Title

Consensus approach for 3D joint space width of metacarpophalangeal joints of rheumatoid arthritis patients using high-resolution peripheral quantitative computed tomography

Journal

Quantitative Imaging in Medicine and Surgery, 10(2)

ISSN

2223-4292

Authors

Stok, Kathryn S
Burghardt, Andrew J
Boutroy, Stephanie
et al.

Publication Date

2020-02-01

DOI

10.21037/qims.2019.12.11

Peer reviewed

Original Article

Consensus approach for 3D joint space width of metacarpophalangeal joints of rheumatoid arthritis patients using high-resolution peripheral quantitative computed tomography

Kathryn S. Stok[1,2], Andrew J. Burghardt[3], Stephanie Boutroy[4], Michiel P. H. Peters[5,6,7], Sarah L. Manske[8,9], Vincent Stadelmann[2,10], Nicolas Vilayphiou[2], Joop P. van den Bergh[5,7,11,12], Piet Geusens[5,6,10], Xiaojuan Li[13], Hubert Marotte[14,15,16], Bert van Rietbergen[17,18], Steven K. Boyd[8,9], Cheryl Barnabe[8,19]; for the SPECTRA Collaboration

[1]Department of Biomedical Engineering, The University of Melbourne, Parkville, Australia; [2]SCANCO Medical AG, Brüttisellen, Switzerland; [3]Department of Radiology and Biomedical Imaging, University of California, San Francisco, USA; [4]INSERM 1033, Hôpital Edouard Herriot, Lyon Cedex, Lyon, France; [5]Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Centre, Maastricht, The Netherlands; [6]Research School CAPHRI, School for Public Health and Primary Care, Maastricht, The Netherlands; [7]NUTRIM School of Nutrition & Translational Research in Metabolism, Maastricht University, Maastricht, The Netherlands; [8]McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, Canada; [9]Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, Canada; [10]Department of Research and Development, Schulthess Klinik, Zürich, Switzerland; [11]Faculty of Medicine and Life Sciences, Hasselt University, Hasselt, Belgium; [12]VieCuri Medical Centre, Venlo, The Netherlands; [13]Program of Advanced Musculoskeletal Imaging (PAMI), Cleveland Clinic, Cleveland, OH, USA; [14]SAINBIOSE, INSERM U1059, University of Lyon, Saint-Etienne, France; [15]Department of Rheumatology, [16]INSERM CIE3 1408, University Hospital of Saint-Etienne, Saint-Etienne, France; [17]Orthopaedic Biomechanics, Department of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands; [18]Department of Orthopaedic Surgery, Research School CAPHRI, Maastricht University Medical Center, Maastricht, The Netherlands; [19]Department of Medicine and Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Canada

Correspondence to: Dr. Kathryn S. Stok. Department of Biomedical Engineering, The University of Melbourne, Parkville, VIC, 3010, Australia. Email: kstok@unimelb.edu.au.

Background: Joint space assessment for rheumatoid arthritis (RA) by ordinal conventional radiographic scales is susceptible to floor and ceiling effects. High-resolution peripheral quantitative computed tomography (HR-pQCT) provides superior resolution, and may detect earlier changes. The goal of this work was to compare existing 3D methods to calculate joint space width (JSW) metrics in human metacarpophalangeal (MCP) joints with HR-pQCT and reach consensus for future studies. Using the consensus method, we established reproducibility with repositioning as well as feasibility for use in second-generation HR-pQCT scanners.

Methods: Three published JSW methods were compared using datasets from individuals with RA from three research centers. A SPECTRA consensus method was developed to take advantage of strengths of the individual methods. Using the SPECTRA method, reproducibility after repositioning was tested and agreement between scanner generations was also established.

Results: When comparing existing JSW methods, excellent agreement was shown for JSW minimum and mean (ICC 0.987–0.996) but not maximum and volume (ICC 0.000–0.897). Differences were identified as variations in volume definitions and algorithmic differences that generated high sensitivity to boundary conditions. The SPECTRA consensus method reduced this sensitivity, demonstrating good scan-rescan reliability (ICC >0.911) except for minimum JSW (ICC 0.656). There was strong agreement between results from first- and second-generation HR-pQCT (ICC >0.833).

Conclusions: The SPECTRA consensus method combines unique strengths of three independently-developed algorithms and leverages underlying software updates to provide a mature analysis to measure 3D JSW. This method is robust with respect to repositioning and scanner generations, suggesting its suitability

for detecting change.

