

# Development of an Ultrasound Imaging Atlas for Grading Osteoarthritis in the First Metatarsophalangeal Joint

Prue Molyneux,<sup>1</sup> Catherine Bowen,<sup>2</sup> Richard Ellis,<sup>1</sup> Keith Rome,<sup>1</sup> Kate Fitzgerald,<sup>3</sup> Phillip Clark,<sup>3</sup> Jackie L. Whittaker,<sup>4</sup> Charlotte Dando,<sup>2</sup> Richard Gee,<sup>5</sup> and Matthew Carroll<sup>1</sup>

**Objective.** Ultrasound (US) imaging may play a fundamental role in the earlier detection and assessment of first metatarsophalangeal joint (MTPJ) osteoarthritis (OA) because of its ability to depict tissue-specific morphologic changes before the point of irreversible structural damage. However, the role of US in supporting the diagnosis of OA in foot joints has not been clearly defined. The aims of this study were to develop a semiquantitative US atlas (the AUTUSI ultrasound imaging [AUTUSI] atlas) to grade the degree of osteoarthritic change in the first MTPJ and to evaluate the intraexaminer and interexaminer reproducibility of using the atlas.

**Methods.** US images were obtained from 57 participants (30 participants with radiographically confirmed first MTPJ OA). The AUTUSI atlas supports the examination of grading joint effusion, synovial hypertrophy, synovitis, osteophytes, joint space narrowing, and cartilage thickness. Six examiners used the atlas to independently grade 24 US images across 2 sessions. Intraexaminer and interexaminer reproducibility were determined using percentage agreement and Gwet's AC2.

**Results.** Observations using the AUTUSI atlas demonstrated almost perfect-to-perfect interexaminer agreement (percentage agreement ranged from 96% to 100%, and Gwet's AC2 values ranged from 0.81 to 1.00) and moderate-to-perfect intraexaminer agreement (percentage agreement ranged from 67% to 100%, and Gwet's AC2 values ranged from 0.54 to 1.00).

**Conclusion.** The AUTUSI atlas demonstrated excellent intraexaminer and interexaminer reproducibility for evaluating first MTPJ joint effusion, synovial hypertrophy, synovitis, joint space narrowing, osteophytes, and cartilage thickness. The AUTUSI atlas affords an opportunity to detect prognostic markers of OA earlier in the disease cascade and has the potential to advance understanding of the pathologic process of first MTPJ OA.

## INTRODUCTION

Osteoarthritis (OA) is a prevalent, chronic, progressive, and disabling joint disease that imposes a remarkable global health burden with notable implications for the individuals affected, health care systems, and wider socioeconomic costs.<sup>1,2</sup> Foot OA results in functional limitations and significant impairments in balance, strength, and locomotor ability and negatively impacts work ability.<sup>3</sup> However, the feet are often overlooked as a site of involvement relative to other joints commonly affected by OA.<sup>4</sup>

Within the foot, the most commonly reported affected foot site is the first metatarsophalangeal joint (MTPJ).<sup>5</sup>

The traditional view of OA as a degenerative disorder of articular cartilage resulting from normal bodily wear and tear is obsolete.<sup>6–8</sup> The contemporary concept of OA emphasizes the complex pathogenesis of the disease as a disorder of the joint as a whole organ, which involves not only hyaline cartilage but an additional and integrated role of bone and synovial tissue.<sup>6</sup> Attention has now turned to the prognostic value and role of markers of inflammation<sup>6,9,10</sup> with several studies reporting

The project is a component of a larger research project funded by the Health Research Council of New Zealand (21/025). This organization had no role in the study design; collection, analysis, or interpretation of the data; or the decision to submit the article for publication.

<sup>1</sup>Prue Molyneux, PhD, Richard Ellis, PhD, Keith Rome, PhD, Matthew Carroll, PhD: Auckland University of Technology, Auckland, New Zealand; <sup>2</sup>Catherine Bowen, PhD, Charlotte Dando, PhD: University of Southampton, Southampton, United Kingdom; <sup>3</sup>Kate Fitzgerald, BHS(MI), DMU, Phillip Clark, MBChB: Beyond Radiology, Auckland, New Zealand; <sup>4</sup>Jackie L. Whittaker, PhD: The University of British Columbia, Vancouver,

British Columbia, Canada; <sup>5</sup>Richard Gee, MBChB: Specialist Radiology Group, Auckland, New Zealand.

Additional supplementary information cited in this article can be found online in the Supporting Information section (<https://acrjournals.onlinelibrary.wiley.com/doi/10.1002/acr.25407>).

Author disclosures are available at <https://onlinelibrary.wiley.com/doi/10.1002/acr.25407>.

Address correspondence via email to Prue Molyneux, PhD, at [prue.susan.molyneux@aut.ac.nz](mailto:prue.susan.molyneux@aut.ac.nz).

Submitted for publication March 20, 2024; accepted in revised form July 15, 2024.

### SIGNIFICANCE & INNOVATIONS

- The AUT ultrasound imaging (AUTUSI) atlas recognizes the heterogeneous involvement of multiple joint tissues and provides a reliable means of evaluating both structural and inflammatory features specific to first metatarsophalangeal joint osteoarthritis (OA).
- The AUTUSI atlas offers the opportunity to identify prognostic inflammatory features earlier in the course of the disease before potentially irreversible damage or disability occurs.
- Application of the AUTUSI atlas may provide an opportunity to detect inflammatory features before the point of pain and before the point of irreversible structural damage. The ability to detect inflammatory features of OA before the point of irreversible structural change may provide a window of opportunity in which individuals may be more responsive to interventions, thus broadening the scope and capabilities of targeted interventions to alter disease progression and symptomatic consequence. Consequently, the AUTUSI atlas will expand people's eligibility for OA research and enable the study of targeted interventions at different stages of the disease process.

an association between active synovitis and structural OA progression.<sup>11–14</sup>

