

research article

Diagnostic accuracy of haemophilia early arthropathy detection with ultrasound (HEAD-US): a comparative magnetic resonance imaging (MRI) study

Domen Plut^{1,2}, Barbara Faganel Kotnik^{2,3}, Irena Preloznik Zupan^{2,4}, Damjana Kljucevsek^{2,3}, Gaj Vidmar^{2,5}, Ziga Snoj^{1,2}, Carlo Martinoli⁶, Vladka Salapura^{1,2}

¹ Clinical Radiology Institute, University Medical Centre Ljubljana, Ljubljana, Slovenia

² Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia

³ Division of Paediatrics, University Medical Centre Ljubljana, Ljubljana, Slovenia

⁴ Division of Internal Medicine, University Medical Centre Ljubljana, Ljubljana, Slovenia

⁵ University Rehabilitation Institute Republic of Slovenia, Ljubljana, Slovenia

⁶ School of Medicine and Pharmacy, University of Genoa, Genoa, Italy

Radiol Oncol 2019; 53(2): 178-186.

Received 21 February 2019

Accepted 25 April 2019

Correspondence to: Assist. Prof. Vladka Salapura, M.D., Ph.D., Clinical Radiology Institute, University Medical Centre Ljubljana, Zaloška cesta 7, SI-1000 Ljubljana, Slovenia. Phone: +386 1 522 85 30; Fax: +386 1 522 24 97; E-mail: salapura@siol.net

Disclosure: No potential conflicts of interest were disclosed.

Background. Repeated haemarthroses affect approximately 90% of patients with severe haemophilia and lead to progressive arthropathy, which is the main cause of morbidity in these patients. Diagnostic imaging can detect even subclinical arthropathy changes and may impact prophylactic treatment. Magnetic resonance imaging (MRI) is generally the gold standard tool for precise evaluation of joints, but it is not easily feasible in regular follow-up of patients with haemophilia. The development of the standardized ultrasound (US) protocol for detection of early changes in haemophilic arthropathy (HEAD-US) opened new perspectives in the use of US in management of these patients. The HEAD-US protocol enables quick evaluation of the six mostly affected joints in a single study. The aim of this prospective study was to determine the diagnostic accuracy of the HEAD-US protocol for the detection and quantification of haemophilic arthropathy in comparison to the MRI.

Patients and methods. The study included 30 patients with severe haemophilia. We evaluated their elbows, ankles and knees (overall 168 joints) by US using the HEAD-US protocol and compared the results with the MRI using the International Prophylaxis Study Group (IPSG) MRI score.

Results. The results showed that the overall HEAD-US score correlated very highly with the overall IPSG MRI score ($r = 0.92$). Correlation was very high for the evaluation of the elbows and knees ($r \approx 0.95$), and slightly lower for the ankles ($r \approx 0.85$).

Conclusions. HEAD-US protocol proved to be a quick, reliable and accurate method for the detection and quantification of haemophilic arthropathy.

Key words: haemophilia; haemophilic arthropathy; HEAD-US; ultrasound; magnetic resonance imaging

Introduction

Intra-articular joint bleeds (haemarthroses) affect approximately 90% of patients with severe haemophilia.¹ The most frequently involved joints are the ankles, knees, and elbows.² Repeated episodes of intra-articular bleeding lead to progressive ar-

thropathy, which is the main cause of morbidity in these patients.³ The prevention of the occurrence of haemarthrosis is therefore important for the prevention of the arthropathy.

Small intra-articular bleeds may be unnoticed at physical examination and the detection of early signs of osteochondral damage is difficult by

clinical evaluation. It is known that osteochondral damage can be present in the joints that are asymptomatic and in which none or just a few bleeding episodes were previously recognized.^{4,5} These subtle articular changes of the subclinical disease can be detected by diagnostic imaging. Consequently, based on the diagnostic findings, appropriate treatment can be introduced or modified to prevent further disease progression and disability.⁶⁻¹¹

Magnetic resonance imaging (MRI) is the modality of choice to evaluate the musculoskeletal system because of its excellent spatial and contrast resolution. By MRI, it is possible to detect disease specific findings and give an accurate visualization of early arthropathy changes. However, MRI is a modality of high cost, its time of examining is long, it is usually poorly accessible and as such, it is not suitable for multi-joint screening. Additionally, it requires sedation in young children.¹²

Ultrasound (US), with the advent of last generation equipment, has excellent spatial resolution for the superficial structures. By US, it is now possible to depict the small, superficial structures of the musculoskeletal system as present in the early stages of haemophilic arthropathy. Contrary to MRI, US has a low cost, the time of examining is short and it is widely accessible. The drawbacks for the use of US in musculoskeletal radiology are poor visualization of inner joint structures and lack of standardized evaluation and reporting. In the field of haemophilic arthropathy, the development of the standardized US protocol for the detection of early changes in haemophilic arthropathy (HEAD-US) by Martinoli *et al.* in 2013 opened new perspectives in the use of US in management of patients with haemophilia. The HEAD-US protocol and scoring method are rapid to perform and enable full screening of the six joints in a single study.⁶

The aim of the present study was to determine the diagnostic accuracy of the HEAD-US protocol and scoring method for the detection and quantification of haemophilic arthropathy in patients with haemophilia in comparison to MRI using the International Prophylaxis Study Group (IPSG) MRI scoring scale.

