ARTHRITIS & RHEUMATOLOGY
Vol. 66, No. 1, January 2014, pp 218–227
DOI 10.1002/art.38197

Detection of Enthesitis in Children With Enthesitis-Related Arthritis

Dolorimetry Compared to Ultrasonography

Pamela F. Weiss,[1] Nancy A. Chauvin,[1] Andrew J. Klink,[1] Russell Localio,[2] Chris Feudtner,[1] Diego Jaramillo,[1] Robert A. Colbert,[3] David D. Sherry,[1] and Ron Keren[1]

Objective. **To evaluate the distribution of enthesitis and the accuracy of physical examination with a dolorimeter for the detection of enthesitis in children, using ultrasound (US) assessment as the reference standard.**

Methods. **We performed a prospective cross-sectional study of 30 patients with enthesitis-related arthritis (ERA) and 30 control subjects. The following tendon insertion sites were assessed by standardized physical examination with a dolorimeter and US: common extensor on the lateral humeral epicondyle, common flexor on the medial humeral epicondyle, quadriceps at the superior patella, patellar ligament at the inferior patella, Achilles, and plantar fascia at the calcaneus.**

Results. **Abnormal findings on US were detected most commonly at the insertion of the quadriceps (30% [18 of 60 sites]), common extensor (12% [7 of 60]), and Achilles (10% [6 of 60]) tendons. The intrarater reliability of US (kappa statistic) was 0.78 (95% confidence interval [95% CI] 0.63–0.93), and the interrater reliability was 0.81 (95% CI 0.67–0.95). Tenderness as detected by standardized dolorimeter examination had poor positive predictive value for US-confirmed enthesitis. In comparison to controls, patients with ERA reported more pain and had lower pain thresholds at every site, including control sites ($P < 0.001$ for all comparisons). The interrater reliability of dolorimeter examination for detection of enthesitis was low ($\kappa = 0.49$ [95% CI 0.33–0.65]).**

Conclusion. **Compared to US, standardized dolorimeter examination for the detection of enthesitis in children has poor accuracy and reliability. The decreased pain threshold of ERA patients likely contributed to the limited accuracy of the physical examination findings. Further research regarding the utility of US for identifying enthesitis at diagnosis of juvenile idiopathic arthritis, accurately predicting disease progression, and guiding therapeutic decisions is warranted.**

Enthesitis refers to inflammation at the attachments of the ligaments, tendons, and joint capsules to the bone. It is a distinct clinical hallmark of the spondyloarthropathies (SpA) in both children and adults. In children, enthesitis is usually diagnosed by clinical findings, including localized pain, tenderness, and swelling. These features, however, are nonspecific and can also be found in normal children (1) and in patients with overuse injuries, apophysitis, and fibromyalgia. In adult SpA, magnetic resonance imaging (MRI) and ultrasound (US) with power Doppler are used to distinguish inflammatory enthesitis from these other conditions (2–4). In studies that used these imaging modalities as the reference standard in the evaluation of adult SpA, physical examination had poor sensitivity for detecting inflammatory enthesitis (2), and US was a useful tool

Dr. Weiss's work was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases grant 1-K23-AR059749-01A1); she is also recipient of a Rheumatology Research Foundation Bridge Funding award. Dr. Colbert's work was supported by the NIH (National Institute of Arthritis and Musculoskeletal and Skin Diseases Intramural Research Program grant Z01-AR-041184).

[1]Pamela F. Weiss, MD, MSCE, Nancy A. Chauvin, MD, Andrew J. Klink, MPH, Chris Feudtner, MD, PhD, MPH, Diego Jaramillo, MD, David D. Sherry, MD, Ron Keren, MD, MPH: Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; [2]Russell Localio, MPH, MS, PhD: University of Pennsylvania, Perelman School of Medicine, Philadelphia, Pennsylvania; [3]Robert A. Colbert, MD, PhD: National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH, Bethesda, Maryland.

Address correspondence to Pamela F. Weiss, MD, MSCE, Children's Hospital of Philadelphia, 3535 Market Street, Room 1526, Philadelphia, PA 19104. E-mail: weisspa@email.chop.edu.

Submitted for publication June 20, 2013; accepted in revised form September 10, 2013.

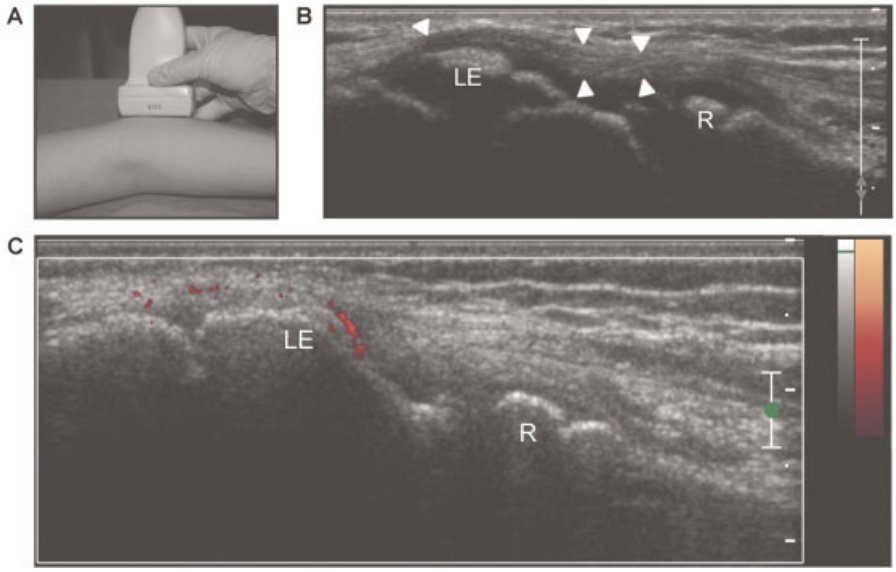

Figure 1. Ultrasound examination of the common extensor tendon insertion at the lateral elbow using a Philips IU22 machine with a high-frequency linear-array 12-MHz transducer. Power Doppler imaging was performed in long and transverse planes. **A,** Images were acquired with the elbow in mild flexion and the forearm pronated. **B,** An image from the gray-scale evaluation shows the thickness of the tendon, as delineated by the **arrowheads**. LE = lateral epicondyle; R = radial head. **C,** An image from the Doppler evaluation shows moderate vascularization. Red regions are increased power Doppler signals, which indicate increased vascularity.

for following the response of enthesitis to treatment (5). In those studies, the sensitivity of the physical examination findings was low, ranging from 0.16 to 0.22, and the specificity was moderate, ranging from 0.80 to 0.87 (2,6).