The development of the radiographic La Trobe Foot Atlas (LFA)<sup>15</sup> led to significant improvements in estimating foot OA disease prevalence,<sup>16</sup> as well as understanding different patterns of foot joint involvement.<sup>17</sup> However, the current method of diagnosing OA by conventional radiography remains reactive and captures OA later in the disease process, once the condition is significantly advanced.<sup>15,18</sup> This necessitates a new model of diagnosis and care for OA that is proactive and preventive. More advanced modalities, including magnetic resonance imaging (MRI) and ultrasound (US) imaging, have emerged as more accurate evaluators of both bone and soft tissue abnormalities in foot OA.<sup>19,20</sup> MRI has played a principal role in changing the understanding of OA pathologies when evaluating OA as a whole organ disease.<sup>10,21</sup> Munteanu et al<sup>19</sup> recently developed an MRI atlas for the assessment of first MTPJ OA that demonstrated excellent intraexaminer and interexaminer reproducibility.<sup>19,22</sup> However, shortcomings for MRI include high costs, prolonged duration of image acquisition, and limited availability in community care, and it is contraindicated in certain conditions such as metal implants.<sup>23,24</sup> Notably, Oo et al<sup>25</sup> demonstrated that the correlations between quantitative knee OA US features and corresponding MRI findings were very strong (intraclass correlation coefficient [ICC] range = 0.85–0.98).<sup>25</sup>

US imaging potentially affords inherent advantages for the diagnosis of first MTPJ OA. US imaging can be performed

chairside, has the ability to detect inflammatory joint pathology that is otherwise not detected by clinical examination,<sup>26,27</sup> and can reliably quantify both bone and soft tissue abnormalities.<sup>28,29</sup>

US imaging may play a fundamental role in the earlier detection and assessment of foot OA,<sup>30,31</sup> thus enabling more targeted and timely interventions that may provide the capacity to alter disease progression. However, no foot-specific US grading system exists to classify the degree of pathologic change in joint tissues. Therefore, the role of US in the diagnosis of OA in foot joints has not been clearly defined. The development of a US imaging atlas could enable earlier detection, a standardization of diagnosis, and a more sensitive method for classifying and grading the disease process. Therefore, the aims of the current study were to develop a semiquantitative US imaging atlas (image-based grading system) to grade the degree of osteoarthritic change in the first MTPJ and to evaluate the intraexaminer and interexaminer reproducibility of using the atlas.

### METHODS

#### Development of the AUTUSI atlas. *Background work.*

The AUTUSI atlas was developed using an evidence-based approach, using findings from a systematic review,<sup>32</sup> a scoping review,<sup>33</sup> a Delphi consensus study,<sup>34</sup> and a US imaging reliability study<sup>29</sup> as the basis for development. The systematic review identified inconsistencies in the assessment of US imaging features, the definition of features, and the grading systems used to determine the degree of OA change in peripheral joints.<sup>32</sup> Inconsistencies were identified by the scoping review against international guidelines and limited implementation of consensus-based recommendations for US imaging procedure guidance when evaluating the first MTPJ.<sup>33</sup> To address these research gaps, an international multispecialty Delphi study was conducted to gain consensus concerning which US imaging features should be assessed and graded and what US imaging procedure should be performed when examining the first MTPJ for features of OA.<sup>34</sup> The Delphi study identified the essential components that the US imaging acquisition procedure should encompass when examining the first MTPJ for features of OA.<sup>34</sup> Our reliability study details the US imaging acquisition procedure and semiquantitative grading system for evaluating first MTPJ OA features in participants with radiologically confirmed OA.<sup>29</sup> The US imaging acquisition procedure and grading system were reliable in evaluating joint effusion, synovial hypertrophy, synovitis, joint space narrowing, osteophytes, and cartilage thickness in participants with radiologically confirmed OA. Two examiners demonstrated good to excellent interexaminer reliability across two sessions (ICC 0.6–1.0 and 0.55–1.0, respectively) and intraexaminer reliability (ICC 0.58–0.92) for all features graded on a semiquantitative scale.<sup>29</sup> The semiquantitative grading of these US features was incorporated into the methodologic development of the AUTUSI atlas. The Southern Health and Disability Ethics

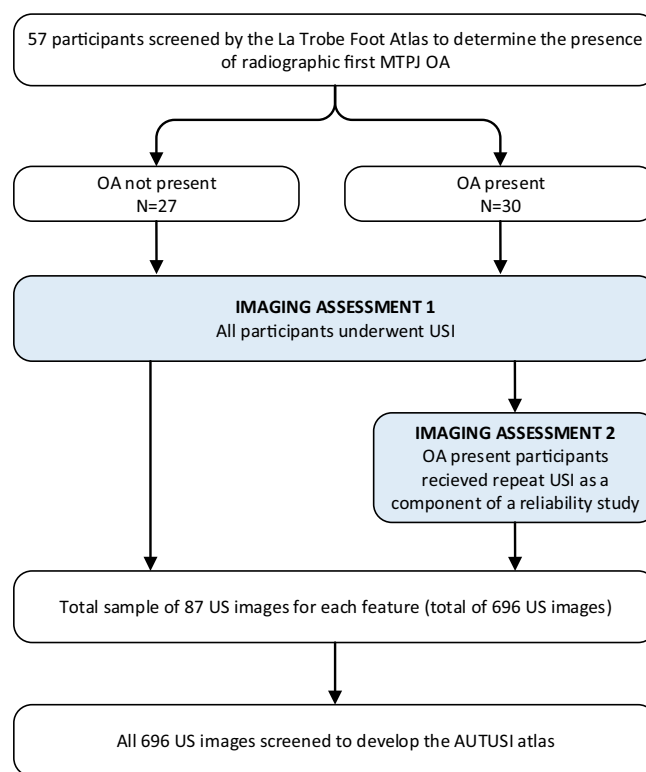
Committee, HDEC Ethics Reference: 2022 FULL 12721, granted ethical approval for the study.