Patients and methods

Patients

All patients were recruited at the Slovenian National Haemophilia Comprehensive Care Centre at the University Medical Centre Ljubljana. The inclusion criteria were age over 16 years, diagnosis

of a severe haemophilia A or B and prophylactic treatment with factor concentrates. Exclusion criteria were non-cooperation and contraindications for the MRI examination. Patients with prosthetic joints were allowed to participate in the study, but the prosthetic joint was not evaluated. The study group included a total of 30 patients (age range 16 to 49, mean age 33) who were willing to participate and met the aforementioned criteria. In 23 patients, six joints (elbows, knees and ankles) were systematically examined by US and MRI according to the protocols. One out of six joints was not examined in six patients due to a prosthetic implant. Two joints were excluded from the evaluation in the patient who had left lower limb amputation. The elbows were excluded in two patients because MRI could not be performed due to patient discomfort. Overall, 168 joints were examined in this study: 59 elbows, 53 knees and 56 ankles. The clinical evaluation of the joints according to the haemophilia joint health score HJHS 2.1 was obtained by a trained haemathologist on the day of the imaging examinations.

This prospective observational study was performed at a single tertiary center from June 2016 to March 2017. Research was conducted following the Helsinki Declaration. All patients included in the study provided a written informed consent for study participation. The National Medical Ethics Committee approved the study (Project number 70/11/15, approved on 11/21/2015).

Diagnostic imaging

In each patient, the US and MRI examinations were performed on the same day.

Ultrasound

US examinations were performed using a 13–5 MHz electronic linear-array transducer on a ProSound F75 scanner (Hitachi Aloka Medical, Ltd. Tokyo, Japan) by an experienced radiologist using the HEAD-US protocol and scoring method described elsewhere.⁶ The total scanning time per patient for all six joints combined was approximately 20 minutes. A series of images from 10 US examinations were reviewed and scored by another radiologist to determine the inter-rater reliability. This latter reviewer was blinded from the original scores of the examinations.

Magnetic resonance imaging

MRI was performed on a 3 Tesla unit (Achieva, Philips Healthcare, Eindhoven, The Netherlands).

TABLE 1. Baseline characteristics of the study population

Age: median; range (years)	33; 16–49					
Age of start of prophylaxis: mean (years)	17.4					
age group: 0–9 (patient count)	7					
age group: 10–19 (patient count)	14					
age group: 20+ (patient count)	9					
Duration of prophylaxis: mean (years)	15.4					
	Ankles		Knees		Elbows	
	Right	Left	Right	Left	Right	Left
No. of joints	30	29	25	28	28	28
No. of lifetime joint bleeds:						
0–5 (joint count)	5	5	12	15	14	13
6–20 (joint count)	12	11	7	10	4	2
> 20 (joint count)	13	12	5	3	9	12
Unknown (joint count)	0	1	1	0	1	1
HJHS 2.1 score: mean; max*	3.3; 12	2.6; 11	1.4; 7	1.2; 8	1.9; 9	1.9; 8

* Minimum was 0 for all the scores

An 8 elements phased array SENSE knee coil was used for the knee imaging, an 8 elements phased-array SENSE foot-ankle coil for the ankle imaging, and two 2 elements phase-array SENSE flex coils for the elbow imaging. The protocol included 3D T2*-weighted water selective gradient echo sequence (FOV, 160×160×108mm; voxel size, 0.58×0.58×0.50mm; flip angle: 15°; TE, 9.2/6.1ms; TR, 26ms), and 3D proton density (PD) weighted turbo spin echo sequence (FOV, 160×160×161mm; voxel size: 0.52×0.52×0.52mm; TE, 33ms; TR, 1000ms). The total scanning time was approximately 15 minutes per joint. In each patient, all joints were scanned in a single session for a total examination time extending up to two hours. After examining each joint, the patient was encouraged to stretch the body while the coils for the imaging of the next joint were setup. As mentioned, two MRI examinations were incomplete due to patient discomfort. All the MRI examinations were scored according to the International Prophylaxis Study Group (IPSG) MRI scale described elsewhere.¹³ The scoring was performed by an experienced musculoskeletal radiologist who was blinded regarding the results of the HEAD-US examinations. Additionally, the datasets of 10 MRI examinations were reviewed and scored by another experienced musculoskeletal radiologist to determine the inter-rater reliability. This latter reader was blinded from the original IPSG scores of the MRI examinations and from the HEAD-US scores.

Statistical analysis

Descriptive statistics were obtained to describe characteristics of the study group. We checked the inter-rater reliability of HEAD-US and MRI for the total scores using intra-class correlation (two-way mixed model, ICC(2,1)) and for all the sub-scores using Cohen's kappa statistics (with quadratic weights). We analyzed the agreement between HEAD-US and MRI using the Pearson correlation coefficient (r) for the total score and separately for the hypertrophic synovium, cartilage degradation, and bone changes (we could not use agreement coefficients because all those scores derive from different scales for HEAD-US and MRI). Regarding the agreement between HEAD-US and MRI for the cartilage degradation, we used Cohen's Kappa (with quadratic weights), because both scores are based on the same (0–4) scale. Agreement was illustrated using the concordance bubble plots.¹⁴

Results

Baseline characteristics of the study group are shown in Table 1. In our series, all patients underwent prophylactic treatment for haemophilia: 7 patients started therapy before the age of 10 years, 14 patients between 10 and 19 years, and 9 patients after the age of 20 years. The mean age at which prophylactic treatment was started was 17.4 years and the mean duration of the prophylaxis was 15.4 years.

