

A Novel Algorithm Using Within-leg Calibration for Enhanced Accuracy of Detection of Arthritis by Infrared Thermal Imaging in Children

Yongdong Zhao¹ , Ramesh S. Iyer², Mahesh Thapa², Debosmita Biswas³, Nivrutti Bhide⁴, Joshua Scheck¹, Kevin Cain⁵, Savannah C. Partridge³, and Carol A. Wallace⁶

ABSTRACT. *Objective.* To standardize and improve the accuracy of detection of arthritis by thermal imaging.

Methods. Children with clinically active arthritis in the knee or ankle, as well as healthy controls, were enrolled to the development cohort; another group of children with knee symptoms was enrolled to the validation cohort. Ultrasound was performed in the arthritis subgroup for the development cohort. Joint exam by certified rheumatologists was used as a reference for the validation cohort. Infrared thermal data were analyzed using custom software. Temperature after within-limb calibration (TAWiC) was defined as the temperature differences between joint and ipsilateral mid-tibia. TAWiC of knees and ankles was evaluated using ANOVA across subgroups. Optimal thresholds were determined by receiver-operating characteristic analysis using Youden index.

Results. There were significant differences in mean and 95th TAWiC of knee in anterior, medial, lateral views, and of ankles in anterior view, between inflamed and uninflamed counterparts ($P < 0.05$). The area under the curve was higher by 30% when using TAWiC_{knee} than that when using absolute temperature. Within the validation cohort, the sensitivity of accurate detection of arthritis in the knees using both mean and 95th TAWiC from individual views or all 3 views combined ranged from 0.60 to 0.70, and the specificity was > 0.90 in all views.

Conclusion. Children with active arthritis or tenosynovitis in knees or ankles exhibited higher TAWiC than healthy joints. Our validation cohort study showed promise for the clinical utility of infrared thermal imaging for arthritis detection.

Key Indexing Terms: ankle, infrared thermal imaging, juvenile idiopathic arthritis, knee, ultrasound

Juvenile idiopathic arthritis (JIA) is the most common rheumatic disease in children.¹ The most commonly affected joints are knees and ankles, followed by wrists and elbows.² Early diagnosis and aggressive treatment are critical for maintaining normal joint

functions in the management of JIA.³ A joint exam performed by a pediatric rheumatologist is considered standard assessment for children with JIA. Musculoskeletal ultrasound (US) is more sensitive for joint synovitis than a physical exam but may also be limited by accessibility to equipment and by operators.⁴

Infrared thermal imaging is a quick and noninvasive tool that can detect temperatures of different body parts with precision. It has been evaluated as a screening or supplementary tool for detecting or following up active arthritis in animal models,^{4,5} osteoarthritis,^{6,7} rheumatoid arthritis,^{8,9} and JIA.^{10,11,12} Studies focusing on larger joints (knees, ankles, wrists)^{8,11,12} defined regions of interest (ROIs) based on anatomic location and reported absolute temperatures for comparison. Lasanen et al showed significantly higher temperatures in inflamed ankles than controls but failed to confirm the difference between inflamed and healthy knee joints.¹¹ Heat distribution index was reported as another approach with a cut-off of 1.3°C to distinguish active arthritis in finger joints and wrists with a sensitivity of 67% and a specificity of 100%.⁸ Ilowite et al¹² used the difference between inflamed joints and adjacent tissues in patients with symmetric arthritis. The temperature difference was associated with disease activity score.¹² However, “adjacent tissue” was not clearly defined in that paper.

This study was supported by Clinical Research Scholar Program and Pilot Study Grant from Seattle Children's Research Institute. The research of YZ is supported by Bristol Myers Squibb, American College of Rheumatology, European Alliance of Associations for Rheumatology, Washington Research Foundation, and CARRA.

¹Y. Zhao, MD, PhD, J. Scheck, BS, Pediatric Rheumatology, Seattle Children's Hospital, Department of Pediatrics, University of Washington, and Center for Clinical and Translational Research, Seattle Children's Research Institute; ²R.S. Iyer, MD, M. Thapa, MD, Department of Radiology, Seattle Children's Hospital; ³D. Biswas, MS, S.C. Partridge, PhD, Department of Radiology, University of Washington; ⁴N. Bhide, Center for Clinical and Translational Research, Seattle Children's Research Institute; ⁵K. Cain, PhD, Department of Statistics, School of Nursing, University of Washington; ⁶C.A. Wallace, MD, Pediatric Rheumatology, Seattle Children's Hospital, Department of Pediatrics, University of Washington, Seattle, Washington, USA.

Address correspondence to Dr. Y. Zhao, MA 7.110, 4800 Sand Point Way NE, Seattle, WA 98105, USA. Email: Yongdong.zhao@seattlechildrens.org
Accepted for publication June 11, 2021.

Our group has developed a standardized ROI definition approach in lower extremities for analysis of thermal images from children with chronic nonbacterial osteomyelitis.¹³ Our objectives were to (1) enhance the sensitivity and reliability of detecting arthritis by using within-limb calibration for thermal imaging analysis; (2) determine the threshold using within-limb calibration in children with known arthritis; and (3) validate this approach in new patients.