**Participants.** US images were obtained from 57 participants who participated in an earlier reliability study.<sup>29</sup> Participants over 20 years of age with a previous imaging diagnosis (confirmed by an imaging modality) or clinically diagnosed (diagnosed based on clinical assessments) first MTPJ OA were recruited from the general population in Auckland, New Zealand. Participants were recruited through professional interactive networks, social media (Twitter/X and Facebook) and local newspaper advertisements. Exclusion criteria were the possibility of pregnancy, the presence of any other inflammatory musculoskeletal condition, history of a first MTPJ surgery, or foot and/or ankle surgery in the last three months. Study participants were categorized into two groups (radiographically confirmed OA and non-OA group). The LFA was used to determine the presence of radiographic first MTPJ OA.<sup>15</sup> A total of 57 participants were screened in the first imaging session, of whom 30 had radiographically confirmed first MTPJ OA. All 30 participants with radiographic first MTPJ OA were invited to return for a repeat US imaging examination as part of the reliability study.<sup>29</sup> Across both US imaging sessions, 87 sets of US images (one static image per feature) were acquired. Each participant was randomly assigned an alphanumeric code upon entry into the study. Written informed consent was obtained from all participants before the study. Figure 1 provides an overview of participant engagement and how US images were acquired to develop the atlas.

**Participant and imaging data.** All imaging was performed at a private medical imaging facility and was conducted sequentially within a 60-minute session in three separate rooms. First, demographic data were obtained for each participant (age, gender, height, weight, body mass index, ethnicity, and first MTPJ affected). Second, the presence of radiographic first MTPJ OA was determined through radiographic screening and assessment. Third, participants underwent a US examination using the US imaging acquisition procedure and grading system.<sup>29</sup>

**Radiographic assessment.** Weightbearing dorsal/plantar and lateral radiographs were obtained by an experienced radiographer. A radiologist (PC, who has more than 20 years of experience) used the LFA to determine the presence of radiographic first MTPJ OA.<sup>15</sup> The LFA considers OA to be present when a score of two or greater for osteophytes or joint space narrowing is documented from either the dorsal/plantar or lateral view.<sup>15</sup> One radiologist (PC) assessed and reported on all radiographs.

**US imaging assessment.** Directly after the radiographic examination, all 57 participants received a US examination. US features included in the examination were joint effusion, synovial hypertrophy, synovitis, osteophytes, joint space narrowing, and cartilage thickness. All US features were assessed in the dorsal view with the probe positioned longitudinally; a transverse



**Figure 1.** Atlas development participant journey. AUTUSI, Auckland University of Technology ultrasound imaging; MTPJ, metatarsophalangeal joint; OA, osteoarthritis; US, ultrasound; USI, ultrasound imaging.

orientation was also applied to examine joint effusion and synovial hypertrophy to ensure the extent of the effusion and synovial thickening was fully investigated. A semiquantitative grading system was applied to all features (0 = absent, 1 = mild, 2 = moderate, 3 = severe). To mitigate problems with adequate discrimination between intermediary grades of cartilage thickness, a zero to two semiquantitative grading system was applied (Supplementary File 1). A sonographer (KF, who has more than 15 years of experience in musculoskeletal US imaging) applied the US imaging acquisition procedure and grading system to evaluate the first MTPJ for features of OA. A radiologist (PC) applied the grading system to the static images acquired by the sonographer. Both examiners were blinded to the radiographic results, clinical data, and each other's grades. All US images were stored for the purpose of developing the AUTUSI atlas.

**Equipment.** A Philips EPIQ Elite HW B.2 US machine, equipped with a multifrequency linear transducer (eL18–4 MHz), was used to acquire images of the first MTPJ. The US device did not change during the session or between sessions. Grayscale was used to examine all features (128-Hz gray map 3, grayscale gain 30%–40%, dynamic range 74%–68%, med, 2D optical resolution, SonoCT, XRES 4) and a power Doppler was applied for the examination of synovitis (PRF 700 Hz, nondirectional flow color map CPA 3, color gain 50%–65%, wall filter 63 Hz, frequency 6.2 MHz). The grayscale and power Doppler settings were

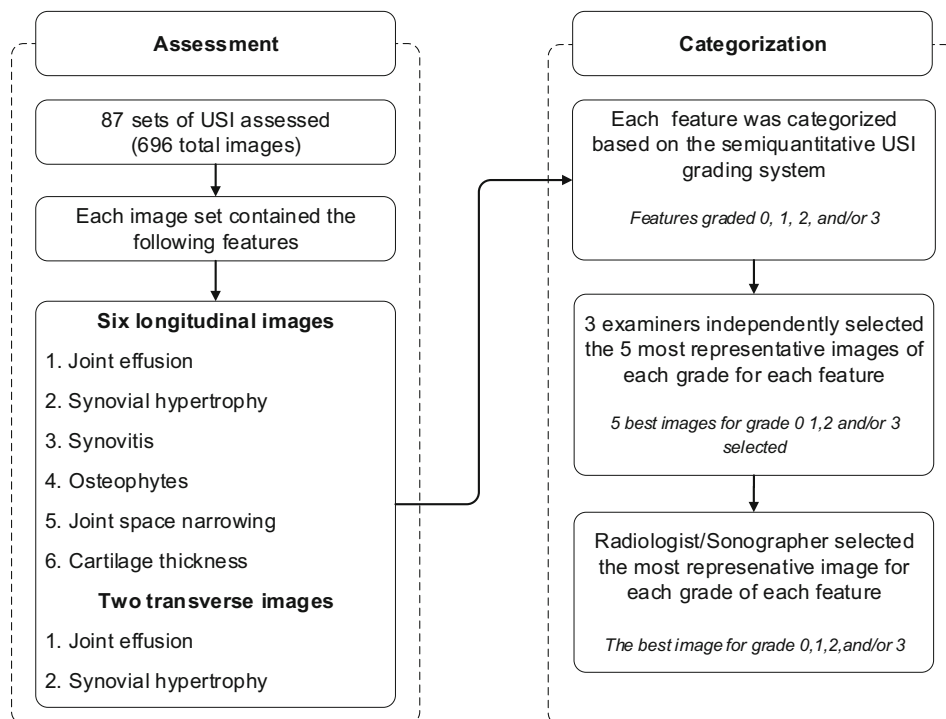
adjusted to optimize image resolution and sensitivity to detect flow. At the beginning of each scanning session, focus was positioned at the level of the region of interest. Color gain was adjusted below the degree that caused the appearance of noise artifacts.