In our series, the disease presentation was quite variable with a mean HJHS 2.1 score of 2.3 (range 0–12). HJHS scores were the highest in the ankle and the lowest in the knee, and correlated well with the lifetime number of joint bleeding episodes. The ankles were the joints with the most often recorded history of prior bleeds: 42% of examined ankles had >20 prior lifetime bleeds recorded. The knees were the least affected joints, with 51% of the examined knees having <5 prior lifetime bleeds recorded.

Inter-rater reliability

The inter-rater reliability of the interpretation was excellent for the US examinations (ICC values 0.960–0.996 for total score, median κ across sub-scores 1.000) and for the MRI (ICC values 0.957–0.990 for total score, median κ across sub-scores 0.815).

Diagnostic accuracy of ultrasound

HEAD-US scores were correlated with the IPSG MRI scores; results are shown in Table 2. A high overall correlation was found between the scores ($r \approx 0.92$). Correlation for the overall scores at the joint level was nearly perfect in the elbows and knees ($r \approx 0.95$) and slightly lower, but still very high in the ankles ($r \approx 0.85$). Separate evaluation of each parameter of the joint (synovial hypertrophy, cartilage degradation, bone changes) showed a medium-high to high agreement for all the parameters. The correlation between the HEAD-US and MRI scores was the lowest for the evaluation of the synovium hypertrophy and cartilage degradation at the ankle level ($r \approx 0.55$). All other parameters showed a high agreement between the methods ($r > 0.70$).

Concordance bubble plot for agreement between the HEAD-US and MRI scores at all three joint levels is shown in Figure 1. The distribution of circles within the plots demonstrates the variable degree of haemophilic arthropathy presentation in our study group at all joint levels. The plots also explicitly demonstrate the high overall correlation between the HEAD-US and MRI scores. The biggest deviation from the perfect line is shown at the ankle level, in the ankles with higher degree of haemophilic arthropathy.

In our series there were 42 joints with no haemophilic arthropathy, that are the joints scored with 0 by the IPSG MRI scoring system: 19 elbows, 20 knees and 3 ankles. In 35 of those joints the HEAD-US score was also 0. An example of a perfect concordance between the US and MRI examination for a knee with no haemophilic arthropathy is shown in Figure 2. In 7 joints with the IPSG MRI score 0 the HEAD-US score was 1. These HEAD-US examinations are false positives for the presence of

TABLE 2. Correlation between the HEAD-US and IPSG MRI scores

	Elbows	Knees	Ankles	All joints
Overall score (r)	0.949	0.941	0.838	0.921
Detailed scores:				
Synovial hypertrophy (r)	0.840	0.710	0.561	
Cartilage degradation (r)	0.734	0.812	0.537	
Bone changes (r)	0.883	0.741	0.725	

Notes: all the reported correlations are statistically significant ($p < 0.001$); the values for the elbows, knees and ankles are the averages over the right and left side values (the differences between them were negligible); the correlations are averaged using Fisher-z transformation.

haemophilic arthropathy. The false positive rate was 16.7%, which means specificity of HEAD-US to diagnose haemophilic arthropathy in our study was 83.3%. Detailed evaluation of the false positive examinations reveals that the findings diagnosed by US and not confirmed by MRI were: mild synovium hypertrophy in one elbow and two knees, small cartilage defect in two elbows and one ankle, and a small osteophyte in one knee.

Conversely, there were 6 joints that were scored with 0 by HEAD-US and scored positive by the IPSG MRI scoring. These HEAD-US examinations are the false negatives for the presence of haemophilic arthropathy. The false negative rate was 4.8%, which means the sensitivity of HEAD-US to diagnose haemophilic arthropathy in our study was 95.2%. Detailed evaluation of the false negative examinations reveals that the findings missed by US were: a cartilage defect at the tibial side of the talocrural joint (Figure 3), a small synovium hypertrophy in the posterior recess in another ankle, two small cartilage defects at the ulnar side of the joint in elbows, and two small osteochondral lesions at the ulnar side of the joint in another two

