiej zdolność diagnostyki uszkodzeń częściowych⁽²¹⁾.

W niniejszym badaniu zgodność ocen pomiędzy obrazami 2D i 3D w klasyfikacji zerwania stożka rotatorów na pełnej grubości była bardzo wysoka ($\kappa = 0,810$) i wysoka ($\kappa = 0,667$) dla zerwania częściowego. Według Haytera i wsp.⁽¹³⁾ przyczyną są trudności w interpretacji niewielkich uszkodzeń częściowych

provides a unique functional information. One weakness of our study is that this dynamic examination option was only given to rater A, who defined the standard sets, but not to rater B and C, who had to deal with sets of static 3D images.

Conclusions

In conclusion, the reliability of 3D US of the supraspinatus tendon depends on the level of experience of the sonographer, according to our data and experience. Experience in 2D US does not seem to be sufficient for the analysis of the 3D US imaging sets. Therefore, we propose to develop 3D diagnostic criteria for the examination of supraspinatus tendon pathologies prior to further 3D ultrasound studies of this shoulder area.

Conflict of interest

Authors do not report any financial or personal links with other persons or organizations, which might negatively affect the content of this publication and/or claim authorship rights to this publication.

Acknowledgments

We thank Andrea Hofstetter, Elisa Provenzano and Christine Paulicke for their assistance in patient recruitment.

w badaniu 3D w miejscu przyczepu ścięgna. Jest to obszar, na który wpływ ma ułożenie sondy USG (anizotropia). Dlatego też należałoby zdefiniować standardy obrazowania USG 3D z ograniczeniem obszarów, na które wpływ ma anizotropia. Konieczna jest standaryzacja obrazów przekrojowych (np. pozyskiwanych automatycznie przy pomocy sondy 3D) w wielu płaszczyznach, a także z uwzględnieniem zmiany pozycji głowicy.

Z drugiej strony badanie USG 2D w porównaniu z badaniem USG 3D ma dodatkową zaletę w kontekście pozyskiwania informacji (np. badanie sonograficzno-palpacyjne, badanie dynamiczne), co umożliwia różnicowanie zerwania częściowego po stronie stawowej od przewlekłego uszkodzenia ścięgna (*tendinosis*) lub anizotropii. Skuteczność badania USG 2D jest w dużej mierze zależna od operatora. Jednakże największą przewagą diagnostyki ultradźwiękowej nad wszystkimi innymi metodami obrazowania (w tym USG 3D) jest możliwość tworzenia zarówno obrazów statycznych, jak i dynamicznych, które dostarczają unikatowych informacji na temat czynności danej struktury. Słabością omawianego badania 3D jest umożliwienie wykonywania obrazowania dynamicznego wyłącznie badaczowi A, który zdefiniował standardy, oraz brak takiej możliwości w przypadku badaczy B i C, którzy musieli pracować z zestawami statycznych obrazów 3D.

Wnioski

Na podstawie zebranych danych i doświadczenia autorów można stwierdzić, że wiarygodność obrazowania USG 3D ścięgna mięśnia nadgrzebieniowego jest zależna od poziomu doświadczenia operatora. Doświadczenie w USG 2D zdaje się niewystarczające do analizy zestawów trójwymiarowych obrazów USG. Wobec tego należałoby określić kryteria diagnostyki trójwymiarowej patologii ścięgna mięśnia nadgrzebieniowego przed dalszymi badaniami USG 3D tego obszaru barku.

Konflikt interesów

Autorzy nie zgłaszają żadnych finansowych ani osobistych powiązań z innymi osobami lub organizacjami, które mogłyby negatywnie wpłynąć na treść publikacji oraz rościć sobie prawo do tej publikacji.

Podziękowania

Autorzy dziękują Andrei Hofstetter, Elisie Provenzano i Christine Paulicke za pomoc w rekrutacji pacjentów.