**AUTUSI atlas image selection and consensus.** The first MTPJ OA AUTUSI atlas was developed by a podiatrist (PM), sonographer (KF), and radiologist (PC), who were all involved in the reliability study.<sup>29</sup> Figure 2 displays the image assessment and categorization process used to develop the AUTUSI atlas. The three atlas developers (PM, KC, and PC) assessed all 87 sets of US images. Each set of US images included six longitudinal images illustrating joint effusion, synovial hypertrophy, synovitis, osteophytes, joint space narrowing, and cartilage thickness and two transverse images illustrating joint effusion and synovial hypertrophy. PM categorized all 696 static US images by feature and severity according to grading by the radiologist and sonographer from the authors' reliability study.<sup>29</sup> All 696 static images representing each feature with varying degrees of severity were reviewed to select the most representative examples of each grade of each feature. PM, KF, and PC independently selected five images that best represented each feature definition and rating description. From here, the sonographer and radiologist independently selected the best image for each grade of each feature. If there was a disagreement, the image in question was replaced with another image that was considered a better representation. This process was repeated until consensus was achieved. This

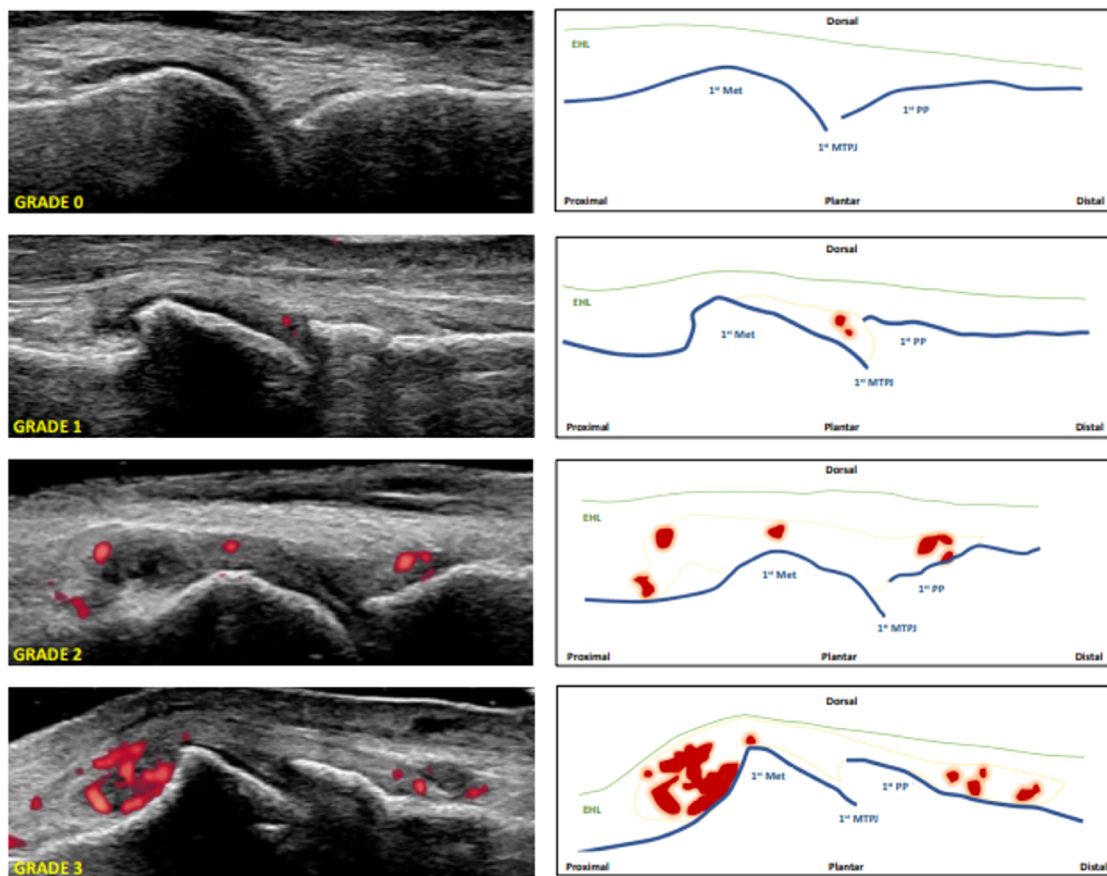
process led to the final presentation of the first MTPJ OA AUTUSI atlas.

**Preparation of US images.** An initial AUTUSI atlas was developed consisting of six features with an associated semi-quantitative grading system. The AUTUSI atlas consisted of 31 static US images. To assist in the differentiation between grades and features, PM designed line drawings to align with each US image. The AUTUSI atlas was distributed to KF and PC for approval. The AUTUSI atlas was independently evaluated and confirmed appropriate by co-authors (MC, CB, RE, and KR). Agreement was reached regarding image selection and format design. Figure 3 shows the AUTUSI atlas images for the assessment of synovitis. Power Doppler signals must be detected within synovial hypertrophy to be considered a sign of synovitis. The scoring of power Doppler signals was performed according to a semiquantitative scale (0 = no flow in the synovium; 1 = mild, single vessel signals (one or more); 2 = moderate, confluent vessel signals in less than half of the area of the synovium; 3 = severe, vessel signals in more than half of the area of the synovium). The full AUTUSI atlas is presented in Supplementary File 2.

**Reproducibility of the AUTUSI atlas. Image selection for grading.** To assess the reproducibility of the AUTUSI atlas, a random sample of 24 images were selected from the total sample of 696 available images for grading. US images were grouped by participant code according to feature and grade using Microsoft Excel (Version 2307, Microsoft Corporation). Participant codes



**Figure 2.** Image assessment and categorisation process used to develop the AUTUSI atlas. AUTUSI, Auckland University of Technology ultrasound imaging; US, ultrasound; USI, ultrasound imaging.



**Figure 3.** Atlas images for synovitis of the first MTPJ. 1st Met, first metatarsal; 1st PP, first proximal phalanx; 1st MTPJ, first metatarsophalangeal joint; EHL, extensor hallucis longus.

representing images included in the AUTUSI atlas were deleted from the spreadsheet to ensure atlas examiners would not grade images that had been previously selected as atlas images. US images were randomly selected by row number using a random number generator. Twenty-four US images (3 images per feature) were randomly selected from the remaining 665 images. Selected images were compiled into a document and sent with the AUTUSI atlas to examiners.