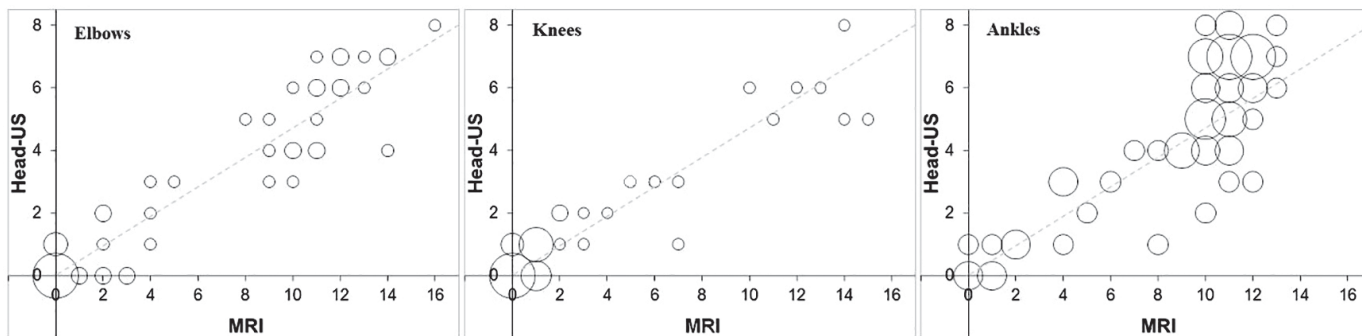


FIGURE 1. Concordance bubble-plot for depicting agreement between HEAD-US and MRI score for all three joints. The circles are centered at the observed combinations of the HEAD-US and MRI scores; their size is proportional to the number of the patients with a given combination. Dashed line represents a perfect agreement.