METHODS

Institutional review board approval (#15350, #1383) was obtained from the authors' tertiary-care, multidisciplinary pediatric hospital prior to the study. For the development cohort, 2 groups, including children with clinically confirmed arthritis in the knee or ankle, and healthy children between the ages of 2 and 18 years, were consented and enrolled. Inclusion criteria for the arthritis group were active arthritis in the knee and/or ankle diagnosed by treating physician (swelling, or pain and limitations in motion if there was no swelling). Inclusion criterion of the healthy control (HC) group was normal skeletal health. Exclusion criteria for both groups were as follows: (1) skin infection in imaged area that could interfere with thermal imaging results; (2) fever; (3) joint contracture > 10°; (4) inability to cooperate with the acquisition of thermal imaging; and (5) recent injury to the areas of interest. For the validation cohort, children with knee pain and/or swelling for at least a week, who sought care from rheumatology for the first time, were enrolled. The same exclusion criteria were applied as for the development cohort.

Thermal imaging acquisition. As previously described,¹³ all subjects received infrared thermal imaging analysis of the lower limbs from 4 views (anterior, posterior, medial, and lateral). Thermal imaging was performed using a Fluke TiR32 Thermal Imager (Fluke Inc.) with 76,800 pixels (320 × 240; detection range -20 to 150 °C, sensitivity ≤ 0.04 °C) by trained staff to ensure sharp focus and consistent camera leveling and stabilization. The entire imaging session for each patient took < 5 minutes. Subjects exposed their feet and entire legs to room air and rested for at least 10 minutes prior to imaging to allow stabilization and equilibration of skin temperature. Ambient temperature was set at 22.2 °C for all patients. Subjects posed in standardized positions to ensure consistency of image acquisition. Imaging was performed with subjects standing on a carpet to avoid influence from the cold floor on body temperature, and away from potentially interfering items such as metal panels, doorknobs, computer screens, and adjacent people. Camera was positioned at the knee level of the subjects. The distance between camera and subject ranged between 4–5 m in order to maximize the spatial resolution of imaged body parts.

US imaging acquisition. Only subjects from the JIA group within the development cohort were scanned with US. Standard B-mode views of the knees (longitudinal and transverse suprapatellar, transverse posterior) with 20–30° of flexion, tibiotalar joints (anterior longitudinal and transverse, medial, and lateral paramalleolus) with plantar flexion, subtalar joints (lateral longitudinal) in neutral position, without compression were collected after thermal imaging in the arthritis group by 1 pediatric rheumatologist (YZ) with Ultrasound School of North American Rheumatologists (USSONAR) certification and 5 years of experience using the GE LOGIQ e US machine (General Electronics Inc.). Matching joint examinations were performed on the same day before US images were obtained.

Subjects within the validation cohort were not scanned by US because the goal of applying this thermal imaging tool is to identify patients from the community to accelerate the referral process, and a joint exam performed by a certified rheumatologist remains the well-accepted standard clinical practice.

Analysis of thermal images. The spatial and temperature data from infrared thermal images were exported from Smartview software (Fluke Inc.). Data were then analyzed using customized semiautomated software developed

in MATLAB (MathWorks) as previously reported.¹³ In brief, lower legs were divided equally into 3 segments (proximal, mid, and distal) longitudinally by placing crosshairs at the medial and lateral sides of the knees and ankles from each view, and distal femur was defined as the same length as the proximal tibia/fibula segment. Then, the proximal tibia/fibula and distal femur segments were merged as the ROI for “knee.” Using the distal tibia/fibula length as a reference, one-third of the reference above the ankle line and one-ninth of the reference below the ankle line were merged as “ankle” for thermal imaging analysis that included tibiotalar and subtalar joints. Mean and 95th percentile temperatures were recorded for each leg or joint segment. A previous study showed high reproducibility for this technique (intraclass correlation coefficient 0.936–0.981).¹³

Temperature after within-limb calibration (TAWiC) was calculated as the summary measure (mean or 95th percentile) for the joint (knee or ankle), minus the summary measure for mid-tibia. Thus, TAWiC measures how much hotter the joint is than the mid-tibia of the same limb.

US reading. A small set of US images from previous patients was reviewed by 2 radiologists (RSI and MT) and a rheumatologist (YZ) for calibration purposes. US images were scored as a consensus between 2 pediatric musculoskeletal radiologists (RSI, MT). When bone and tendon landmarks were not well visualized, images were excluded. A joint effusion was defined as anechoic material within the joint space or within the suprapatellar bursa (knee), or that displaced a fat pad in the tibiotalar and subtalar joints, as previously reported.^{14,15} Grading of joint effusion was performed as previously published.^{15,16} Synovial thickening was defined as hypoechoic material within the joint space that was not compressible. Tenosynovitis was defined as anechoic or hypoechoic material within the tendon sheath that circumscribed the tendon. Presence or absence of these variables was recorded. Arthritis was defined as the presence of synovial thickening, or at least a moderate effusion without synovial thickening. Since it was difficult to distinguish tibiotalar and subtalar joints on thermal imaging, these are combined: the “ankle” was considered to be inflamed if either tibiotalar or subtalar joint (or both) was inflamed, or isolated tenosynovitis was present.

Demographic, clinical, and laboratory data collection. Demographic information including sex, age, ethnicity, and race, and clinical data including body height, weight, and oral or temporal temperatures were collected in all subjects. Within the JIA group in the development cohort and new patients in the validation cohort, the presence or absence of joint swelling, pain or warmth, physician global assessment (0–10), Childhood Health Assessment Questionnaire score (0–3), patient/parent assessment of arthritis activity (0–10), patient/parent assessment of overall health (0–10), and current medications were recorded. Laboratory data were also collected if available.