**Atlas examiners.** Six examiners with experience in US imaging from different countries (New Zealand [n = 4], United Kingdom [n = 1], and Canada [n = 1]) and different health professions (radiologists [n = 2], podiatrists [n = 2], sonographers [n = 1], and physical therapists [n = 1]) independently rated the randomly selected US images using the AUTUSI atlas. All examiners were provided with written instructions detailing the AUTUSI atlas, the US imaging grading sheet, and the tasks required of them for grading sessions one and two. An additional video outlining the process was also provided. This step included three examiners (PM, KF, and PC) who developed the atlas and three examiners who were not involved in the US imaging acquisition or the atlas development—a radiologist, a physical therapist, and a podiatrist. To determine the interexaminer reproducibility of the AUTUSI atlas, all six examiners (PM, KF, PC, RG, JW, and CD) independently rated the

random sample of 24 US images (3 images per feature of varying severity) using the atlas. To determine intraexaminer reproducibility, this process was repeated three to four weeks later (without reference to previous ratings).

**Statistical analysis.** Gwet’s AC2 values were calculated using the irrCAC package in R, version 1.0, to determine interexaminer and intraexaminer reliability. Quadratic weighting was used for all calculations because of the ordinal nature of scoring across the US features. Gwet’s AC2 values were interpreted using the following cutoffs: <0 poor; 0.01 to 0.20 slight; 0.21 to 0.40 fair; 0.41 to 0.60 moderate; 0.61 to 0.80 substantial; and 0.81 to 1.00 almost perfect or 1.00 perfect.<sup>35</sup> Gwet’s AC1 values were calculated because this approach has been shown to provide a more robust measure of reliability than the traditionally used kappa statistic.<sup>36</sup>

**RESULTS**

Table 1 details the demographic characteristics of the 57 participants (25 women, 5 men) included in the AUTUSI atlas development study. Demographic characteristics are categorized with and without radiographic first MTPJ OA.

**Table 1.** Participant demographic data\*

Participant characteristics	Non-OA	Radiographically confirmed OA
Number	27	30
Age, mean $\pm$ SD, y	40.7 $\pm$ 9.7	54.9 $\pm$ 10.4
Sex, n (%)		
Male	6 (11)	5 (9)
Female	21 (37)	25 (44)
Ethnicity, n (%)		
NZ European	18 (32)	22 (39)
Māori	1 (2)	4 (7)
Asian	3 (5)	1 (2)
White British	2 (4)	2 (4)
Australian	1 (2)	0 (0)
Canadian	1 (2)	0 (0)
Russian	0 (2)	1 (2)
South African	1 (2)	0 (0)
Body mass index, mean $\pm$ SD, kg/m <sup>2</sup>	28.4 $\pm$ 4.5	30.3 $\pm$ 6.7
First MTPJ affected, n (%)		
Left	12 (21)	13 (23)
Right	15 (26)	17 (30)

\* OA, osteoarthritis; MTPJ, first metatarsophalangeal joint.

**Interexaminer reproducibility.** Table 2 displays the reproducibility statistics for interexaminer analyses. Percentage exact agreement ranged from 96% to 100%. Gwet's AC2 values ranged from 0.81 to 1.00, indicating almost perfect-to-perfect levels of agreement. Interexaminer reproducibility was similar between session one and session two grading. Although acceptable, the grading of synovitis was not as consistent as other US features.

**Intraexaminer reproducibility.** Table 3 displays the reproducibility statistics for intraexaminer analyses. Percentage exact agreement ranged from 67% to 100%. Gwet's AC2 values

ranged from 0.54 to 1.00, indicating moderate-to-perfect levels of agreement. Intraexaminer reproducibility was similar for the examiners 1, 2, 3, and 5. Although acceptable, the grading of synovitis was not as consistent as other US features for two examiners.

## DISCUSSION

This study developed and assessed the reproducibility of the AUTUSI atlas to evaluate the presence and severity of joint effusion, synovial hypertrophy, synovitis, osteophytes, joint space narrowing, and cartilage thickness in the first MTPJ. The AUTUSI atlas demonstrated excellent intraexaminer and interexaminer reproducibility for evaluating first MTPJ joint effusion, synovial hypertrophy, joint space narrowing, osteophytes, and cartilage thickness. Although acceptable, the grading of synovitis was not as consistent as other US features.

The uncertainty in differentiating between images representing different grades of synovitis may be partly attributable to images depicting a similar volume of power Doppler signal. As outlined in the synovitis description of the AUTUSI atlas, power Doppler signals must be detected within the synovium to be considered a sign of synovitis (0 = no flow in the synovium; 1 = mild, single vessel signals [one or more]; 2 = moderate, confluent vessel signals in less than half of the area of the synovium; 3 = severe, vessel signals in more than half of the area of the synovium). The randomly selected synovitis images chosen for the US imaging grading sheet represent different grades of severity. However, when assessing the power Doppler signal across the entire image and not just the Doppler signal contained to the synovium, the volume may have appeared similar. Therefore, there may have been cases in which the Doppler signal was graded outside the synovium, consequently increasing the severity grade. This finding

**Table 2.** Interexaminer reliability of grading sessions 1 and 2\*

Interexaminer reliability	Percentage agreement	Gwet's AC2 value (95% CI)	Interpretation
Session 1			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	98.02%	0.94 (0.58–1.00)	Almost perfect
Synovial hypertrophy (longitudinal)	95.80%	0.86 (0.41–1.00)	Almost perfect
Synovial hypertrophy (transverse)	98.76%	0.96 (0.80–1.00)	Almost perfect
Synovitis	91.11%	0.82 (0.20–1.00)	Almost perfect
Osteophytes	96.30%	0.89 (0.84–0.93)	Almost perfect
Joint space narrowing	97.22%	0.92 (0.52–1.00)	Almost perfect
Cartilage	97.22%	0.92 (0.52–1.00)	Almost perfect
Session 2			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	98.02%	0.94 (0.58–1.00)	Almost perfect
Synovial hypertrophy (longitudinal)	97.53%	0.92 (0.73–1.00)	Almost perfect
Synovial hypertrophy (transverse)	97.53%	0.93 (0.69–1.00)	Almost perfect
Synovitis	92.22%	0.81 (0.11–1.00)	Almost perfect
Osteophytes	97.22%	0.92 (0.52–1.00)	Almost perfect
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	97.22%	0.92 (0.52–1.00)	Almost perfect

\* 95% CI, 95% confidence interval.