The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of clinical examination for the detection of enthesitis have been sparsely studied in juvenile idiopathic arthritis (JIA) (7,8). The accurate diagnosis of enthesitis is important because its presence has implications with regard to the classification of JIA, which in turn, influences treatment decisions and monitoring for extra-articular manifestations of the disease. Currently, physicians vary with regard to which entheses they evaluate, the amount of pressure they apply, and what criteria they use to diagnose positive findings (e.g., verbal indication of pain or withdrawal upon pain). The International League of Associations for Rheumatology (ILAR) criteria for JIA classifies many children with SpA as having enthesitis-related arthritis (ERA) (9) based solely on the presence of enthesitis.

In the present study, we sought to evaluate the distribution of enthesitis and the accuracy of physical examination for the detection of enthesitis in children. We examined 6 bilateral entheses in children with ERA, the category of JIA with the highest prevalence of enthesitis, as well as in healthy controls. We used a dolorimeter to objectively measure the pressure applied during the examination and performed US as the reference standard test at the same visit.

PATIENTS AND METHODS

Human subjects protections. The protocol for the conduct of this study was approved by the Committee for the Protection of Human Subjects at the Children's Hospital of Philadelphia (CHOP).

Study subjects. The study subjects were a convenience sample drawn without regard to disease duration, severity, or current activity or therapy from a source population of children with a diagnosis of ERA who were 5–18 years of age and had been evaluated at the CHOP Rheumatology Clinic between March 2011 and February 2013. All patients met the ILAR criteria for ERA (9), according to the treating physician. The presence of at least 1 tender enthesis on physical examination performed at the screening visit was required in order for an ERA patient to be eligible for study. Exclusion criteria were contraindications for performing US, including local malignancy, metal implants below the area of examination, localized tissue or bone infections, or vascular abnormalities. Healthy control subjects were recruited from a local primary care practice and were age- and sex-matched to the patients.

Clinical data. We collected self-reported information on sex, age, race, ethnicity, and family history of rheumatic disease. We abstracted information about medications (past and current), disease duration, HLA–B27 status, and antinuclear antibody (ANA) status from the electronic health record.

Health status measures. The patients (if at least 13 years of age) or their parents or legal guardians completed the following 5 health status questionnaires: the Pediatric Rheumatology Quality of Life (PRQL) (10), the Childhood Health

Table 1. Demographic and clinical features of the study subjects*

	ERA patients (n = 30)	Control subjects (n = 30)
Age at visit, median (IQR) years	13 (11–15)	12 (7–15)
Disease duration, median (IQR) years	0.6 (0.1–2.6)	–
No. (%) male	18 (60)	16 (52)
Criteria for ERA diagnosis, no. (%)		
Enthesitis	30 (100)	–
Arthritis	29 (97)	–
Inflammatory back pain	14 (47)	–
Acute symptomatic uveitis	2 (7)	–
Onset of arthritis in a male ≥6 years of age†	17 (94)	–
HLA–B27 positive	9 (30)	–
Family history of AS, ERA, IBD with associated sacroiliitis, acute uveitis, or ReA in a first-degree relative	7 (23)	–
No. of joints with active arthritis, median (IQR)‡	2 (1–2)	0 (0)
Erythrocyte sedimentation rate, median (IQR) mm/hour‡	0 (0–12)	–
C-reactive protein, median (IQR) mg/dl	0 (0–0.5)	–
PRQL score, median (IQR)	6.5 (3–10)	0 (0–1)
Patient's/parent's assessment of pain over the previous week, by VAS, median (IQR) score	4 (2–7)	0 (0–0)
C-HAQ score, median (IQR)	0.31 (0.13–0.88)	0 (0–0)
BASFI, median (IQR)	1.6 (0.7–4.9)	0 (0–0)
BASDAI, median (IQR)	3.8 (2.6–6.4)	0 (0–0)
Current medications, no. (%)		
Daily NSAIDs	19 (63)	–
DMARDs	7 (23)	–
Anti-TNF agents		
Etanercept	8 (27)	–
Infliximab	2 (7)	–
Adalimumab	1 (3)	–

* ERA = enthesitis-related arthritis; IQR = interquartile range; AS = ankylosing spondylitis; IBD = inflammatory bowel disease; ReA = reactive arthritis; PRQL = Pediatric Rheumatology Quality of Life; VAS = visual analog scale; C-HAQ = Childhood Health Assessment Questionnaire; BASFI = Bath Ankylosing Spondylitis Functional Index; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; NSAIDs = nonsteroidal antiinflammatory drugs; DMARDs = disease-modifying antirheumatic drugs; anti-TNF = anti–tumor necrosis factor.

† Results apply only to the 18 male patients.

‡ Results obtained within 4 weeks of the study visit were available for only 18 of the 30 patients.

Assessment Questionnaire (C-HAQ) (11), the patient's/parent's assessment of pain over the previous week (using a 10-cm visual analog scale [VAS], anchored at 0 = no pain and 10 = the most severe pain), the Bath Ankylosing Spondylitis Functional Index (BASFI) (12), and the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI) (13).

US examination. An ultrasound technologist or pediatric radiologist (NAC) with experience in musculoskeletal imaging performed all US examinations using a Philips IU22 machine with a high-frequency linear-array 12-MHz transducer (Figure 1). The sonographer, who was blinded to the clinical details, performed gray-scale and power Doppler evaluations at each insertion site using standardized imaging parameters (2). Gray-scale evaluation of the entheses included assessments of the appearance of the tendon and the presence or absence of bony erosions, enthesophytes, or bursitis. Tendon appearance was considered abnormal if there was loss of fibrillar pattern, regions of hypoechogenicity, or fusiform thickening. Bursitis of the knee was defined as abnormal anechoic or hypoechoic intraarticular material that distended the suprapatellar bursa at least 3 mm in the anteroposterior dimension. On real-time imaging, the material had to be compressible, displaceable, and without internal Doppler signal (14).