Data analysis. Histograms were examined for outliers and nonnormality. Demographic variables were summarized and compared between children with JIA and healthy subjects using chi-square tests for categorical measures and *t* tests or Mann-Whitney *U* tests for numerical measures, depending on whether the measure was approximately normally distributed. Generalized estimating equations analysis was used to compare inflamed to uninflamed joints, while accounting for the fact that the 2 joints within a child were not independent observations and using the sandwich estimator of standard error, which is robust to nonnormality. Absolute temperatures and TAWiC were dependent variables; the predictor of interest was whether or not the joint was inflamed. Analyses were done separately for each view. Receiver-operating characteristic (ROC) curve analyses were used to describe how well the different summary measures can predict whether a joint is inflamed. Optimal thresholds were determined by ROC analysis using Youden index; these were then applied to the validation cohort. Sensitivity and specificity of detecting knee arthritis in the validation cohort was determined using derived thresholds. Pearson correlation was used to describe the association between TAWiC 95 and demographic measures of sex, age, height, weight, and BMI. A *P* value < 0.05 was considered statistically significant. All analyses were done using IBM SPSS version 19 (IBM Corp.).

A conservative power analysis showed that a sample size of 25 subjects per group would give over 90% power for detecting group differences as long as the true standardized effect size (difference in means divided by within-group SD) was at least 1.0. This effect size corresponds approximately to sensitivity of 70% and specificity of 70%. Since a measure with sensitivity and specificity smaller than this would not be useful clinically, this study has adequate power for detecting any clinically useful difference.

RESULTS

Demographic characteristics. Fifty-three children from the JIA group, 49 from the HC group from the development cohort, and 43 children with knee symptoms from the validation

cohort were enrolled. Fifty-one children within the JIA group completed US examinations and had evaluable thermal imaging and were included in the analysis. Forty-eight children from the HC group had evaluable thermal imaging and were included in the analysis. Patient characteristics from each group were summarized and compared in Table 1. There was no statistically significant difference in demographic characteristics between JIA and control groups. Within the JIA group, the mean disease duration was 3 years, and a majority of subjects were not on systemic medications. The majority of patients with JIA (65%) in the development cohort were categorized as oligoarticular.

Table 1. Baseline characteristics of children in the JIA group and control group.

	Development Cohort		P	Validation Cohort
	JIA, n = 51	HCs, n = 48		Knee Symptoms, n = 43
Age at enrollment, yrs	9.0 ± 4.1	9.7 ± 3.4	0.40	10.7 ± 4.1
Disease duration at enrollment, yrs	3.0 ± 3.6	NA		NA
Female	33 (65)	33 (69)	0.42	27 (63)
Weight, kg	36.6 ± 21.4	36.2 ± 18.7	0.91	43.8 ± 18.9
Height, cm	133.1 ± 25.6	137.9 ± 20.4	0.31	144.2 ± 22.1
BMI	18.9 ± 4.7	17.9 ± 4.1	0.26	17.9 ± 4.1
Oral temperature, °C	36.8 ± 0.4 (n = 15)	36.9 ± 0.3 (n = 37)	0.41	37.0 ± 0.2
Race			0.06	
White	39 (77)	35 (73)		36 (84)
African American	3 (6)	3 (6)		1 (2)
Asian	2 (4)	9 (19)		0 (0)
Native American	0 (0)	1 (2)		1 (2)
Pacific islander	1 (2)	0 (0)		0 (0)
Other	6 (12)	0 (0)		5 (12)
ILAR category				
Persistent oligoarticular	29 (57)	NA		3 (7)
Extended oligoarticular	4 (8)	NA		0 (0)
RF-negative polyarthritis	9 (18)	NA		1 (2)
RF-positive polyarthritis	1 (2)	NA		1 (2)
Enthesitis-related arthritis	3 (6)	NA		0 (0)
Psoriatic JIA	4 (8)	NA		0 (0)
Systemic JIA	1 (2)	NA		0 (0)
Undifferentiated JIA	0 (0)	NA		1 (2)
Laboratory findings				
ANA positive	25 (49), (n = 45)	NA		12 (48), (n = 25)
RF positive	3 (6), (n = 38)	NA		1 (5), (n = 19)
CCP positive	1 (2), (n = 33)	NA		2 (15), (n = 13)
HLA-B27 positive	2 (4), (n = 33)	NA		4 (31), (n = 13)
ESR, mm/h (normal 0–20; n = 34)	25.8 ± 26.0	NA		NA
Patient/parent-reported measures				
Patient/parent global assessment of overall health (0–10)	2.3 ± 2.3 (n = 48)	NA		NA
Patient/parent assessment of arthritis activity (0–10)	3.9 ± 2.8 (n = 48)	NA		NA
CHAQ	0.5 ± 0.5 (n = 50)	NA		0.8 ± 0.6 (n = 43)
PGA (range 0–10)	2.6 ± 1.5 (n = 40)	NA		NA
Treatment at study entry				
NSAIDs	18 (35)	NA		27 (63)
DMARDs	15 (29)	NA		NA
Biologics	8 (16)	NA		NA
Systemic glucocorticoids	4 (8)	NA		1 (2)

Values are expressed as mean ± SD or n (%). * Statistics between 2 groups within development cohort. ANA: antinuclear antibody; CCP: cyclic citrullinated peptide; CHAQ: Childhood Health Assessment Questionnaire; DMARD: disease-modifying antirheumatic drug; ESR: erythrocyte sedimentation rate; HC: healthy control; ILAR: International League of Associations for Rheumatology; JIA: juvenile idiopathic arthritis; NA: not applicable; NSAID: nonsteroidal antiinflammatory drug; PGA: physician global assessment; RF: rheumatoid factor.