**Table 3.** Intraexaminer (between-session) reliability\*

Intrarater (between-session) reliability	Percentage agreement	Gwet's AC2 value (95% CI)	Interpretation
<b>Examiner 1</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovitis	91.67%	0.79 (–0.29 to 1.00)	Substantial
Osteophytes	100.00%	1.00 (1.00–1.00)	Perfect
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect
<b>Examiner 2</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovitis	100.00%	1.00 (1.00–1.00)	Perfect
Osteophytes	100.00%	1.00 (1.00–1.00)	Perfect
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect
<b>Examiner 3</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovitis	100.00%	1.00 (1.00–1.00)	Perfect
Osteophytes	100.00%	1.00 (1.00–1.00)	Perfect
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect
<b>Examiner 4</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	96.29%	0.88 (0.17–1.00)	Almost perfect
Synovial hypertrophy (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovitis	66.67%	0.54 (–2.06 to 1.00)	Moderate
Osteophytes	88.89%	0.64 (0.36–0.91)	Substantial
Joint space narrowing	91.67%	0.79 (–0.29 to 1.00)	Substantial
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect
<b>Examiner 5</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	92.59%	0.76 (0.05–1.00)	Substantial
Synovial hypertrophy (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovitis	100.00%	1.00 (1.00–1.00)	Perfect
Osteophytes	100.00%	1.00 (1.00–1.00)	Perfect
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect
<b>Examiner 6</b>			
Joint effusion (longitudinal)	100.00%	1.00 (1.00–1.00)	Perfect
Joint effusion (transverse)	100.00%	1.00 (1.00–1.00)	Perfect
Synovial hypertrophy (longitudinal)	92.59%	0.76 (0.05–1.00)	Substantial
Synovial hypertrophy (transverse)	96.30%	0.88 (0.17–1.00)	Almost perfect
Synovitis	100.00%	1.00 (1.00–1.00)	Perfect
Osteophytes	91.67%	0.79 (–0.29 to 1.00)	Substantial
Joint space narrowing	100.00%	1.00 (1.00–1.00)	Perfect
Cartilage	100.00%	1.00 (1.00–1.00)	Perfect

\* 95% CI, 95% confidence interval.

may also reflect the use of previous US imaging grading systems originally constructed from rheumatoid arthritis (RA) populations, where the degree of inflammation is distinctly much higher.<sup>37,38</sup> Additionally, the marked variations across interpretations of how the synovial pathology has been defined and categorized in

previous investigations may also be reflected in the scoring variation.<sup>32</sup> Variation may also be explained by the number of choices with a semiquantitative grading scale; synovitis grading variation may be attributed to one examiner overscoring and one underscoring when the image was doubtful. Therefore, additional

training of all examiners before use of the AUTUSI atlas may improve performance and ensure understanding of grading descriptions.

To elucidate the role of inflammation in foot OA, grayscale inflammatory abnormalities for synovial hypertrophy and joint effusion were graded separately. Synovitis was also examined as a separate entity by power Doppler signal (flow signal detected within synovial hypertrophy was considered a sign of synovitis).<sup>39,40</sup> Differentiating between inflammatory features addresses past inconsistencies and interpretation of the different entities of synovial pathology. Previous discrepancies in how synovitis, synovial hypertrophy, and joint effusion have been defined and categorized make it unclear if synovial pathology is best represented as separate entities (joint effusion and synovial hypertrophy) or combined as proxy measures for synovitis.<sup>32</sup> Because of the marked variation in prevalence between grayscale and Doppler-detected inflammatory features demonstrated in hand OA,<sup>40</sup> including grayscale features indicative of inflammation solely may result in OA being underestimated. The distinction between inflammatory pathology is further enhanced by our semi-quantitative grading system specific to OA. To date, US imaging grading systems applied to OA have been largely extrapolated from those originally designed and validated to quantify inflammatory change in RA.<sup>32</sup> The distinct difference of inflammation experienced in OA compared with RA<sup>12,37</sup> reinforced the need for an OA-specific grading that truly depicts disease progression, as in this work.

The AUTUSI atlas will allow for quantification of the presence and severity of pathologic change occurring in joint tissues in first MTPJ OA. The AUTUSI atlas has the potential to advance understanding of the pathologic process of first MTPJ OA, provide capacity for earlier detection, support standardization of diagnosis, and subsequently broaden the scope and capabilities of targeted interventions. From this work, it is evident that there is a role for using US imaging in detecting different stages of first MTPJ OA. The AUTUSI atlas may be the catalyst in developing a classification criteria specific to the first MTPJ OA, which could be used both clinically and in future research. The AUTUSI atlas will enhance the capability to detect and characterize first MTPJ OA across various stages of disease progression rather than being restricted solely to the point of irreversible end-stage disease.

The AUTUSI atlas is the first step in a longitudinal analysis to determine the predictive value of US features in participants without radiographically confirmed OA. Incorporating the AUTUSI atlas as an outcome measure in population-based cohort studies could provide valuable insight into early detection. Furthermore, it provides the capacity to determine which US imaging features may progress to radiographic change and increased symptomatic consequences. Finally, further work is needed to determine which AUTUSI atlas features would be most accurate in the diagnosis of first MTPJ OA compared with radiographic grading, the

relationship between the atlas and symptom severity, and how atlas features respond to clinical interventions.

A particular strength of the present study is the inclusion of six examiners from different countries and health professions who are all experienced users of US imaging. Additionally, three reviewers were masked to the image acquisition and atlas development process. The justification for including six examiners from different professional backgrounds ensures the atlas is reproducible across a range of health professions that are likely to implement the atlas, such as podiatrists, physical therapists, radiologists, sonographers, and researchers. In comparison, an MRI atlas and the LFA included the same two examiners for atlas development and reproducibility.<sup>15,19</sup> A further strength of our study is the accompaniment of supporting images in our atlas that clearly depict the US imaging (patient and probe positioning) acquisition procedure. US imaging has previously been demonstrated as being operator-dependent.<sup>41,42</sup> However, substantial developments in US software and transducers have advanced the resolution to better identify superficial structures, such as the first MTPJ, allowing more detailed assessment of associated pathology.<sup>28,43,44</sup> To help mitigate operator dependence, the authors wanted to ensure that a standardized procedure could be performed for all atlas users. Supporting images would ensure correct interpretation and allow subsequent studies to replicate our procedure.