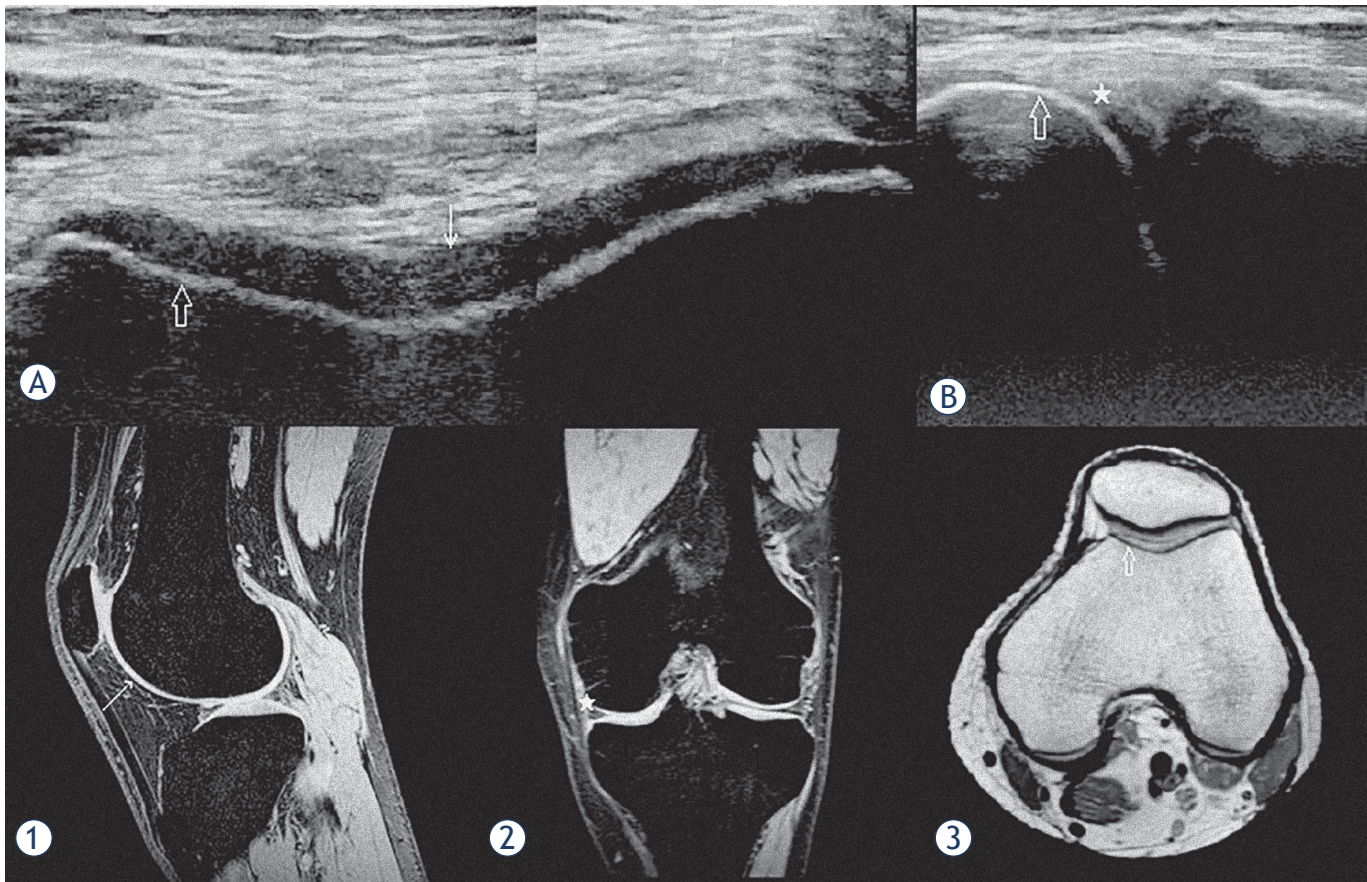


FIGURE 2. An example of a good concordance between HEAD-US and MRI. US images of the femoral trochlea in the transverse plane (**A**) and the medial femorotibial space in the coronal plane (**B**) are shown. T2* weighted MR images in the sagittal (**1**) and coronal (**2**) planes and a PD weighted MR image in the transverse plane (**3**) of the same knee are shown for comparison of the corresponding structures. The smooth surface, normal thickness and homogenous structure of the trochlear joint cartilage are shown in the US image (**A**) (white arrow); the corresponding intact cartilage is shown on MR image (**1**). The intact cortical bone of the medial femoral condyle (hollow arrow) is shown by US in the images (**A**) and (**B**); the corresponding cortical bone is depicted by MRI in the image (**3**). No signs of synovium hypertrophy are shown by US in the medial femorotibial recess (white star) in the image B; the corresponding recess confirming no synovium hypertrophy is shown by MRI in the image (**2**). On MRI, there were also no arthropathic changes in the parts of the joint not visualized by US. The images show a perfect concordance between US and MRI findings in this knee with no signs of haemophilic arthropathy.

elbows (images not shown). All the described osteochondral pathologic changes of haemophilic arthropathy were outside the area of visualization by US.

The positive predictive value for the presence of haemophilic arthropathy by HEAD-US in our study was 94.5% and the negative predictive value 85.4%.

Discussion

The aim of the study was to evaluate the diagnostic accuracy of US for the detection and quantification of haemophilic arthropathy using the HEAD-US protocol and scoring method in comparison to MRI using the IPSPG MRI scoring scale. We evalu-

ated the three most commonly affected joints in patients with severe haemophilia: the elbows, knees and ankles. Overall, we evaluated 168 joints: 59 elbows, 53 knees and 56 ankles. The sensitivity of HEAD-US to detect the signs of haemophilic arthropathy in our study was 95.2%; the specificity was 83.3%. The results of the correlation analysis showed a very high correlation for the quantification of haemophilic arthropathy between US using the HEAD-US protocol and scoring method and MRI using the IPSPG MRI score ($r = 0.921$, $p < 0.001$). The correlation was nearly perfect for the elbows and knees ($r \approx 0.95$), and very high for the ankles ($r \approx 0.85$). Excellent inter-rater reliability in our study for both MRI and US further support the validity of both scales (ICC values 0.960–0.996 for US, and ICC values 0.957–0.990 for MRI).

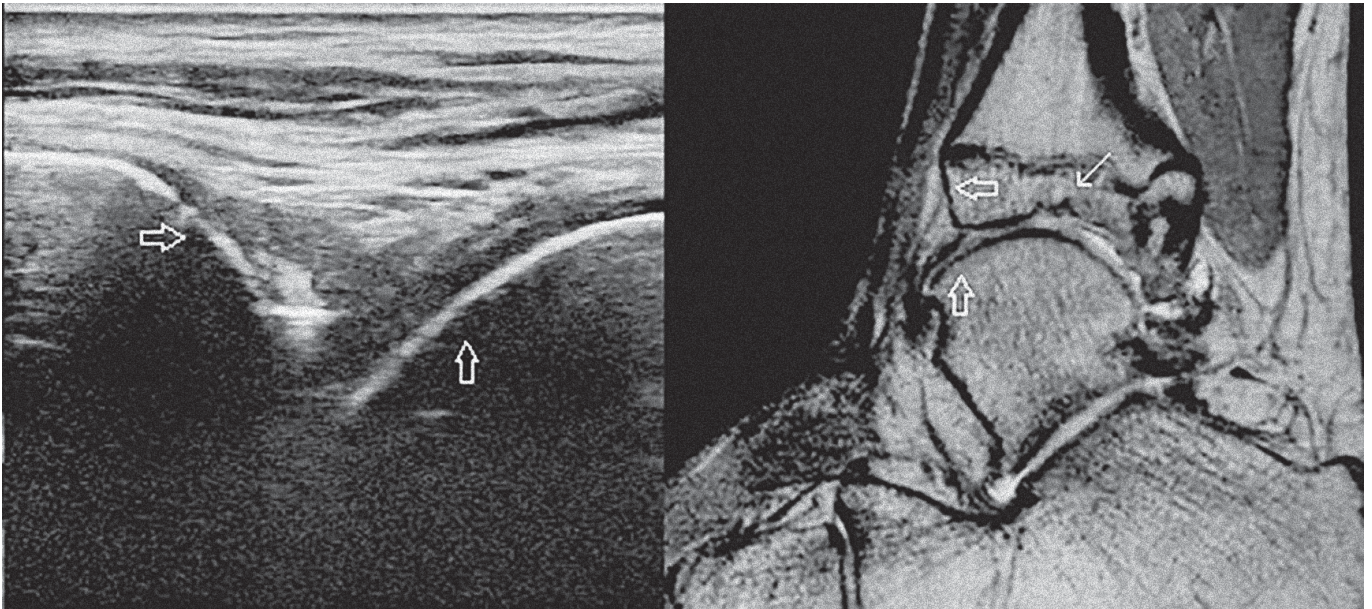


FIGURE 3. An example of a discordance between US and MRI. An US image of the tibiotalar joint in the sagittal plane is shown on the left, a PD weighted MR image of the same ankle in the sagittal plane is shown on the right for comparison of the corresponding structures. In both images, the smooth surface of the tibial cortical bone is marked by the horizontal hollow arrow and the smooth surface of the talar cortical bone is marked by the vertical hollow arrow. MRI demonstrates an osteochondral defect at the tibial side of the talocrural joint (white arrow), which is outside of the visualization area of US.