Physical exam and US results. Active arthritis on joint exam was defined as pain of motion (POM) plus limitation of motion (LOM), or swelling, for knee and ankle; and tenderness and POM or tenderness and LOM for subtalar joint.¹⁷ Active arthritis on US was based on the presence of synovial thickening with or without effusion, or at least moderate effusion if without synovial thickening for all 3 joints. Tenosynovitis around ankle and subtalar joints was classified as “inflammation of ankle” on US. Within the JIA group, 49 (48%) knee, 24 (22%) tibiotalar, and 15 (15%) subtalar joints had active arthritis on physical examination. Meanwhile, 45 (44%) knee, 15 (14%) tibiotalar, 11 (11%) subtalar joints had active arthritis and 8 (8%) ankle joints had active tenosynovitis on US. The final count of inflammatory knees was 45, and number of inflammatory ankles was 19. A total of 11 joints (knees or ankles) from the development cohort were excluded from the analysis due to physical exam findings of arthritis but a normal US. Among 43 children with knee complaints within the validation cohort, 7 patients had arthritis in a total of 10 knees, whereas only 3 patients had arthritis in 5 ankles (tibiotalar and/or subtalar joint) determined by physical exam alone.

Performance of thermal imaging analysis for detecting arthritis in development cohort. Within the development cohort, all joint segments (knee and ankle) were divided into 3 groups: the HC

group, joints in the JIA group with inflammation, and joints in the JIA group without inflammation. Joints that were classified as arthritis by joint exam but not confirmed by US were excluded (for results when joint exam was the only confirmation of arthritis, see Supplementary Tables 1–3, available with the online version of this article). Preliminary analyses showed little difference between uninflamed joints in children with JIA and in HC children, so these 2 groups were combined into 1 group for all analyses, referred to as the uninflamed joint group. Figure 1 revealed a representative patient with thermal image, absolute temperatures and TAWiC of knees, ankles, and mid-tibia, as well as corresponding US findings confirming active arthritis in a knee and a tibiotalar joint. Table 2 shows means and SDs of the absolute and calibrated temperature summaries by inflammation status of joints from the development cohort. The size of ROI showed trends of increase in inflamed limbs but no statistically significant difference (Supplementary Table 4).

Comparison of area under the curve using TAWiC vs absolute temperature. In general, absolute and TAWiC temperatures were higher in inflamed knees and ankles than in uninflamed counterparts. Compared to absolute values, TAWiC showed a greater temperature difference between groups, with smaller SD within each group and more significant *P* values. Posterior view showed a considerably smaller difference between groups than did the