This study must be viewed in the context of its limitations. First, despite examining 24 US images with the atlas, the reviewers only evaluated 3 images per feature. Therefore, future research is needed to determine the reproducibility of the atlas when grading a larger sample of images. Second, there may have been examiner bias because the same radiologist reported on radiographic screening and graded the acquired US images. To minimize the risk of bias, we randomly assigned an alphanumeric code to all participants upon entry into the study. Finally, the prevalence and burden of OA are not uniform across demographic groups. Minority populations, especially African American, Hispanic, and Māori and Pacific people, experience poorer health outcomes (such as pain and disability).<sup>45,46</sup> Despite efforts to proactively recruit an ethnically diverse population representative of the broader New Zealand population, no Pacific peoples were included. Pacific peoples suffer from significant and long-standing health inequalities and poorer health outcomes compared with other New Zealanders. Therefore, the inclusion of ethnicity related to OA data is needed to better understand what factors contribute to these inequalities and to provide the capacity to measure progress.

The AUTUSI atlas demonstrated excellent intraexaminer and interexaminer reproducibility for evaluating first MTPJ joint effusion, synovial hypertrophy, synovitis, joint space narrowing, osteophytes, and cartilage thickness. The AUTUSI atlas affords an opportunity to detect prognostic markers of OA earlier in the disease cascade, which would provide a window of opportunity

and broaden the scope and capabilities of targeted interventions to alter disease progression. The AUTUSI atlas has the potential to advance understanding of the pathologic process of first MTPJ OA.

## ACKNOWLEDGMENTS

The authors thank all participants who volunteered for the study and the staff of Beyond Radiology for their assistance. Open access publishing facilitated by Auckland University of Technology, as part of the Wiley - Auckland University of Technology agreement via the Council of Australian University Librarians.

## 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 Molyneux 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.** Molyneux, Bowen, Ellis, Rome, Carroll.

**Acquisition of data.** Molyneux, Fitzgerald, Clark.

**Analysis and interpretation of data.** Molyneux, Fitzgerald, Clark, Whitaker, Dando, Gee.

## REFERENCES

- Hunter DJ, Schofield D, Callander E. The individual and socio-economic impact of osteoarthritis. *Nat Rev Rheumatol* 2014;10(7):437–441.
- Samvelyan HJ, Hughes D, Stevens C, et al. Models of osteoarthritis: relevance and new insights. *Calcif Tissue Int* 2021;109(3):243–256.
- Paterson KL, Gates L. Clinical assessment and management of foot and ankle osteoarthritis: a review of current evidence and focus on pharmacological treatment. *Drugs Aging* 2019;36(3):203–211.
- Mathiessen A, Fjellstad C, Slatkowsky-Christensen B, et al. The reliability and distribution of foot osteoarthritis assessed by ultrasound and clinical examination in a hand osteoarthritis cohort. *Osteoarthritis Cartilage* 2021;29:S353–S354.
- Roddy E, Thomas MJ, Marshall M, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: cross-sectional findings from the clinical assessment study of the foot. *Ann Rheum Dis* 2015;74(1):156–163.
- Sellam J, Berenbaum F. The role of synovitis in pathophysiology and clinical symptoms of osteoarthritis. *Nat Rev Rheumatol* 2010;6(11):625–635.
- Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. *Lancet* 2011;377(9783):2115–2126.
- Lane NE, Brandt K, Hawker G, et al. OARSI-FDA initiative: defining the disease state of osteoarthritis. *Osteoarthritis Cartilage* 2011;19(5):478–482.
- Robinson WH, Lepus CM, Wang Q, et al. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. *Nat Rev Rheumatol* 2016;12(10):580–592.
- Loeser RF, Goldring SR, Scanzello CR, et al. Osteoarthritis: a disease of the joint as an organ. *Arthritis Rheum* 2012;64(6):1697–1707.
- Mancarella L, Addimanda O, Pelotti P, et al. Ultrasound detected inflammation is associated with the development of new bone erosions in hand osteoarthritis: a longitudinal study over 3.9 years. *Osteoarthritis Cartilage* 2015;23(11):1925–1932.
- Kortekaas MC, Kwok WY, Reijnen M, et al. Inflammatory ultrasound features show independent associations with progression of structural damage after over 2 years of follow-up in patients with hand osteoarthritis. *Ann Rheum Dis* 2015;74(9):1720–1724.
- Hammer HB, Iagnocco A, Mathiessen A, et al. Global ultrasound assessment of structural lesions in osteoarthritis: a reliability study by the OMERACT ultrasonography group on scoring cartilage and osteophytes in finger joints. *Ann Rheum Dis* 2016;75(2):402–407.
- Berenbaum F. Osteoarthritis as an inflammatory disease (osteoarthritis is not osteoarthrosis!). *Osteoarthritis Cartilage* 2013;21(1):16–21.
- Menz HB, Munteanu SE, Landorf KB, et al. Radiographic classification of osteoarthritis in commonly affected joints of the foot. *Osteoarthritis Cartilage* 2007;15(11):1333–1338.
- Menz HB, Roddy E, Marshall M, et al. Demographic and clinical factors associated with radiographic severity of first metatarsophalangeal joint osteoarthritis: cross-sectional findings from the Clinical Assessment Study of the Foot. *Osteoarthritis Cartilage* 2015;23(1):77–82.
- Rathod T, Marshall M, Thomas MJ, et al. Investigations of potential phenotypes of foot osteoarthritis: cross-sectional analysis from the clinical assessment study of the foot. *Arthritis Care Res (Hoboken)* 2016;68(2):217–227.
- Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;16(4):494–502.
- Munteanu SE, Auhl M, Tan JM, et al. Development and reproducibility of a first metatarsophalangeal joint osteoarthritis magnetic resonance imaging scoring system. *Arthritis Care Res (Hoboken)* 2020;72(9):1205–1212.
- Zabotti A, Filippou G, Canzoni M, et al. OMERACT Ultrasound Task Force members. OMERACT agreement and reliability study of ultrasonographic elementary lesions in osteoarthritis of the foot. *RMD Open* 2019;5(1):e000795.
- Sharma L, Chmiel JS, Almagor O, et al. Significance of preradiographic magnetic resonance imaging lesions in persons at increased risk of knee osteoarthritis. *Arthritis Rheumatol* 2014;66(7):1811–1819.
- Munteanu SE, Auhl M, Tan JM, et al. Characterisation of first metatarsophalangeal joint osteoarthritis using magnetic resonance imaging. *Clin Rheumatol* 2021;40(12):5067–5076.
- Hayashi D, Roemer FW, Guermazi A. Recent advances in research imaging of osteoarthritis with focus on MRI, ultrasound and hybrid imaging. *Clin Exp Rheumatol* 2018;36 (suppl 114[5]):43–52.
- Wang X, Oo WM, Linklater JM. What is the role of imaging in the clinical diagnosis of osteoarthritis and disease management? *Rheumatology (Oxford)* 2018;57(suppl 4):iv51–iv60.
- Oo WM, Linklater JM, Bennell KL, et al. Reliability and convergent construct validity of quantitative ultrasound for synovitis, meniscal extrusion, and osteophyte in knee osteoarthritis with MRI. *J Ultrasound Med* 2022;41(6):1559–1573.
- Wakefield RJ, Green MJ, Marzo-Ortega H, et al. Should oligoarthritis be reclassified? Ultrasound reveals a high prevalence of subclinical disease. *Ann Rheum Dis* 2004;63(4):382–385.
- Delle Sedie A, Riente L, Filippucci E, et al. Ultrasound imaging for the rheumatologist. XXXII. Sonographic assessment of the foot in patients with psoriatic arthritis. *Clin Exp Rheumatol* 2011;29(2):217–222.
- Sakellariou G, Conaghan PG, Zhang W, et al. EULAR recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis. *Ann Rheum Dis* 2017;76(9):1484–1494.
- Molyneux P, Bowen C, Ellis R, et al. Reliability of an ultrasound imaging acquisition procedure for examining osteoarthritis in the first metatarsophalangeal joint. *Res Sq*. Published online April 1, 2024. doi: [10.21203/rs.3.rs-3326416/v1](https://doi.org/10.21203/rs.3.rs-3326416/v1)
- Podlipská J, Guermazi A, Lehenkari P, et al. Comparison of diagnostic performance of semi-quantitative knee ultrasound and knee