Power Doppler imaging was standardized with a repetition frequency of 500–750 Hz and gain adjusted to the highest level without background noise artifact (15). For each enthesis, the sonographer assessed power Doppler at the cortical bone insertion, in both the long and the transverse imaging planes, and graded the findings on a scale of 0–3, where 0 = absent, 1 = minimal (1 spot), 2 = moderate (2 spots), and 3 = severe (≥3 spots) (7). Since minimal power Doppler findings have been identified in normal children (7,16), we considered positive findings to be grade 2 or above. Enthesitis by US was defined as power Doppler findings of grade 2 or 3 at the cortical bone insertion, abnormal tendon appearance (loss of fibrillar pattern, regions of hypoechogenicity, or fusiform thickening), or structural abnormalities (calcification, enthesophytes, or erosions) (7). The complete US examination took an average of 45 minutes.

The medial epicondyles were evaluated with the arm in full extension. The lateral epicondyles were imaged with mild flexion of the elbow and pronation of the forearm. Examinations of the quadriceps tendon insertion and the patellar ligament insertion were performed with the patient in the supine position with the knee in 30° of flexion. The Achilles tendon and the plantar aponeurosis were examined with the patient lying prone with the feet hanging over the edge of the

examination table in neutral position. A stand-off pad was used to image the Achilles tendon. The same radiologist (NAC) interpreted all US images.

The intraobserver and interobserver reliability of US for the detection of enthesitis was estimated with a reference set of 64 images representing different patterns seen in the study population (ERA patients and control subjects). These images were deidentified, randomly ordered, and submitted to a second reading by the original examiner as well as by a second radiologist (DJ) ~6 months after the study visit. To avoid interpretation fatigue, the second readings were done in sets of 8 images, with a maximum of 1 set interpreted at any sitting. In addition to evaluating the patterns of agreement by site and reader, we estimated the kappa coefficient (17).

Physical examination. At the study visit, the same pediatric rheumatologist (PFW) conducted a complete joint examination and an assessment of the following entheses bilaterally, using a 22-lb–gauge dolorimeter (Baseline Dolorimeter; Fabrication Enterprises): common extensor tendon on the lateral humeral epicondyle, common flexor tendon on the medial humeral epicondyle, quadriceps at the superior patella, patellar ligament at the inferior patella, Achilles tendon, and plantar fascia at the calcaneus. Also examined were the thumbnail on the nondominant hand and the following 3 control sites bilaterally: mid-trapezius muscle, 2 cm distal to the lateral epicondyle, and 1 cm posterior to the greater trochanter.

Pressure was applied gradually at a steady rate (~0.5 kg/second) to each site using a 1.5-cm^2 metal tip fixed to the dolorimeter. Children were blinded to the amount of applied pressure. Once a subject verbally indicated pain, the pressure application was stopped and the pain threshold was recorded. The maximum amount of pressure applied was 4.5 kg; if the subject had not yet indicated pain, then the pain threshold was recorded as 4.5 kg. A second pediatric rheumatologist (DDS) repeated the examination at each site on 8 randomly selected ERA patients (96 sites) with the same dolorimeter to document the interobserver reliability of the clinical examination findings.

Peripheral arthritis was defined as joint swelling or, in the absence of swelling, limitation of motion accompanied by pain or tenderness. Sacroiliitis was defined as tenderness on direct compression over the joint or abnormal findings on imaging (9).

Statistical analysis. Demographic information and baseline characteristics were summarized by frequencies and percentages for categorical variables (e.g., sex, race, HLA–B27 status) and by mean, SD, median, and interquartile range (IQR) for continuous or count variables (e.g., number of sites with enthesitis). The sensitivity, specificity, PPV, and NPV of physical examination for the detection of US-confirmed enthesitis were calculated for each enthesis, using cutoff values for positivity and negativity ranging from 2.0 to 4.0 kg, with adjustment for the potential correlation of repeated measures within each subject. The reliability of the dolorimeter examination and the US examination were measured with kappa coefficients. According to Landis and Koch (18), a kappa value of <0.40 indicates poor agreement, 0.40–0.59 fair, 0.60–0.74 good, and 0.75–1.00 indicates excellent agreement. The 95% confidence intervals (95% CIs) for the kappa coefficients were calculated using percentiles obtained from bootstrap methods with 999 resamplings. We also calculated agreement between raters for both US and dolorimeter examination as described by Localio et al (19). The numerator of this statistic is the number of cases in which both reviewers documented abnormalities, and the denominator is the sum of the numerator and the number of cases where the reviewers disagreed. This statistic does not consider cases for which both raters agreed there were no abnormalities, but focuses on how often cases in which any rater finds disease are cases in which other, and perhaps all, raters find disease. Statistical analyses were performed using Stata software version 12.1 (StataCorp).

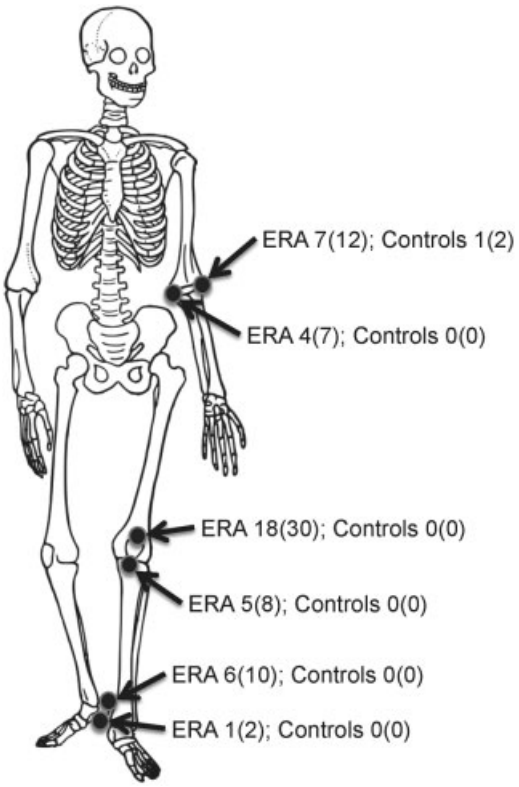

Figure 2. Sites of enthesitis identified by ultrasound (US). A total of 60 entheses were examined for each of the 6 sites in 30 patients with enthesitis-related arthritis (ERA) and 30 healthy control subjects. Enthesitis by US was defined as power Doppler grade 2 or 3 at the cortical bone insertion, abnormal tendon appearance (loss of fibrillar pattern, regions of hypoechogenicity, or fusiform thickening), or structural abnormalities (calcification, enthesophytes, or erosions). Values are the no. (%) of sites with enthesitis. Skeleton drawing reproduced from Wikimedia Commons; from a drawing originally published in Collier's new encyclopedia. Vol. VIII. New York: PF Collier & Son; 1921. p. 446.