Keywords: X-ray computed tomography; 3D imaging; high-resolution peripheral quantitative computed tomography (HR-pQCT); arthritis; rheumatoid; metacarpophalangeal joint; OMERACT; peripheral quantitative CT; joint space width (JSW); three-dimensional; *in vivo*

Submitted Aug 06, 2019. Accepted for publication Dec 09, 2019.
doi: 10.21037/qims.2019.12.11
View this article at: http://dx.doi.org/10.21037/qims.2019.12.11

Introduction

Inflammation in rheumatoid arthritis (RA) results in disruption of the balance between osteoblasts and osteoclasts at periarticular regions resulting in bone erosion, and synovitis, activation of proinflammatory cytokines and antibodies which disrupt cartilage homeostasis leading to joint space thinning (1). These hallmarks of joint damage are closely linked to function (2), and for this reason, erosion and joint space width (JSW) are accepted outcome measures in rheumatology. The 2nd and 3rd metacarpophalangeal (MCP) joints are some of the most affected by RA, and the gold standard for assessing these outcomes are through a semi-quantitative analysis of radiograph images where trained readers score erosion and JSW using ordinal scales (3-6). This method is predisposed to floor and ceiling effects and significant progression is required to reach the incremental worsening threshold to detect score progression (7). Furthermore, projection errors due to overlapping bone features can confuse the definition of joint margins and the subsequent results.

Imaging methods, such as magnetic resonance imaging and computed tomography, provide three-dimensional (3D) images and suggest an opportunity for 3D JSW quantification that could be sensitive to disease change. Due to its high spatial resolution (90–140 microns) and superior contrast of bone margins compared to MRI, high-resolution peripheral quantitative computed tomography (HR-pQCT) has potential to provide precise quantification and visualization of morphologic changes for clinical trials and longitudinal observational studies of RA in peripheral joints, and potentially improve detection and evaluation of pre-erosive and pre-narrowing bony changes. While further refinement and testing of automated algorithms is in progress for erosion determination and measurement (4-6,8-10), significant developments have already occurred in the automated analysis of JSW for MCP joints (11-14).

In a recent systematic literature review (15) of the current approaches to joint space measurement, three techniques have been identified for fully automated, quantitative 3D measurements of JSW from HR-pQCT derived data (11-13). Two of these have demonstrated high-throughput, robust, reproducible capacity (11,13), while the third was a preliminary investigation of JSW width changes with finger flexion published as an abstract (12). The methods report not only mean JSW, but also measures for the minimum and maximum JSW, and the ratio of maximum to minimum JSW (asymmetry). JSW asymmetry is larger for joints where the joint space is thinner on one side than the other, or where there is evidence of joint subluxation. All three published methods were developed for single institution use with the first-generation HR-pQCT scanner (XtremeCT, SCANCO Medical AG, Brüttisellen, Switzerland). Two of the studies investigated reproducibility with repositioning (11,13) estimating precision errors of 4.8% and 12.5% in minimum JSW, respectively. When measuring the same joint in various angles of flexion (11,12) the studies highlighted the significant dependence of JSW measurement on position. However, there is no comparison of the methods to argue whether or not data can be interchangeably analysed by any of the three methods.

The SPECTRA collaboration is a global network of rheumatologists, clinicians, epidemiologists, engineers, radiologists, fellows, students, physicists, paediatricians, and industry partners. The principal aims of SPECTRA are to (I) investigate HR-pQCT for arthritis assessment in clinical trials, (II) validate HR-pQCT as a new imaging modality in clinical trials, and (III) harmonize SPECTRA's global efforts for efficiency. Previously SPECTRA has reached agreement on a protocol for imaging MCP joints in inflammatory arthritis with HR-pQCT (16), as well as a RA erosion definition for detection and measurement (17).

In this spirit of collaboration and expanding the availability of the algorithms to the global research community, SPECTRA has been working towards a consensus on a 3D method for calculating JSW from

human fingers measured with HR-pQCT (18). The goal of this work was to describe in detail the consensus approach, methodology and performance, namely, to compare the existing algorithms for face validity, reliability and ease of use, and reach consensus on a 3D method for calculating JSW from human fingers measured with the first generation HR-pQCT scanner (XtremeCT I, SCANCO Medical AG) that should be applied in ongoing and future observational studies and clinical trials for rheumatology. Using the consensus method, we established scan-rescan reproducibility and demonstrated the feasibility of extending the method for use in the second-generation HR-pQCT scanner (XtremeCT II, SCANCO Medical AG) which has since become available.

Methods

Study populations

Ethical approval was obtained from the respective institutional review boards; specifically, the Conjoint Health Research Ethics Board, University of Calgary (REB 15-0582) along with patient consent, the Human Research Protection Program at UCSF, and CPP Sud-Est in Lyon, France. Donors from Maastricht dedicated their body by testament signed during life to the Department of Anatomy and Embryology of the University of Amsterdam, The Netherlands. A handwritten and signed codicil from each donor, posed when still alive and well, is kept at the Department of Anatomy and Embryology, University of Amsterdam, Amsterdam, The Netherlands. The work presents a Level 3 cross-sectional analytic study design focussing on methods and technological development.