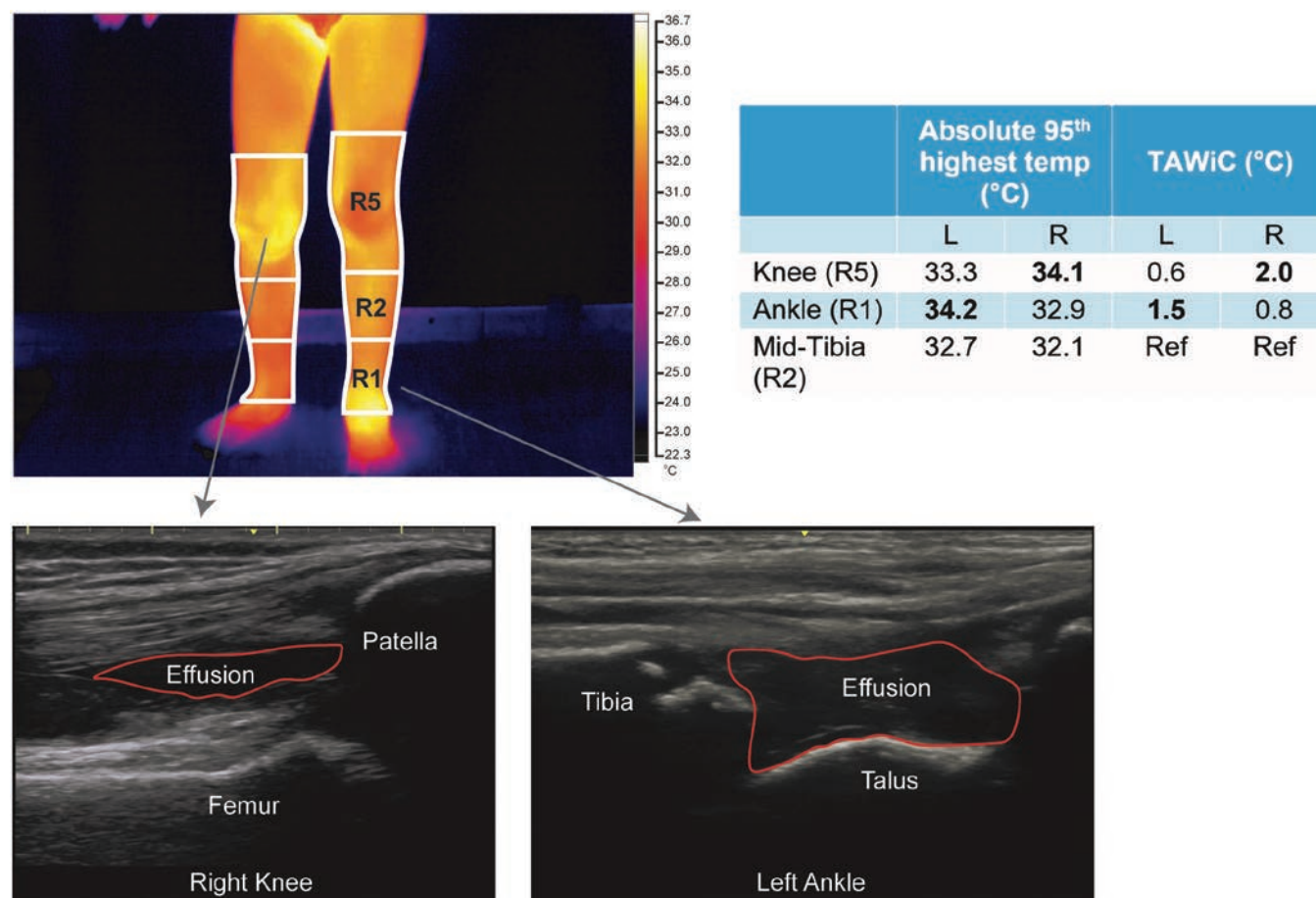


Figure 1. Representative thermal image, ultrasound images, and thermal analysis results from a child with arthritis (swelling on exam, effusion in ultrasound) in right knee and left ankle. Enclosed red lines on ultrasound indicate joint effusion. Significantly elevated temperatures vs reference region (mid-tibia) are shown in bold. TAWiC: temperature after within-limb calibration.

Table 2. TAWiC from knee and ankle ROI in children with JIA group and healthy children.

	Inflamed Joints (JIA)	Uninflamed Joints (JIA and Controls)	<i>P</i>
Analyses of knees and corresponding mid-tibia			
	n = 45	n = 142	
Knee			
95th absolute (anterior)	34.46 ± 1.20	33.76 ± 1.26	0.002
95th absolute (lateral)	34.03 ± 0.97	33.53 ± 1.05	0.006
95th absolute (medial)	33.96 ± 1.04	33.60 ± 1.20	0.09
95th absolute (posterior)	34.24 ± 1.03	34.28 ± 1.00	0.81
95th TAWiC (anterior)	1.02 ± 0.83	-0.08 ± 0.50	< 0.001
95th TAWiC (lateral)	1.15 ± 0.90	0.17 ± 0.50	< 0.001
95th TAWiC (medial)	1.27 ± 0.77	0.28 ± 0.60	< 0.001
95th TAWiC (posterior)	1.93 ± 0.85	1.64 ± 0.56	0.08
Mean absolute (anterior)	33.33 ± 1.27	32.71 ± 1.32	0.01
Mean absolute (lateral)	32.76 ± 1.01	32.56 ± 1.11	0.30
Mean absolute (medial)	32.85 ± 1.07	32.65 ± 1.17	0.35
Mean absolute (posterior)	32.95 ± 1.02	33.01 ± 1.09	0.75
Mean TAWiC (anterior)	0.67 ± 0.97	-0.39 ± 0.56	< 0.001
Mean TAWiC (lateral)	0.72 ± 0.74	-0.13 ± 0.52	< 0.001
Mean TAWiC (medial)	0.89 ± 0.76	0.02 ± 0.60	< 0.001
Mean TAWiC (posterior)	1.53 ± 0.73	1.22 ± 0.52	0.03
Mid-tibia			
95th absolute (anterior)	33.44 ± 1.23	33.85 ± 1.21	0.07
Mean absolute (anterior)	32.66 ± 1.26	33.10 ± 1.26	0.06
Analyses of ankles and corresponding mid-tibia			
	n = 20	n = 168	
Ankle			
95th absolute (anterior)	33.83 ± 1.55	33.55 ± 1.76	0.50
95th TAWiC (anterior)	0.60 ± 1.19	-0.28 ± 1.11	0.002
Mean absolute (anterior)	32.14 ± 1.55	32.23 ± 1.73	0.80
Mean TAWiC (anterior)	-0.42 ± 0.73	-0.84 ± 0.96	0.02
Mid-tibia			
95th absolute (anterior)	33.23 ± 1.43	33.83 ± 1.15	0.06
Mean absolute (anterior)	32.56 ± 1.40	33.08 ± 1.20	0.11

Values are expressed as mean ± SD. JIA: juvenile idiopathic arthritis; ROI: region of interest; TAWiC: temperature after within-limb calibration.

other views. Both TAWiC 95th percentile and mean temperatures of the inflamed knees from the anterior, lateral, and medial views differed from the uninflamed knees by about 1 °C. However, TAWiC 95th percentile temperatures of the inflamed ankles differed from the uninflamed ankles more than TAWiC mean temperature did (0.88 vs 0.42 °C). The temperatures of mid-tibia (a reference ROI for computing TAWiC) were slightly cooler in limbs corresponding to an inflamed joint, although this difference was not statistically significant.