RESULTS

ERA patients. In this study, we enrolled 30 patients with ERA and 30 control subjects. The median age of the ERA patients was 13 years (IQR 11–15 years). Sixty percent of the patients were male, and 30% were HLA–B27+ (Table 1). This was a relatively healthy ERA cohort, with a median BASDAI score for disease activity of 3.8 (IQR 2.6–6.4). Function was also good,

Table 2. Abnormal findings on ultrasound examination of children and adolescents with enthesitis-related arthritis*

Abnormal finding	No. (%) of tendon insertion sites with abnormal findings		
	Quadriceps	Common extensor	Achilles
Power Doppler grade ≥2 at cortical insertion	11 (18)	7 (12)	2 (3)
Fusiform tendon hypoechogenicity	5 (8)	0 (0)	4 (7)
Tendon thickening	4 (7)	0 (0)	0 (0)
Bursitis	5 (28)	1 (2)	2 (3)
Intratendinous calcifications	0 (0)	0 (0)	0 (0)
Enthesophytes	0 (0)	0 (0)	0 (0)
Bony erosions	0 (0)	0 (0)	0 (0)
Bony cortex irregularities	0 (0)	0 (0)	0 (0)

* A total of 60 entheses were examined in the 30 study patients. More than one abnormal finding may be present at the same site.

with median C-HAQ and BASFI scores of 0.31 (IQR 0.13–0.88) and 1.6 (IQR 0.7–4.9), respectively. The median pain score over the previous week, as reported by the patient/parent using a VAS, was 4 (IQR 2–7). Only 7% of the children with ERA (2 of 30 patients) reported no pain (VAS score 0) over the past week, in comparison to 93% of the healthy controls (28 of 30 subjects). On the day of the study visit, the median number of joints with active arthritis was 2. The most frequent joints with active arthritis were the sacroiliac joints (40 of 60 joints [67%]) and knees (9 of 60 joints [15%]). The hips, wrists, ankles, subtalar joints, midfoot joints, and the metatarsophalangeal, proximal interphalangeal, and distal interphalangeal joints of the fifth toe had active arthritis in 3% of the patients (2 of 60).

Ultrasound findings. *Control subjects.* Abnormal US power Doppler findings were detectable at 2% of the common extensor tendon insertions at the lateral humeral epicondyle (1 of 60 sites), and at none of the remaining insertion sites. Tendon appearance was normal in all control subjects at all sites examined.

ERA patients. Fifty-seven percent of the patients with ERA (17 of 30) had abnormal findings on US examination at 1 or more entheses, and 33% (10 of 30) at 2 or more entheses. The most frequently affected site was the quadriceps insertion at the superior patella (30% [18 of 60 sites]), common extensor (12% [7 of 60 sites]), and Achilles (10% [6 of 60 sites]) tendons; the least frequently affected site was the plantar fascia insertion at the calcaneus (2% [1 of 60 sites]). Figure 2 shows the distribution of US-confirmed sites with enthesitis in this pediatric cohort, and Table 2 lists the abnormalities identified on US examination. The abnormal US findings included increased vascularity, tendon hypoechogenicity, tendon thickening, and bursitis. Of the quadriceps insertions with US-confirmed enthesitis (n = 18), 11 sites (61%) had increased vascularity, 7 (39%) had abnormal tendon appearance, and 0 sites had both increased vascularity and abnormal tendon appearance. ERA patients did not have any findings of chronic damage from enthesitis, including intratendinous calcifications, enthesophytes, bony erosions, or bony cortex irregularities (Table 2).

Fluid within the suprapatellar bursa was present

Table 3. Findings of dolorimeter examination in ERA patients and control subjects*

	ERA patients		Control subjects	
	No. (%) of entheses with tenderness to 4 kg/cm^2 of pressure	Pain threshold, mean ± SD kg/cm^2	No. (%) of entheses with tenderness to 4 kg/cm^2 of pressure	Pain threshold, mean ± SD kg/cm^2
Tendon insertion sites				
Quadriceps	47 (78)	2.5 ± 1.2†	16 (27)	4.0 ± 0.9
Common extensor	33 (55)	3.5 ± 1.1†	11 (18)	4.2 ± 0.7
Achilles	21 (35)	4.0 ± 0.8†	3 (5)	4.4 ± 0.3
Patellar ligament	48 (80)	2.6 ± 1.2†	17 (28)	4.3 ± 0.6
Common flexor	31 (52)	3.5 ± 1.1†	10 (17)	4.2 ± 0.7
Plantar fascia	18 (30)	4.1 ± 0.7†	2 (3)	4.4 ± 0.4
Control sites				
Nondominant thumbnail	18 (60)	3.5 ± 1.0‡	8 (27)	4.1 ± 0.7
Mid-trapezius	46 (77)	2.9 ± 1.1†	16 (27)	4.1 ± 0.9
2 cm distal to the lateral epicondyle	46 (77)	3.0 ± 1.1†	21 (35)	4.1 ± 0.8
1 cm posterior to the greater trochanter	38 (63)	3.5 ± 0.9†	6 (10)	4.3 ± 0.6

* A total of 60 entheses (30 for the nondominant thumbnail) were evaluated in the 30 patients with enthesitis-related arthritis (ERA) as well as in the 30 healthy control subjects.
† $P < 0.001$ versus controls, by *t*-test.
‡ $P = 0.003$ versus controls, by *t*-test.