These results show that the HEAD-US method is highly accurate and reliable in comparison to MRI for detection and quantification of haemophilic arthropathy changes. High sensitivity and specificity verify the method is a dependable tool for the recognition of presence of arthropathy changes. High correlation proves a great value of the method for evaluation of pathology at any different stage of the disease. Our results indicate that HEAD-US is equally applicable for all evaluated joints. However, we noticed some differences in evaluation of specific parameters in the ankle joint. The agreement between the methods for the evaluation of the ankles in comparison to the other joints was lower for the osteochondral changes (especially cartilage degradation, $r = 0.54$ for the ankles, 0.81 for knees, and 0.73 for elbows) and for the synovium hypertrophy ($r = 0.56$ for the ankles, 0.71 for knees and 0.84 for elbows). The lower agreement for the osteochondral changes in the ankles is likely due to the limited visualization of the weight-bearing part of the osteochondral surface of the ankle joint by US. By US, only a part of the osteochondral surface of the talar dome is accessible for the evaluation, whereas MRI enables optimal visualization of the whole osteochondral surfaces of the talocrural and talocalcaneal joints. The lower agreement for the synovium hypertro-

phy in the ankles between US and MRI could be because of the ability of MRI to depict haemosiderin deposits in the synovia that appear due to the susceptibility artifact. This artifact allows great visualization of the synovia with haemosiderin deposits on MRI even when the synovia is not extensively hypertrophied. The ankles were in average the most affected joints in our study (HJHS average score of 8.9 , compared to 5.5 for the elbows and 2.7 for the knees); consequently, this fact may have the biggest impact in the ankles. Another reason for the lower agreement for the synovial hypertrophy is probably in the scoring scales: minimal hypertrophy of the synovium is scored 0 by the HEAD-US scoring scale, while it is scored 1 on the IPSG MRI scale.

In the literature, some studies investigated the diagnostic accuracy of US in the assessment of haemophilic arthropathy in comparison with MRI.^{5,15-18} These studies showed promising results for US and stressed the importance of further research in this field.^{5,19} In comparison to these studies, our study included a larger cohort of 168 joints and included all three mostly involved joints with a wide range of haemophilic arthropathy presentation. These characteristics make our study the biggest and most comprehensive study for the evaluation of diagnostic accuracy of US for detection and quanti-

fication of haemophilic arthropathy in comparison to MRI to date.

In a recent study published in May 2018 by Foppen *et al.*, the HEAD-US system was compared with MRI using the IPSPG MRI scoring scale. Their study evaluated knees and ankles of 24 patients with haemophilia with no or minimal arthropathy. The study demonstrated a strong correlation between the HEAD-US score and the IPSPG MRI score for the evaluation of the synovium hypertrophy ($r = 0.90$, $p < 0.01$), cartilage degradation ($r = 0.73$, $p < 0.01$) and bone changes ($r = 0.88$, $p < 0.01$).¹⁵ Comparing their results to the results of our study, the biggest discrepancy is in the correlation between the methods for the evaluation of the synovium hypertrophy. In their study, the correlation between US and MRI was the highest for the evaluation of synovium hypertrophy of all evaluated parameters of the joint, whereas in our study, the correlation between US and MRI was the lowest for the evaluation of synovium hypertrophy in comparison to other joint parameters, especially at the ankle level. We believe the key for the differences in the results between our studies is in the evaluation of the synovium hypertrophy by MRI. The assessment of the degree of synovium hypertrophy is not clearly defined in the IPSPG MRI scoring scale; it is defined as no, small, moderate, or large hypertrophy, with no objective measurements to define the different degrees. In their study, the authors determined the lower cut-off values for the synovium measurements to define the presence/absence of synovium hypertrophy on MRI.¹⁵ In our study, any impression of synovium hypertrophy, due to a focal increase in thickness or loss of signal due to the susceptibility artifact was evaluated as synovium hypertrophy by the MRI reviewers. Consequently, in our study some more minimal synovium hypertrophy was scored 1 by the IPSPG MRI scale, which was not scored by the HEAD-US. As already mentioned earlier, minimal synovial hypertrophy is scored 0 by the HEAD-US scoring scale. Nevertheless, the results of both studies prove that the HEAD-US method is a reliable tool for the detection and quantification of haemophilic arthropathy with significant correlation to the evaluation by MRI.