Dataset 1

For the initial comparison of outputs from the three published MCP JSW evaluation methods and determination of ease of use of the algorithms, thirty datasets of individuals with RA from the three sites [University of Lyon (LYN; n=9), University of California San Francisco (SFR; n=10), and University of Calgary (CLG; n=11)] were provided to the first author (Kathryn S. Stok) for use. The images represented a range of patients with no visible damage through to extensive erosive change, joint space loss and subluxation.

Dataset 2

To test repositioning reproducibility, twenty-one datasets from patients who were recruited to an HR-pQCT imaging study of early RA (diagnosis <1 year, mean ± SD age 49±11 years) through the Early Inflammatory Arthritis Clinics of the Division of Rheumatology at the University of Calgary were provided. All patients with RA fulfilled the 2010 American College of Rheumatology (ACR)/European League Against Rheumatism (EULAR) classification criteria for RA (19).

Dataset 3

To test reproducibility of the algorithm between first- and second-generation HR-pQCT scanner (XtremeCT I and XtremeCT II), ten finger samples from cadaveric donors—1 male, 9 female—from Maastricht University Medical Centre were used; age 86±9 years. The medical history of the donors is unknown.

HR-pQCT scanning procedures

All subjects were imaged using HR-pQCT (XtremeCT I, SCANCO Medical AG, Switzerland). Specifically, the first-generation XtremeCT was used for all data capture, except for Dataset 3 comparing both models of the HR-pQCT scanners (XtremeCT I and XtremeCT II); where all donor samples (n=10) were scanned on both machines.

For XtremeCT I measurements, the SPECTRA measurement protocol was used (16). Briefly, clinical *in vivo* settings of 60 kVp tube voltage, 900 µA tube current, 100 ms integration time, and an 82 µm isotropic voxel size were used for image acquisition. The reference line to determine the scan start location was placed at the midpoint of the concave articular surface at the base of the second or third proximal phalanx (whichever was the most distal between the two). The scan covered a length of 9.02 mm in the distal direction and 18.04 mm in the proximal direction (total scan length 27.06 mm, 330 slices, dose 9 µSv). The total scanning time was approximately 9 minutes.

For XtremeCT II measurements, clinical *in vivo* settings of 68 kVp tube voltage, 1,470 µA tube current, 43 ms integration time, and a 61 µm isotropic voxel size were used. The full hand was scanned, and then the 2nd and 3rd MCP joints were individually selected to match volumes defined on XtremeCT I scans.

For Dataset 2, each subject was scanned twice with repositioning: the subject removed their hand from the fixation device, had a 5-minute rest, and were repositioned for a second scan.

Subject scans were assessed for image quality, and images

Table 1 Summary of parametric differences, advantages and disadvantages between the three published methods for JSW measurement: CLG (11), LYN (12), and SFR (13)

JSW method	Joint space definition		Distance transform		
	Dilation	Erosion	Function	Advantages	Disadvantages
CLG	25**	25**	/dt_background	/close function faster than /dilation & / erosion sequence but limited to similar amplitudes	Same as SFR
LYN	25	31	/dt_object	Conservative definition of joint space via larger erosion than dilation reduces sensitivity to joint angle (see *Figure 1*)	/dt_object function sensitive to boundary conditions. Values on boundary of joint space may be truncated
SFR	20	20	/dt_background	/dt_background function reduces sensitivity to boundary effects; dilation/ erosion of similar amplitudes create a joint space volume that looks closer to intuitive definition	Larger joint space definition induces more sensitivity of DT to joint alignment/angle in DT at boundaries (see *Figure 1*)
Spectra consensus	20	26	/dt_spacing*	Conservative joint definition less sensitive to joint alignment/placement: new function /dt_spacing less sensitive to boundary conditions	

*, indicates function available since 2018 in the manufacturer's Image Processing Language IPL v5.42; **, indicates use of a single-step closing function in place of separate dilation/erosion steps. JSW, joint space width; SFR, University of California San Francisco; CLG, University of Calgary; LYN, University of Lyon.

of poor quality (grade >3) were excluded from further analysis.

Data processing

All data from both HR-pQCT models were processed using the manufacturer's Image Processing Language IPL v5.42. MCPs 2 and 3 were manually isolated from each dataset, and then each MCP joint was filtered using a 3D Laplace Hamming filter (20) (Hamming cut-off frequency =0.4). The resulting image was put through a custom algorithm written for rabbit tibio-femoral joints (21) which are a similar physical size to human MCP joints. The algorithm has been used in a number of different animals producing reproducible outcomes (21-25). It defines solid masks for the distal metacarpal head and proximal phalangeal base using a dual-threshold approach, where initially a high threshold value is used to identify the two joint parts and mask them, followed by a normal threshold to segment bone within that masked volume (21-23,25). These segmented and solid images were used as the same input for all subsequent JSW methods.