References / Piśmiennictwo

1. Yablon CM, Bedi A, Morag Y, Jacobson JA: Ultrasonography of the shoulder with arthroscopic correlation. *Clin Sports Med* 2013; 32: 391–408.
2. Wiener SN, Seitz WH Jr.: Sonography of the shoulder in patients with tears of the rotator cuff: accuracy and value for selecting surgical options. *AJR Am J Roentgenol* 1993; 160: 103–1010.
3. van Holsbeeck MT, Kolowich PA, Eyer WR, Craig JG, Shirazi KK, Habra GK, Vanderschueren GM, Bouffard JA: US depiction of partial-thickness tear of the rotator cuff. *Radiology* 1995; 197: 443–446.
4. Bachmann GF, Melzer C, Heinrichs CM, Möhring B, Rominger MB: Diagnosis of rotator cuff lesions: comparison of US and MRI on 38 joint specimens. *Eur Radiol* 1997; 7: 192–197.
5. Teeffey SA, Hasan SA, Middleton WD, Patel M, Wright RW, Yamaguchi K: Ultrasonography of the rotator cuff. A comparison of ultrasonographic and arthroscopic findings in one hundred consecutive cases. *J Bone Joint Surg Am* 2000; 82: 498–504.
6. O'Connor PJ, Rankine J, Gibbon WW, Richardson A, Winter F, Miller JH: Interobserver variation in sonography of the painful shoulder. *J Clin Ultrasound* 2005; 33: 53–56.
7. Rutten MJ, Jager GJ, Blickman JG: From the RSNA refresher courses: US of the rotator cuff: pitfalls, limitations, and artifacts. *Radiographics* 2006; 26: 589–604.
8. Kelly IM, Gardener JE, Lees WR: Three-dimensional fetal ultrasound. *Lancet* 1992; 339: 1062–1064.
9. Sohn Ch, Lenz GP, Thies M: Die dreidimensionale Ultraschalldarstellung der Säuglingshüfte. *Ultraschall in Med* 1990; 11: 302–305.
10. Riedl S, Kühner C, Tauscher A, Gühring U, Sohn C, Meeder PJ: Experimentelle Untersuchung von Meniskusläsionen. *Chirurg* 1997; 68: 1150–1155.
11. Co S, Bhalla S, Rowan K, Aippersbach S, Bicknell S: Comparison of 2- and 3-dimensional shoulder ultrasound to magnetic resonance imaging in a community hospital for the detection of supraspinatus rotator cuff tears with improved worktime room efficiency. *Can Assoc Radiol J* 2012; 63: 170–176.
12. Wallny T, Schild RL, Perlick L, Schulze Bertelsbeck D, Schmitt O: Die dreidimensionale Ultraschalluntersuchung der Rotatorenmanschette. *Ultraschall in Med* 2000; 21: 180–185.
13. Hayter CL, Miller TT, Nguyen JT, Adler RS: Comparative analysis of 2- versus 3-dimensional sonography of the supraspinatus tendon. *J Ultrasound Med* 2012; 31: 449–453.
14. Tamborrini G, Marx C, Müller AM, Szöllösy GPL, Draghi F, Bianchi S: *Ultrasound, Anatomy and Arthroscopy of the Shoulder*. Books on Demand, Norderstedt 2015.
15. European Society of Musculoskeletal Radiology: http://www.essr.org/cms/website.php?id=/en/index/educational_material.htm/.
16. Landis JR, Koch GG: An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977; 33: 363–374.
17. Scheel AK, Schmidt WA, Hermann KG, Bruyn GA, D'Agostino MA, Grassi W *et al.*: Interobserver reliability of rheumatologists performing musculoskeletal ultrasonography: results from a EULAR "Train the trainers" course. *Ann Rheum Dis* 2005; 64: 1043–1049.
18. Naredo E, Möller I, Moragues C, de Agustín JJ, Scheel AK, Grassi W *et al.*: Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. *Ann Rheum Dis* 2006; 65: 14–19.
19. Le Corroller T, Cohen M, Aswad R, Pauly V, Champsaur P: Sonography of the painful shoulder: role of the operator's experience. *Skeletal Radiol* 2008; 37: 979–986.
20. Wallny TA, Theuerkauf I, Schild RL, Perlick L, Bertelsbeck DS: The three-dimensional ultrasound evaluation of the rotator cuff – an experimental study. *Eur J Ultrasound* 2000; 11: 135–141.
21. Spencer EE Jr, Dunn WR, Wright RW, Wolf BR, Spindler KP, McCarty E *et al.*: Shoulder Multicenter Orthopaedic Outcomes Network: Interobserver agreement in the classification of rotator cuff tears using magnetic resonance imaging. *Am J Sports Med* 2008; 36: 99–103.