ROC analyses using TAWiC_{knee} showed that the area under the curve (AUC) was similar among anterior, medial, and lateral views, but much lower in posterior views (Table 3). AUC was increased by 0.2 (30%) when TAWiC_{knee} was used comparing to absolute temperature (range 0.544–0.659). The thresholds of TAWiC_{knee}, which maximizes Youden index, were similar for anterior, medial, and lateral views (Table 3). The sensitivity of detecting arthritis in the knee varied from 0.644 to 0.778, and the specificity ranged between 0.793 and 0.923 excluding posterior view. The sensitivity of detecting inflammation in the ankle region from anterior view was 0.800 and the specificity was 0.601. Other views of ankle, using our ROI definition, repeatedly spilled outside of limb contours and therefore were

not evaluable. These results were similar to those from analyses completed with joint exam as the gold standard (Supplementary Tables 1–3, available with the online version of this article) or US alone as the gold standard (Supplementary Tables 5–6).

Factors correlated with TAWiC. Correlations of 95th percentile TAWiC with sex, age, height, weight, and BMI are shown in Table 4. Within the inflamed knee group, females had higher TAWiC_{knee} than males; younger children and shorter children had higher TAWiC_{knee} than their older and taller counterparts, respectively. Within the inflamed ankle group, there was no correlation of TAWiC_{ankle} with sex, age, height, weight, or BMI. However, within the HC group, males, younger children, and those with higher BMI had higher TAWiC_{ankle}, whereas younger and shorter children without inflamed ankles in JIA group had higher TAWiC_{ankle}.

Validation of using TAWiC to detect arthritis in the knee. Within the validation cohort, a knee joint was considered inflamed when both mean and 95th TAWiC_{knee} of each knee were greater than the corresponding thresholds from each view in the development cohort. Comparing to the results of physical exam as the gold standard, the sensitivity of accurate detection of arthritis from

Table 3. AUC when using TAWiC or absolute temperatures from knee and ankle, and thresholds derived by using TAWiC.

	AUC by TAWiC	AUC by Absolute Temperature	Youden-maximizing Threshold (°C)	Youden Index ^a	Sensitivity	Specificity
Knee						
95th TAWiC (anterior)	0.867	0.659	0.540	0.628	0.733	0.894
95th TAWiC (lateral)	0.840	0.653	0.710	0.560	0.689	0.871
95th TAWiC (medial)	0.846	0.587	0.805	0.583	0.733	0.850
95th TAWiC (posterior)	0.590	0.474	1.900	0.207	0.488	0.728
Mean TAWiC (anterior)	0.843	0.640	0.305	0.567	0.644	0.923
Mean TAWiC (lateral)	0.843	0.558	0.270	0.585	0.778	0.800
Mean TAWiC (medial)	0.815	0.544	0.463	0.504	0.711	0.793
Mean TAWiC (posterior)	0.618	0.470	1.485	0.226	0.465	0.761
Ankle						
95th TAWiC (anterior)	0.699	0.542	0.130	0.407	0.800	0.601
Mean TAWiC (anterior)	0.637	0.425	-0.720	0.276	0.800	0.476

^a Youden index = sensitivity + specificity - 1. AUC: area under the curve; TAWiC: temperature after within-limb calibration.

Table 4. Correlation of sex, age, height, weight, BMI, and body temperature with TAWiC_{knee} and TAWiC_{ankle}.

	HCs	JIA	
		Uninflamed Joint	Inflamed Joint
Anterior knee	n = 96	n = 46	n = 45
Sex	0.71 (0.49)	0.13 (0.41)	0.37 (0.01)
Age	0.07 (0.51)	-0.16 (0.30)	-0.23 (0.13)
Height	0.09 (0.38)	-0.20 (0.19)	-0.20 (0.18)
Weight	0.13 (0.22)	-0.18 (0.24)	0.07 (0.64)
BMI	0.07 (0.51)	-0.15 (0.33)	0.01 (0.93)
Lateral knee	n = 94	n = 46	n = 45
Sex	-0.11 (0.30)	0.15 (0.33)	0.45 (0.002)
Age	0.03 (0.76)	-0.28 (0.06)	-0.32 (0.03)
Height	0.12 (0.24)	-0.26 (0.08)	-0.30 (0.04)
Weight	0.16 (0.13)	-0.20 (0.19)	-0.23 (0.13)
BMI	0.11 (0.30)	-0.07 (0.63)	-0.18 (0.25)
Medial knee	n = 94	n = 46	n = 45
Sex	0.09 (0.37)	-0.06 (0.70)	0.03 (0.86)
Age	0.14 (0.17)	0.03 (0.86)	0.01 (0.93)
Height	0.13 (0.21)	-0.04 (0.79)	0.11 (0.49)
Weight	0.06 (0.57)	0.03 (0.85)	0.23 (0.13)
BMI	-0.06 (0.55)	0.01 (0.94)	0.21 (0.17)
Posterior knee	n = 96	n = 46	n = 43
Sex	0.24 (0.02)	0.06 (0.71)	-0.35 (0.02)
Age	0.08 (0.45)	-0.05 (0.75)	0.12 (0.45)
Height	0.07 (0.52)	-0.13 (0.41)	0.23 (0.14)
Weight	0.04 (0.74)	0.02 (0.89)	0.29 (0.06)
BMI	-0.04 (0.70)	0.08 (0.60)	0.15 (0.34)
Anterior ankle	n = 96	n = 83	n = 19
Sex	-0.21 (0.045)	0.006 (0.95)	0.24 (0.32)
Age	-0.22 (0.03)	-0.29 (0.008)	-0.23 (0.34)
Height	-0.09 (0.39)	-0.27 (0.02)	-0.20 (0.41)
Weight	0.15 (0.16)	-0.10 (0.37)	-0.05 (0.83)
BMI	0.31 (0.002)	0.13 (0.23)	0.11 (0.66)

Values are expressed as correlation coefficient (*P* value). HC: healthy control; JIA: juvenile idiopathic arthritis; TAWiC: temperature after within-limb calibration.

individual views ranged from 0.60 to 0.70 and the specificity was > 0.90 in all views (Table 5). When arthritis was defined as all mean and 95th TAWiC readings from every view being greater

than the corresponding thresholds of corresponding views, the sensitivity and specificity were similar to using individual views (Table 5). Although the study was not designed to validate

Table 5. Sensitivity and specificity of detection of arthritis in knee (n = 43)^a.