Table 4. PPVs and NPVs of findings on standardized dolorimeter examination, using ultrasound as the reference standard test*

Tendon insertion site, pressure applied	PPV (95% CI)	NPV (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Quadriceps tendon at the superior patella				
2.0 kg/1.5 cm^2	0.23 (0.04–0.43)	0.63 (0.43–0.84)	0.39 (0.11–0.66)	0.45 (0.26–0.64)
2.5 kg/1.5 cm^2	0.27 (0.09–0.45)	0.65 (0.40–0.91)	0.56 (0.25–0.86)	0.36 (0.17–0.54)
3.0 kg/1.5 cm^2	0.26 (0.10–0.42)	0.61 (0.32–0.90)	0.61 (0.31–0.91)	0.26 (0.10–0.43)
3.5 kg/1.5 cm^2	0.26 (0.10–0.41)	0.59 (0.28–0.89)	0.61 (0.31–0.91)	0.24 (0.07–0.40)
4.0 kg/1.5 cm^2	0.28 (0.12–0.43)	0.62 (0.28–0.95)	0.72 (0.49–0.95)	0.19 (0.03–0.35)
Common extensor tendon at the lateral humeral epicondyle				
2.0 kg/1.5 cm^2	0	0.87 (0.75–0.98)	0	0.13 (0.03–0.23)
2.5 kg/1.5 cm^2	0	0.85 (0.72–0.98)	0	0.26 (0.12–0.41)
3.0 kg/1.5 cm^2	0.04 (0.0–0.13)	0.84 (0.68–1.0)	0.14 (0.0–0.42)	0.42 (0.24–0.59)
3.5 kg/1.5 cm^2	0.07 (0.0–0.16)	0.84 (0.66–1.0)	0.29 (0.0–0.66)	0.51 (0.34–0.68)
4.0 kg/1.5 cm^2	0.06 (0.0–0.14)	0.81 (0.62–1.0)	0.29 (0.0–0.66)	0.58 (0.42–0.75)
Achilles tendon				
2.0 kg/1.5 cm^2	0	0.90 (0.81–0.99)	0	0.02 (0.0–0.05)
2.5 kg/1.5 cm^2	0.29 (0.0–0.65)	0.92 (0.85–1.00)	0.33 (0.02–0.65)	0.09 (0.0–0.18)
3.0 kg/1.5 cm^2	0.22 (0.0–0.51)	0.92 (0.85–1.00)	0.33 (0.02–0.65)	0.13 (0.03–0.23)
3.5 kg/1.5 cm^2	0.13 (0.0–0.32)	0.91 (0.83–1.00)	0.33 (0.02–0.65)	0.24 (0.10–0.39)
4.0 kg/1.5 cm^2	0.10 (0.0–0.23)	0.90 (0.80–1.00)	0.33 (0.02–0.65)	0.35 (0.19–0.51)
Patellar ligament at the inferior patella				
2.0 kg/1.5 cm^2	0	0.85 (0.70–0.99)	0	0.51 (0.35–0.67)
2.5 kg/1.5 cm^2	0	0.81 (0.63–0.98)	0	0.38 (0.23–0.54)
3.0 kg/1.5 cm^2	0	0.77 (0.57–0.97)	0	0.31 (0.18–0.44)
3.5 kg/1.5 cm^2	0	0.64 (0.34–0.94)	0	0.16 (0.05–0.28)
4.0 kg/1.5 cm^2	0.02 (0.0–0.06)	0.67 (0.33–1.0)	0.20 (0.0–0.57)	0.15 (0.03–0.26)
Common flexor tendon at the medial humeral epicondyle				
2.0 kg/1.5 cm^2	0.09 (0.0–0.27)	0.94 (0.85–1.0)	0.25 (0.0–0.72)	0.18 (0.05–0.30)
2.5 kg/1.5 cm^2	0.11 (0.0–0.26)	0.95 (0.89–1.0)	0.50 (0.15–0.85)	0.29 (0.13–0.44)
3.0 kg/1.5 cm^2	0.18 (0.0–0.39)	1.0	1.0	0.32 (0.16–0.48)
3.5 kg/1.5 cm^2	0.14 (0.0–0.31)	1.0	1.0	0.43 (0.25–0.60)
4.0 kg/1.5 cm^2	0.13 (0.0–0.28)	1.0	1.0	0.48 (0.30–0.66)
Plantar fascia at the calcaneus				
2.0 kg/1.5 cm^2	0	0.98 (0.95–1.00)	0	0.02 (0.0–0.05)
2.5 kg/1.5 cm^2	0	0.98 (0.95–1.00)	0	0.07 (0.01–0.13)
3.0 kg/1.5 cm^2	0	0.98 (0.94–1.00)	0	0.12 (0.03–0.21)
3.5 kg/1.5 cm^2	0	0.98 (0.94–1.00)	0	0.24 (0.09–0.39)
4.0 kg/1.5 cm^2	0	0.98 (0.93–1.00)	0	0.31 (0.16–0.45)

* The positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity, along with their 95% confidence intervals (95% CIs), were determined for each enthesis at the indicated pressure cutoff.

in 13% of the sites (8 of 60). Of the 18 sites with enthesitis at the quadriceps insertion, 28% (5 of 18) had accompanying bursitis. Fluid within the suprapatellar bursa was observed in the absence of enthesitis at 3% of the sites (2 of 60). In comparison, none of the controls had bursitis.

Reliability of US. The kappa coefficients (18) for the intrarater and interrater reliability of US for the evaluation of enthesitis were 0.78 (95% CI 0.63–0.93) and 0.81 (95% CI 0.67–0.95), respectively. There was also substantial agreement when each enthesis was assessed individually. At 9% of sites (6 of 64), one of the raters found disease while the other did not. Rater 1 rated 58% of the sites (37 of 64) as normal and 42% (27 of 64) as diseased. Rater 2 rated 48% of the sites (31 of 64) as normal and 52% (33 of 64) as diseased. For 82% of the sites at which a rater found disease, both raters noted this finding ($\kappa = 0.82$ [95% CI 0.65–0.93]).

Standardized dolorimeter examination. *Control subjects.* Using a dolorimeter cutoff value of 4 kg/cm^2 for positivity and negativity, 30% of the control subjects (10 of 30) had tenderness at 1 or more entheses, 23% (7 of 30) at 3 or more entheses, and 13% (4 of 30) at 6 or more entheses. Table 3 shows the distribution of entheses and control sites with tenderness in the ERA patients and control subjects. Tenderness in control subjects occurred most commonly at the quadriceps insertion (27% [16 of 60 sites]) and least commonly at the plantar fascia insertion (3% [2 of 60]). Tenderness at the control sites was less frequent in control subjects than in patients.