- radiography with MRI: Oulu knee osteoarthritis study. *Sci Rep* 2016; 6(1):22365.
31. Hunter DJ. Osteoarthritis. *Best Pract Res Clin Rheumatol* 2011;25(6): 801–814.
  32. Molyneux P, Bowen C, Ellis R, et al. Evaluation of osteoarthritic features in peripheral joints by ultrasound imaging: a systematic review. *Osteoarthr Cartil Open* 2021;3(3):100194.
  33. Molyneux P, Bowen C, Ellis R, et al. Ultrasound imaging acquisition procedures for evaluating the first metatarsophalangeal joint: a scoping review. *Ultrasound Med Biol* 2022;48(3):397–405.
  34. Molyneux P, Bowen C, Ellis R, et al. International multispecialty consensus on how to image, define, and grade ultrasound imaging features of first metatarsophalangeal joint osteoarthritis, a Delphi consensus study. *Osteoarthr Cartil Open* 2023;5(1):100336.
  35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–174.
  36. Wongpakaran N, Wongpakaran T, Wedding D, et al. A comparison of Cohen's kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol* 2013;13(1):61.
  37. Steen Pettersen P, Neogi T, Magnusson K, et al. Associations between radiographic and ultrasound-detected features in hand osteoarthritis and local pressure pain thresholds. *Arthritis Rheumatol* 2020;72(6):966–971.
  38. Besselink NJ, Jacobs JWG, Westgeest AAA, et al. Can optical spectral transmission assess ultrasound-detected synovitis in hand osteoarthritis? *PLoS One* 2019;14(2):e0209761.
  39. D'Agostino MA, Terslev L, Aegerter P, et al. Scoring ultrasound synovitis in rheumatoid arthritis: a EULAR-OMERACT ultrasound taskforce-Part 1: definition and development of a standardised, consensus-based scoring system. *RMD Open* 2017;3(1):e000428.
  40. Mathiessen A, Hammer HB, Terslev L, et al. Ultrasonography of inflammatory and structural lesions in hand osteoarthritis: an OMERACT agreement and reliability study. *Arthritis Care Res (Hoboken)* 2021;74(12):2005–2012.
  41. Karim Z, Wakefield RJ, Quinn M, et al. Validation and reproducibility of ultrasonography in the detection of synovitis in the knee: a comparison with arthroscopy and clinical examination. *Arthritis Rheum* 2004; 50(2):387–394.
  42. Naredo E, Möller I, Moragues C, et al. EULAR Working Group for Musculoskeletal Ultrasound. Interobserver reliability in musculoskeletal ultrasonography: results from a "Teach the Teachers" rheumatologist course. *Ann Rheum Dis* 2006;65(1):14–19.
  43. Möller I, Janta I, Backhaus M, et al. The 2017 EULAR standardised procedures for ultrasound imaging in rheumatology. *Ann Rheum Dis* 2017;76(12):1974–1979.
  44. Messina C, Bignotti B, Tagliafico A, et al. A critical appraisal of the quality of adult musculoskeletal ultrasound guidelines using the AGREE II tool: an EuroAIM initiative. *Insights Imaging* 2017;8(5): 491–497.
  45. Flowers P, Nelson AE, Hannan MT, et al. Foot osteoarthritis frequency and associated factors in a community-based cross-sectional study of White and African American adults. *Arthritis Care Res (Hoboken)* 2021;73(12):1784–1788.
  46. Van der Merwe L, Duke C, Ung Y, et al. The Economic Cost of Arthritis in New Zealand in 2018. Deloitte Touche Tohmatsu Limited; 2018. <https://www.arthritis.org.nz/wp-content/uploads/2018/09/Economic-cost-of-Arthritis-in-New-Zealand-2018.pdf>