Views	Sensitivity (95% CI)	Specificity (95% CI)
Anterior	0.60 (0.26–0.88)	0.91 (0.82–0.96)
Lateral	0.70 (0.35–0.93)	0.95 (0.87–0.99)
Medial	0.70 (0.35–0.93)	0.93 (0.85–0.98)
Anterior + lateral + medial ^b	0.60 (0.26–0.88)	0.99 (0.93–1.00)

^a Definition of accurate detection by thermal imaging is made when the mean TAWiC of the knee is greater than the threshold for mean AND the 95th TAWiC of the knee is greater than the threshold for 95th temperature.

^b All mean and 95th TAWiC readings from every view must be greater than the corresponding thresholds of corresponding views. TAWiC: temperature after within-limb calibration.

the detection of ankle inflammation, the sensitivity of using TAWiC_{ankle} for detection was 0.80 and the specificity was 0.68.

DISCUSSION

To our knowledge, this is the first study to propose a novel algorithm to reliably detect active arthritis in children using infrared thermal imaging. Our approach to analyzing the thermal images from children with JIA and healthy children is reproducible and semiautomated, making it potentially useful in a wide range of situations to detect active arthritis. The addition of the within-leg internal control in this investigation improved the capacity to distinguish between inflamed and uninfamed joint area over an absolute temperature measure of the area of interest. This was a proof-of-concept study that focused on lower extremities due to the high prevalence of arthritis in knees and ankles. Further refinement of this approach may be applied to disease monitoring of chronic arthritis in both adults and children.

We not only identified significantly increased temperatures in both inflamed knee and ankle joints by absolute temperature, as in other studies,^{10,11} but also reduced the variation significantly by applying within-limb calibration. Therefore, our algorithm greatly improved the distinguishing ability of arthritis by thermal imaging. In addition, the definition of the knee joint and ankle were based on anatomy; this principle can be applied to other joints such as elbow, wrist, and digit joints. Another advantage of applying an internal control is to allow identification of joint inflammation in both legs of an affected individual.

Among all views, anterior, medial, and lateral views provided similar sensitivity to distinguish knees with inflammation from those without inflammation, and this is consistent with previous studies.^{10,11} For the ankle joint, due to greater anatomical complexity, articular or tendon sheath inflammation may cause temperature changes that are detectable only on certain views. In this analysis, only the anterior view showed a significant difference in TAWiC_{ankle} between inflamed and uninfamed ankles. Optimization of ROI for ankle joints from medial, lateral, and posterior views might allow us to determine the specificity of view-specific changes of temperatures that correspond to inflammation from specific anatomical structures. For example, isolated inflammation within lateral tendons may reveal elevated TAWiC_{ankle} only from a lateral view and not from other views.

Definition of ankle ROI and patterns of heat distribution from other views may be defined and evaluated through a machine learning approach in the future.

The significant effects of age, sex, and height on 95th TAWiC_{knee} and TAWiC_{ankle} in subgroups suggest that our method needs to be validated in various age groups, and that thresholds may be different depending on age and sex. It is also possible that the increase of TAWiC is dependent on the severity of joint swelling such that more subtle swelling is less detectable by thermal imaging. Using the current dataset from the development cohort, we identified thresholds of TAWiC for equally maximized sensitivity and specificity. For practical use, one may select a higher threshold for greater specificity when the pretest probability is low, such as screening of healthy children. In contrast, a lower threshold may be chosen for higher sensitivity when the pretest probability is high, such as in a child with history of JIA who has knee pain.

We validated the new algorithm and preliminary thresholds of TAWiC_{knee} in a separate cohort that demonstrated reasonable sensitivity and high specificity. With modification of the threshold of mean TAWiC_{knee}, sensitivity can be increased from 0.70 to 0.90 without sacrificing specificity. These results showed promise of potentially applying thermal imaging in screening and monitoring knee arthritis in children, especially during the era of increasing telehealth when joint exams are not performed in person. However, in-person visits and established imaging such as MRI and US are still needed when persistent symptoms are concerning despite normal thermal imaging results.

Our study had several limitations. Our sample size was small but comparable to previous studies, and exploratory statistics was performed without adjusting for multiple comparisons. Ankle ROI definition was not suitable for views other than anterior, which limits broader applicability. Finally, reproducibility of the temperature measurements over several days was not assessed due to difficulty in retaining subjects for repeat evaluations. However, we were able to prove that capacity for determining inflammation of knees and ankles by thermal imaging was increased when using internal calibration. Further, the determined thresholds can effectively screen for arthritis with a reasonable sensitivity and high specificity. These findings, if validated in a large population with optimization, will be highly applicable to patient care, especially during telehealth.