The other comparable studies also presented a high correlation between US and MRI, with some differences in the methodology used. Doria *et al.* evaluated 34 ankles and 25 knees in two centers, using IPSPG MRI scale as a reference standard. Their US examination was comprehensive with an US scoring scale corresponding to the items evalu-

ated in the IPSPG MRI scale. Their results showed that when data was acquired by radiologists, US was highly reliable for assessing soft-tissue changes (ICC 0.98 for ankles and 0.97 for knees) and substantially to highly reliable for assessing osteochondral changes (ICC 0.61 for ankles and 0.89 for knees).¹⁶ Sierra Aisa *et al.* evaluated 30 joints (knees and ankles), comparing findings detected by a comprehensive US exam to findings found on MRI. The parameters used for MRI analysis were the same as analyzed by US. Their results showed a good positive correlation between US and MRI in detecting synovial hypertrophy ($\kappa = 0.839$ – 1.000) and erosions ($\kappa = 0.850$ – 1.000). They showed lower correlation between the methods for detecting bone cysts ($\kappa = 0.643$ – 0.552) and cartilage loss ($\kappa = 0.643$ – 0.462).¹⁷ Di Minno *et al.* evaluated 40 clinically asymptomatic joints (knees, elbows and ankles) in patients with haemophilia, evaluating ability of US to detect early haemophilic arthropathy. A comprehensive US examination with Doppler was performed for joint evaluation, progressive and additive MRI scales were used as a reference standard. Their results showed significant correlation between US and MRI ($r = 0.732$ for the additive MRI scale, $r = 0.598$ for the progressive MRI scale).⁵ Acharya *et al.* compared power Doppler and grey-scale US findings with dynamic contrast enhanced MRI in 33 joints (elbows, knees and ankles). Their results showed a strong correlation between US and MRI in assessment of synovial hypertrophy ($r = 0.70$) and vascularity ($r = 0.73$).¹⁸

All the aforementioned studies were based on comprehensive US protocols that were very complex, with only trained readers potentially able to get an acceptable reproducibility. In the present paper, we used a simplified HEAD-US scanning protocol and scoring method which was developed for non-imaging specialists to enable them to analyze the joints at the time and place of patient care (point-of-care). The HEAD-US system includes systematic evaluation of the recesses of the elbow, knee, and ankle for the detection of synovium hypertrophy and evaluation of a single osteochondral surface in the elbow (anterior aspect of the distal humeral epiphysis), knee (femoral trochlea) and ankle (the anterior aspect of the talar dome) for the detection of osteochondral damage. The rationale of selecting one reference surface is based on the evidence that the diffuse establishment of osteochondral damage in haemophilic arthropathy may warrant the policy of considering one surface representative of the overall status of the joint without significantly reducing the sensitivity of the meth-

od. This keeps the HEAD-US method easy and fast to perform, thus enabling the examiner to screen the six joints of interest in a single study.⁶ In our study, the HEAD-US was performed by an experienced radiologist. The results of our study prove, that even this simplified US examination is reliable and accurate for the detection and quantification of haemophilic arthropathy.

Regarding the inter-rater reliability of US and MRI scores, an excellent agreement between readers was noted in our study, a finding that is supported by previous studies with US^{15,20} and MRI.^{13,15}

Generally speaking, MRI is still the only widely validated gold standard tool for precise evaluation of joints. MRI scoring scales have been widely used and approved as reference standards in haemophilic arthropathy trials, although barely applied in clinical practice for diagnosis and outcome, because of their complexity, time commitment and cost. In our study, the protocol used for MR imaging was optimized for the detection of early haemophilic arthropathy. It achieved good visualization of the cortical bone, articular cartilage and synovium hypertrophy, including haemosiderin deposits, while keeping the examination time as short as possible. Despite keeping the examination time very short, we encountered the typical problems related to this modality. One patient was claustrophobic, so imaging of all joints was not possible on MRI. In several patients, immobilization was uncomfortable, caused pain and the process of image acquisition was impaired by motion artifact and imaging had to be repeated. On the contrary, US due to its great availability, low cost and feasibility is a great method for diagnostic evaluation and regular follow-up of patients with haemophilia. The development of the standardized and simplified HEAD-US protocol and scoring method made a crucial step towards wider clinical application with ability to perform it quickly and with high reproducibility. Regarding the evaluation by HEAD-US, the joints affected by severe arthropathy with prominent osteophytes and narrowed joint space were somewhat more demanding to evaluate. However, the interpretation that the joint status was severely compromised was straightforward. Our study results prove that this fast and simplified protocol is diagnostically accurate and reliable. For this reason we believe that HEAD-US is the most appropriate method for regular screening and patient follow-up. We implemented this method into our regular clinical practice during yearly follow-up of pediatric patients with haemophilia in 2017. Since then we managed to include all pediatric patients with

haemophilia in our country into the regular screening program.

Although, our study group included all patients with severe haemophilia in Slovenia aged 16–50 years, we had a relatively small number of patients with an early haemophilic arthropathy. This gave us only a limited insight into depiction of early arthropathy changes.

Conclusions

In our study, the HEAD-US protocol and scoring method proved to be a quick, reliable and accurate method for the detection and quantification of haemophilic arthropathy in comparison to MRI. HEAD-US shows great promise in diagnostics and regular follow-up of patients with haemophilia, possibly influencing prophylactic treatment to prevent occurrence of haemophilic arthropathy, especially in children, or prevent disease progression. Further long-term studies are needed to evaluate the role of US in modification of prophylactic therapy and prevention of haemophilic arthropathy or its progression.