ERA patients. Using a dolorimeter cutoff value of 4 kg/cm^2 for positivity and negativity, 93% of the ERA

patients (28 of 30) had tenderness at 1 or more entheses, 87% (26 of 30) at 3 or more entheses, and 57% (17 of 30) at 6 or more entheses. Tenderness occurred most commonly at the patellar ligament (80% [48 of 60 sites]) and the quadriceps insertion (78% [47 of 60 sites]) and least commonly at the plantar fascia insertion (30% [18 of 60 sites]) (Table 3).

ERA patients reported tenderness significantly more often than did the control subjects at all entheses and at the 4 control sites (Table 3). ERA patients also had a significantly lower threshold of pain than did the controls at each enthesis and each control site ($P < 0.001$ for all sites) (Table 3). There was no significant difference in the threshold of pain at the quadriceps insertion among ERA patients with ($\kappa = 2.84$ [95% CI 2.21–3.48]) and those without ($\kappa = 2.40$ [95% CI 2.18–2.77]) US-confirmed enthesitis ($P = 0.90$).

Reliability of standardized dolorimeter examination. Using a dolorimeter cutoff value of 4 kg/cm^2 for positivity and negativity, there was 75% agreement between the 2 examiners on 8 randomly selected study subjects (96 sites). The interrater reliability of the dolorimeter examination for the evaluation of tenderness was moderate ($\kappa = 0.49$ [95% CI 0.15–0.78]). At 25% of the sites (24 of 96), one rater elicited tenderness while the other did not. Rater 1 rated 48% of the entheses (46 of 96) as normal and 52% (50 of 96) as tender. Rater 2 rated 27% of the entheses (26 of 96) as normal and 73% (70 of 96) as tender. For 67% of the sites at which a rater found tenderness, both raters had this finding ($\kappa = 0.67$ [95% CI 0.55–0.77]).

Standardized physical examination versus ultrasonography. The PPV and NPV of physical examination using dolorimeter pressures ranging from 2 kg/cm^2 to 4 kg/cm^2 for the detection of enthesitis at each enthesis are shown in Table 4. For the entheses with the highest prevalence of abnormalities identified on US (quadriceps, common extensor, and Achilles tendon insertions) a cutoff value of 4 kg/cm^2 for positivity and negativity gave the best PPV, NPV, sensitivity, and specificity. Determination of the optimum cutoff value for the other entheses is less clear given the low prevalence of abnormal findings at those sites.

DISCUSSION

In our relatively healthy cross-sectional ERA cohort, we found that US-confirmed enthesitis is common, particularly at the insertions of the quadriceps, the common extensor, and the Achilles tendons. US abnormalities included increased vascularity, tendon hypoechogenicity, tendon thickening, and bursitis. US with power Doppler, a reference standard for the detection of enthesitis in adults, had excellent interrater and intrarater reliability. In this study, tenderness detected by standardized dolorimeter examination had poor PPV for US-confirmed enthesitis. We tested several cutoff values for positivity and negativity for tenderness at each enthesis, ranging from 2 to 4 kg/1.5 cm^2, with some improvement in the sensitivity and specificity, but minimal improvement in the PPV and NPV, at the higher pressures. In comparison to US, physical examination with a dolorimeter had poor accuracy and reliability. Additionally, in comparison to healthy controls, children with ERA reported pain significantly more often and had lower mean pain thresholds at every site examined, including the control sites.

These results are consistent with those of previous studies exploring the use of US for the detection of enthesitis in patients with JIA. Jousse-Joulin and colleagues (7) used US to evaluate the insertions of the quadriceps, proximal and distal patellar tendons, Achilles tendon, and plantar fascia in 26 consecutive patients with JIA, 9 (35%) of whom had ERA. US-confirmed enthesitis was found in 12.5% of entheses in the JIA group, with enthesitis occurring most commonly at the insertions of the distal patellar, Achilles, quadriceps, and proximal patellar ligament; overall, the sensitivity of physical examination for the detection of US-confirmed disease was 0.50 (7). Another study evaluated the proximal gluteus medius insertion at the posterior iliac crest in 38 children with JIA (71% with ERA); the PPV of physical examination for the detection of US-confirmed enthesitis was 0.53 (8).

Our findings also suggest that in comparison to US, physical examination with a dolorimeter is a poor diagnostic test for enthesitis. The intrarater and interrater reliability of US for the evaluation of enthesitis was substantial. Compared to US, the interrater reliability of the standardized dolorimeter examination was poor. The reliability of physical examination is dependent not only upon consistent performance of the research protocol dolorimeter examination by the rater, but also upon consistent pain reporting by the subject. Other factors that may have affected the reliability of the physical examination include the order in which the 2 examiners performed the dolorimeter examination, small differences in the speed with which pressure was added with the dolorimeter, and each subject's variable and possibly inconsistent threshold for pain. If the interrater reliability of a research protocol standardized dolorimeter examination is low, it is quite likely that the interrater reliability of multiple physicians simply pressing on entheses is much worse.

For a condition such as enthesitis, in which pain is the main physical examination finding that is considered indicative of active inflammation, a more objective measure, such as US, may be extremely useful. This is particularly true at the time of diagnosis, when the presence of enthesitis is needed to ensure accurate JIA classification and to monitor response to therapy over time, especially since prior studies have demonstrated an unclear relationship between the presence of pain, the pain thresholds, and the disease activity (20,21).

Our results should be interpreted in the context of several limitations. To our knowledge, this is the first study to evaluate the use of a standardized examination and US for the detection of enthesitis in a cohort of children with ERA and the largest study of US and enthesitis in JIA to date. However, our sample size is limited, with only 30 ERA patients and 30 control subjects. Additional studies with larger sample sizes are needed to confirm our findings.

Second, this study examined the accuracy and reliability of one examination method, dolorimetry. We made every attempt to standardize this examination, including use of the same dolorimeter for both raters. Our findings may not be generalizable to other physical examination methods. There is also a risk for spectrum and/or selection bias, as our institution is a highly specialized, tertiary care medical center with specialists in the care of inflammatory arthritis. Nevertheless, our institution treats patients with wide variations in symptoms and disease severity, and thus, our patients should represent the full breadth of cases that might present across many institutions.