The use of a novel algorithm of infrared thermal imaging in children with active arthritis, or tenosynovitis, in knees or ankles revealed higher TAWiC than healthy unaffected joints. Our validation cohort study showed promise of the clinical utility of infrared thermal imaging for arthritis detection.

ACKNOWLEDGMENT

The authors would like to thank participants of the study; Ms. Ching Hung, Ms. Mary Eckert, and Mr. Christopher Budech for their help with infrared thermal imaging; and Dr. Jeffrey Ojemann for graciously allowing the study team to use the infrared thermal camera. We appreciate the guidance from Drs. Margret Rosenfeld, Dennis Shaw, Anne Stevens, Susan Halbach, and David Suskind; and the referrals of patients from Drs. Matthew Basiaga, Sri Grevich, Kristen Hayward, Shaun Jackson, Kabita Nanda, Sarah Ringold, Susan Sheno, Clayton Sontheimer, Anne Stevens, and Jennifer Turner.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. Manners PJ, Diepeveen DA. Prevalence of juvenile chronic arthritis in a population of 12-year-old children in urban Australia. *Pediatrics* 1996;98:84-90.
2. Huemer C, Malleson PN, Cabral DA, Huemer M, Falger J, Zidek T, et al. Patterns of joint involvement at onset differentiate oligoarticular juvenile psoriatic arthritis from pauciarticular juvenile rheumatoid arthritis. *J Rheumatol* 2002;29:1531-5.
3. Wallace CA, Giannini EH, Spalding SJ, Hashkes PJ, Neil KMO, Zeff AS, et al; Childhood Arthritis and Rheumatology Research Alliance. Trial of early aggressive therapy in polyarticular juvenile idiopathic arthritis. *Arthritis Rheum* 2012;64:2012-21.
4. Brenner M, Braun C, Oster M, Gulko PS. Thermal signature analysis as a novel method for evaluating inflammatory arthritis activity. *Ann Rheum Dis* 2006;65:306-11.
5. Sanchez BM, Lesch M, Brammer D, Bove SE, Thiel M, Kilgore KS. Use of a portable thermal imaging unit as a rapid, quantitative method of evaluating inflammation and experimental arthritis. *J Pharmacol Toxicol Methods* 2008;57:169-75.
6. Varju G, Pieper CF, Renner JB, Kraus VB. Assessment of hand osteoarthritis: correlation between thermographic and radiographic methods. *Rheumatology* 2004;43:915-9.
7. Denoble AE, Hall N, Pieper CF, Kraus VB. Patellar skin surface temperature by thermography reflects knee osteoarthritis severity. *Clin Med Insights Arthritis Musculoskelet Disord* 2010;3:69-75.
8. Spalding SJ, Kwok CK, Boudreau R, Enama J, Lunich J, Huber D, et al. Three-dimensional and thermal surface imaging produces reliable measures of joint shape and temperature: a potential tool for quantifying arthritis. *Arthritis Res Ther* 2008;10:R10.
9. Snehalatha U, Anburajan M, Sowmiya V, Venkatraman B, Menaka M. Automated hand thermal image segmentation and feature extraction in the evaluation of rheumatoid arthritis. *Proc Inst Mech Eng H* 2015;229:319-31.
10. Viitanen SM, Laaksonen AL. Thermography in juvenile rheumatoid arthritis. *Scand J Rheumatol* 1987;16:91-8.
11. Lasanen R, Piippo-Savolainen E, Remes-Pakarinen T, Kroger L, Heikkila A, Julkunen P, et al. Thermal imaging in screening of joint inflammation and rheumatoid arthritis in children. *Physiol Meas* 2015;36:273-82.
12. Ilowite NT, Walco GA, Pochaczewsky R. Assessment of pain in patients with juvenile rheumatoid arthritis: relation between pain intensity and degree of joint inflammation. *Ann Rheum Dis* 1992;51:343-6.
13. Zhao Y, Iyer RS, Reichley L, Oron AP, Gove NE, Kitsch AE, et al. A pilot study of infrared thermal imaging to detect active bone lesions in children with chronic nonbacterial osteomyelitis. *Arthritis Care Res* 2019;71:1430-5.
14. Laurell L, Court-Payen M, Nielsen S, Zak M, Fasth A. Ultrasonography and color Doppler in juvenile idiopathic arthritis: diagnosis and follow-up of ultrasound-guided steroid injection in the ankle region. A descriptive interventional study. *Pediatr Rheumatol Online J* 2011;9:4.
15. Zhao Y, Rascoff NE, Iyer RS, Thapa M, Reichley L, Oron AP, et al. Flares of disease in children with clinically inactive juvenile idiopathic arthritis were not correlated with ultrasound findings. *J Rheumatol* 2018;45:851-7.
16. Ting T V, Vega-Fernandez P, Oberle EJ, De Ranieri D, Bukulmez H, Lin C, et al; Childhood Arthritis and Rheumatology Research Alliance Juvenile Idiopathic Arthritis Ultrasound Workgroup. Novel ultrasound image acquisition protocol and scoring system for the pediatric knee. *Arthritis Care Res* 2019;71:977-85.
17. Wallace CA, Giannini EH, Huang B, Irtter L, Ruperto N; Childhood Arthritis Rheumatology Research Alliance; Pediatric Rheumatology Collaborative Study Group; Paediatric Rheumatology International Trials Organisation. American College of Rheumatology provisional criteria for defining clinical inactive disease in select categories of juvenile idiopathic arthritis. *Arthritis Care Res* 2011;63:929-36.