References

- Hilgartner MW. Current treatment of hemophilic arthropathy. *Curr Opin Pediatr* 2002; **14**: 46-9. PMID: 11880733
- Dalyan M, Tuncer S, Kemahli S. Hemophilic arthropathy: evaluation of clinical and radiological characteristics and disability. *Turk J Pediatr* 2000; **42**: 205-9. PMID: 11105618
- Roosendaal G, Lafeber FP. Pathogenesis of haemophilic arthropathy. *Haemophilia* 2006; **12**: 117-21. doi: 10.1111/j.1365-2516.2006.01268.x
- Manco-Johnson MJ, Abshire TC, Schapiro AD, Riske B, Hacker MR, Kilcoyne R, et al. Prophylaxis versus episodic treatment to prevent joint disease in boys with severe hemophilia. *N Engl J Med* 2007; **357**: 535-44. doi: 10.1056/NEJMoa067659
- Di Minno MND, Iervolino S, Soscia E, Tosetto A, Coppola A, Schiavulli M et al. Magnetic resonance imaging and ultrasound evaluation of "healthy" joints in young subjects with severe haemophilia A. *Haemophilia* 2013; **19**: e167-73. doi: 10.1111/hae.12107
- Martinoli C, Della Casa Alberighi O, Di Minno G, Graziano E, Molinari AC, Pasta G, et al. Development and definition of a simplified scanning procedure and scoring method for haemophilia early arthropathy detection with ultrasound (HEAD-US). *J Thromb Haemost* 2013; **109**: 1170-9. doi: 10.1160/th12-11-0874
- Berntorp E, Boulyjenkov V, Brettler D, Chandy M, Jones P, Lee C, et al. Modern treatment of haemophilia. *Bull World Health Organ* 1995; **73**: 691-701. PMID: 8846496
- Ljung R. Paediatric care of the child with haemophilia. *Haemophilia* 2002; **8**: 178-82. doi: 10.1046/j.1365-2516.2002.00631.x
- Nilsson IM, Berntorp E, Lofqvist T, Pettersson H. Twenty-five years' experience of prophylactic treatment in severe haemophilia A and B. *J Intern Med* 1992; **232**: 25-32. doi: 10.1111/j.1365-2796.1992.tb00546.x
- Rosendaal FR, Smit C, Varekamp I, Brocker-Vriends AH, van Dijk H, Suurmeijer TP, et al. Modern haemophilia treatment: medical improvements and quality of life. *J Intern Med* 1990; **228**: 633-40. doi: 10.1111/j.1365-2796.1990.tb00291.x

11. Berntorp E. Joint outcomes in patients with haemophilia: the importance of adherence to preventive regimens. *Haemophilia* 2009; **15**: 1219-27. doi: 10.1111/j.1365-2516.2009.02077.x
12. Kilcoyne RF, Nuss R. Radiological assessment of haemophilic arthropathy with emphasis on MRI findings. *Haemophilia* 2003; **9**: 57-63. doi: 10.1046/j.1365-2516.9.s1.11.x
13. Lundin B, Manco-Johnson ML, Ignas DM, Moineddin R, Blanchette VS, Dunn AL, et al; International Prophylaxis Study Group. An MRI scale for assessment of haemophilic arthropathy from the International Prophylaxis Study Group. *Haemophilia* 2012; **18**: 962-70. doi: 10.1111/j.1365-2516.2012.02883.x
14. Vidmar G, Rode N. Visualizing concordance. *Comput Stat* 2007; **22**: 499-509. doi: 10.1007/s00180-007-0057-9
15. Foppen W, der Schaaf ICV, Beek FJA, Mali WPTM, Fischer K. Diagnostic accuracy of point-of-care ultrasound for evaluation of early blood-induced joint changes: comparison with MRI. *Haemophilia*; **24**: 971-9. doi: 10.1111/hae.13524
16. Doria AS, Keshava SN, Mohanta A, Jarrin J, Blanchette V, Srivastava A, et al. Diagnostic accuracy of ultrasound for assessment of hemophilic arthropathy: MRI correlation. *Am J Roentgenol* 2015; **204**: W336-47. doi: 10.2214/AJR.14.12501
17. Sierra Aisa C, Lucía Cuesta JF, Rubio Martínez A, Fernández Mosteirín N, Iborra Muñoz A, Abio Calvete M, et al. Comparison of ultrasound and magnetic resonance imaging for diagnosis and follow-up of joint lesions in patients with haemophilia. *Haemophilia* 2014; **20**: e51-7. doi: 10.1111/hae.12268
18. Acharya SS, Schloss R, Dyke JP, Mintz DN, Christos P, Dimichele DM, et al. Power Doppler sonography in the diagnosis of hemophilic synovitis - a promising tool. *J Thromb Haemost* 2008; **6**: 2055-61. doi: 10.1111/j.1538-7836.2008.03160.x
19. Ligocki CC, Abadeh A, Wang KC, Adams-Webber T, Blanchette VS, Doria AS. A systematic review of ultrasound imaging as a tool for evaluating haemophilic arthropathy in children and adults. *Haemophilia* 2017; **23**: 598-612. doi: 10.1111/hae.13163
20. Stephensen D, Classey S, Harbidge H, Patel V, Taylor S, Wells A. Physiotherapist inter-rater reliability of the haemophilia early arthropathy detection with ultrasound protocol. *Haemophilia* 2018; **24**: 471-6. doi: 10.1111/hae.13440