Last, our definition of US-defined enthesitis was power Doppler findings of grade 2 or above, abnormal tendon appearance (loss of fibrillar pattern, regions of hypoechogenicity, or fusiform thickening), or structural abnormalities (calcification, enthesophytes, or erosions). This definition is consistent with those of previous studies of enthesitis in JIA and adult spondyloarthritis (2,7,8). However, the published Outcome Measures in Rheumatology (OMERACT) definition of enthesopathy is abnormal tendon appearance, with or without power Doppler findings (22). Heterogeneity in the definition of US-defined enthesitis clearly exists (23), and future work should address development of standardized definitions and validation of these definitions in the adult spondyloarthritis and the pediatric JIA/spondyloarthritis populations.

Three major findings from this study warrant further discussion. First, our findings highlight some differences between enthesitis in children and adults. In adults, the most common sites of enthesitis are the insertions of the Achilles, plantar fascia, and patellar ligaments. US-confirmed enthesitis at the Achilles tendon insertion is detectable in 79–96% (2,24) of adult entheses, as compared to only 10% in our cohort. Findings of enthesitis at the insertions of the patellar ligament and plantar fascia were rare in our cohort of children, 8% and 2%, respectively. In comparison to findings in adults, our cohort had no US findings of chronic damage from enthesitis, including intratendinous calcifications, enthesophytes, bony erosions, or bony cortex irregularities. These differences are most likely secondary to the differences in disease duration between the published adult cohorts (10–16 years) (2,24) and our pediatric cohort (0.6 years).

Second, in comparison to US, physical examination with a dolorimeter had low accuracy and reliability for the detection of enthesitis. Enthesitis, as defined by tenderness on examination, is significantly and independently associated with increased pain intensity in children with ERA (25) and has a major impact on function and quality of life in adults with ankylosing spondylitis (26). However, our results suggest that tenderness on physical examination with a dolorimeter is not helpful in determining which entheses are actually inflamed and will require monitoring over time for damage in the form of calcifications, enthesophytes, erosions, and bony cortex irregularities or, conversely, for resolution of disease.

The low accuracy of physical examination for the detection of US-confirmed enthesitis raises a fundamental question regarding the diagnostic criteria for ERA, as clinically defined enthesitis is one of the major criteria. It is possible that children with other categories of JIA and tenderness at one or more entheses but without any evidence of inflammation on imaging are being assigned inappropriately into the ERA category. This finding has implications for treatment decisions and long-term monitoring for comorbidities. At our institution, physicians have a lower threshold at which to initiate an anti–tumor necrosis factor (anti-TNF) medication in children with ERA versus the other categories of JIA. Monitoring practices for comorbidities, such as uveitis and sacroiliitis, also differ in children with ERA versus the other JIA categories. Additionally, children with true ERA may be misclassified as not having ERA if they do not report tenderness on examination.

Third, ERA patients in this study had a significantly lower threshold of pain than did the control subjects at each enthesis and control site, and there was no significant difference in the threshold of pain at sites with and those without US-confirmed enthesitis. It is also possible that these patients have areas that are so subtly inflamed that they feel tenderness, but we cannot

yet detect inflammation. Several studies have reported decreased pain thresholds in children with chronic pain conditions (27–29) and JIA (20,21). These lowered pain thresholds are present not only at sites directly affected by disease, but also at sites presumably unaffected by disease (20,27,29). Studies have demonstrated that adults with rheumatoid arthritis have increased activity in areas of the brain implicated in pain perception and control of emotions (30), supporting a theory of altered central pain perception and sensitization (31,32). In comparison to rheumatoid arthritis patients with low central nervous system pain activity, those with elevated pain activity in the brain respond better to TNF inhibition (33). Further, the CNS response to TNF inhibition occurred before detectable objective clinical measures of disease activity. These data suggest that the patient's perception of disease and pain may have a strong influence on responsiveness to antiinflammatory therapy (33).

There is also some evidence of the role of substance P–immunoreactive nerve fibers and naked, myelinated, and unmyelinated fibers in the development of pain in the entheses (34). The lower pain threshold at the entheses and control sites in ERA patients in this study almost certainly contributes to the low PPV of the physical examination for the detection of US-confirmed enthesitis. Future evaluation of changes in pain perception and threshold in response to biologic treatment in children with ERA may be indicated.

In summary, the presence of enthesitis (and whether it is defined by physical examination or US) has potential implications with regard to JIA classification, which in turn, influences treatment decisions and monitoring for extraarticular manifestations of disease. US had excellent intrarater and interrater reliability for the detection of enthesitis. Compared to US, standardized dolorimeter examination for the detection of enthesitis in children had poor accuracy and reliability. Future research is warranted concerning the utility of US for identifying enthesitis at the time of JIA diagnosis, accurately predicting disease progression, and guiding therapeutic decisions.

ACKNOWLEDGMENTS

We thank the network of primary care clinicians and their patients and families for their contribution to this project and clinical research facilitated through the Pediatric Research Consortium at the Children's Hospital of Philadelphia. We would also like to thank Allyson Gutstein and Jenna Tress for their assistance with recruitment of subjects and Colden Broderick for her assistance with data entry.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be published. Dr. Weiss had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Study conception and design. Weiss, Chauvin, Localio, Feudtner, Jaramillo, Colbert, Sherry, Keren.

Acquisition of data. Weiss, Chauvin, Klink, Jaramillo, Sherry.

Analysis and interpretation of data. Weiss, Chauvin, Klink, Localio, Feudtner, Jaramillo, Colbert, Keren.

REFERENCES

1. Sherry DD, Sapp LR. Enthesalgia in childhood: site-specific tenderness in healthy subjects and in patients with seronegative enthesopathic arthropathy. J Rheumatol 2003;30:1335–40.
2. D'Agostino MA, Said-Nahal R, Hacquard-Bouder C, Brasseur JL, Dougados M, Breban M. Assessment of peripheral enthesitis in the spondylarthropathies by ultrasonography combined with power Doppler: a cross-sectional study. Arthritis Rheum 2003;48: 523–33.
3. Sprott H, Jeschonneck M, Grohmann G, Hein G. Microcirculatory changes over the tender points in fibromyalgia patients after acupuncture therapy (measured with laser-Doppler flowmetry). Wien Klin Wochenschr 2000;112:580–6. In German.
4. D'Agostino MA, Breban M, Brasseur JL, Canale MP, Magaro M. Ultrasonographic (US) assessment of spondyloarthropathy (SpA) enthesitis, displays vascularisation change specific of inflammatory process [abstract]. Arthritis Rheum 2001;44 Suppl:S95.
5. Aydin SZ, Karadag O, Filippucci E, Atagunduz P, Akdogan A, Kalyoncu U, et al. Monitoring Achilles enthesitis in ankylosing spondylitis during TNF-α antagonist therapy: an ultrasound study. Rheumatology (Oxford) 2010;49:578–82.
6. Balint PV, Kane D, Wilson H, McInnes IB, Sturrock RD. Ultrasonography of entheseal insertions in the lower limb in spondyloarthropathy. Ann Rheum Dis 2002;61:905–10.
7. Jousse-Joulin S, Breton S, Cangemi C, Fenoll B, Bressolette L, de Parscau L, et al. Ultrasonography for detecting enthesitis in juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2011; 63:849–55.
8. Laurell L, Court-Payen M, Nielsen S, Zak M, Thomsen C, Miguel-Perez M, et al. Ultrasonography and color Doppler of proximal gluteal enthesitis in juvenile idiopathic arthritis: a descriptive study. Pediatr Rheumatol Online J 2011;9:22.
9. Petty RE, Southwood TR, Manners P, Baum J, Glass DN, Goldenberg J, et al. International League of Associations for Rheumatology classification of juvenile idiopathic arthritis: second revision, Edmonton, 2001. J Rheumatol 2004;31:390–2.
10. Filocamo G, Schiappapietra B, Bertamino M, Pistorio A, Ruperto N, Magni-Manzoni S, et al. A new short and simple health-related quality of life measurement for paediatric rheumatic diseases: initial validation in juvenile idiopathic arthritis. Rheumatology (Oxford) 2010;49:1272–80.
11. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum 1994;37:1761–9.
12. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Spondylitis Functional Index. J Rheumatol 1994;21:2281–5.
13. Garrett S, Jenkinson T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis: the Bath Ankylosing Spondylitis Disease Activity Index. J Rheumatol 1994;21:2286–91.
14. Vlad V, Iagnocco A. Ultrasound of the knee in rheumatology. Med Ultrason 2012;14:318–25.

15. Iagnocco A, Riente L, Delle Sedie A, Filippucci E, Salaffi F, Meenagh G, et al. Ultrasound imaging for the rheumatologist. XXII. Achilles tendon involvement in spondyloarthritis: a multicentre study using high frequency volumetric probe. Clin Exp Rheumatol 2009;27:547–51.
16. Toib D, French AR, Dahiya N, Middleton WD, Brasington R, White AJ. The role of ultrasound in B mode and power Doppler sonography in the diagnosis of enthesitis in children [abstract]. Arthritis Rheum 2010;62 Suppl 10:S680.
17. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas 1960;20:37–46.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159–74.
19. Localio AR, Weaver SL, Landis JR, Lawthers AG, Brenhan TA, Hebert L, et al. Identifying adverse events caused by medical care: degree of physician agreement in a retrospective chart review. Ann Intern Med 1996;125:457–64.
20. Leegaard A, Lomholt JJ, Thastum M, Herlin T. Decreased pain threshold in juvenile idiopathic arthritis: a cross-sectional study. J Rheumatol 2013;40:1212–7.
21. Hogeweg JA, Kuis W, Huygen AC, de Jong-de vos van Steenwijk C, Bernards AT, Oostendorp RA, et al. The pain threshold in juvenile chronic arthritis. Br J Rheumatol 1995;34:61–7.
22. Wakefield RJ, Balint PV, Szkudlarek M, Filippucci E, Backhaus M, D'Agostino MA, et al. Musculoskeletal ultrasound including definitions for ultrasonographic pathology. J Rheumatol 2005;32: 2485–7.
23. Gandjbakhch F, Terslev L, Joshua F, Wakefield RJ, Naredo E, D'Agostino MA. Ultrasound in the evaluation of enthesitis: status and perspectives. Arthritis Res Ther 2011;13:R188.
24. Kiris A, Kaya A, Ozgocmen S, Kocakoc E. Assessment of enthesitis in ankylosing spondylitis by power Doppler ultrasonography. Skeletal Radiol 2006;35:522–8.
25. Weiss P, Beukelman T, Schanberg LE, Kimura Y, Colbert RA, CARRAnet Investigators. Enthesitis is a significant predictor of decreased quality of life, function, and arthritis-specific pain across juvenile arthritis (JIA) categories: preliminary analyses from the CARRAnet registry [abstract]. Arthritis Rheum 2011;63 Suppl: S105.
26. Maksymowych WP, Mallon C, Morrow S, Shojania K, Olszynski WP, Wong RL, et al. Development and validation of the Spondyloarthritis Research Consortium of Canada (SPARCC) Enthesitis Index. Ann Rheum Dis 2009;68:948–53.
27. Buskila D, Press J, Gedalia A, Klein M, Neumann L, Boehm R, et al. Assessment of nonarticular tenderness and prevalence of fibromyalgia in children. J Rheumatol 1993;20:368–70.
28. Duarte MA, Goulart EM, Penna FJ. Pressure pain threshold in children with recurrent abdominal pain. J Pediatr Gastroenterol Nutr 2000;31:280–5.
29. Hashkes PJ, Friedland O, Jaber L, Cohen HA, Wolach B, Uziel Y. Decreased pain threshold in children with growing pains. J Rheumatol 2004;31:610–3.
30. Hess A, Axmann R, Rech J, Finzel S, Heindl C, Kreitz S, et al. Blockade of TNF-a rapidly inhibits pain responses in the central nervous system. Proc Natl Acad Sci U S A 2011;108:3731–6.
31. Reinold H, Ahmadi S, Depner UB, Layh B, Heindl C, Hamza M, et al. Spinal inflammatory hyperalgesia is mediated by prostaglandin E receptors of the EP2 subtype. J Clin Invest 2005;115:673–9.
32. Lee YC, Nassikas NJ, Clauw DJ. The role of the central nervous system in the generation and maintenance of chronic pain in rheumatoid arthritis, osteoarthritis and fibromyalgia. Arthritis Res Ther 2011;13:211.
33. Rech J, Hess A, Finzel S, Kreitz S, Sergeeva M, Englbrecht M, et al. Association of brain functional magnetic resonance activity with response to tumor necrosis factor inhibition in rheumatoid arthritis. Arthritis Rheum 2013;65:325–33.
34. Palesy PD. Tendon and ligament insertions—a possible source of musculoskeletal pain. Cranio 1997;15:194–202.