Statistical comparison of three JSW methods from literature

The code for three published methods [CLG (11), LYN (12), and SFR (13)] for JSW measurement was provided for this study to KS, and investigated by running Dataset 1 separately through each one. All three scripts make use of morphometric closing/opening operations to define the joint space in 3D. An image mask, defining the boundaries of the joint space is generated and then a distance transform (26) performed to obtain measures of width. *Table 1* summarises the key parametric differences between the three methods. Each algorithm provides a measure of average JSW (mm), and standard deviation, JSW.SD (mm), minimum JSW, JSW.min (mm), maximum JSW, JSW.max (mm), and joint space volume, JSV (mm^3), as described previously (13). A two-way mixed model intraclass correlation coefficient (ICC) (27) with a 95% confidence interval for all variables was used to test the consistency and absolute agreement between the individual outcomes of the three methods, while precision errors were calculated, expressed as both an absolute error, PE(SD), and as a percentage of the coefficient of variation,

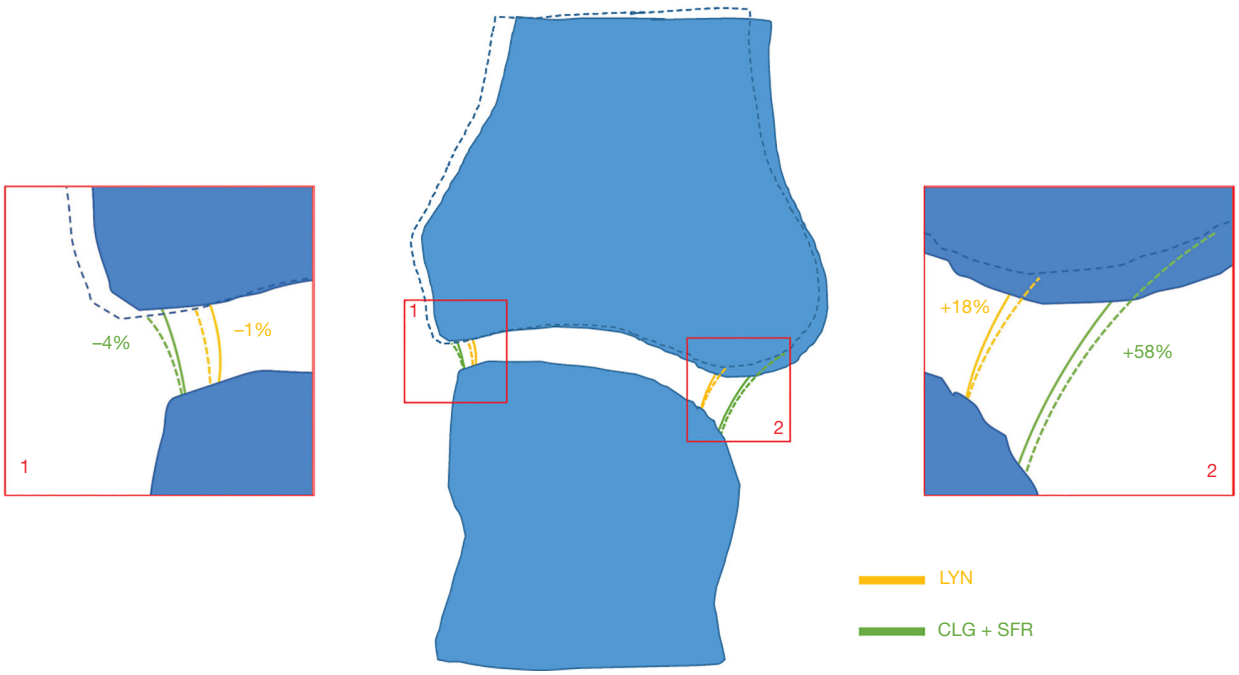

Figure 1 Conceptual illustration of the sensitivity of joint space volume definitions between the three original methods. A 2.5° rotation of the phalange (dotted lines) on joint space definition and border length has little effect on square edges (inset 1), while on rounded edges (inset 2) it has a large effect on CLG/SFR methods. SFR, University of California San Francisco; CLG, University of Calgary; LYN, University of Lyon.

PE (%CV) (28). Confidence intervals at 95% were calculated for the percent coefficient of variation assuming a chi-squared distribution as described (28). A univariate analysis of variance (with Bonferroni post-hoc) was used to test for significant differences, $P<0.05$, between the three methods, including consideration of finger (MCP2 or MCP3), and interaction between method and finger. Bland-Altman plots were drawn to check for proportional bias between the two generations of HR-pQCT scanner. All statistical analyses were conducted using SPSS (22.0, IBM, NY, USA), except Bland-Altman plots produced with R (v3.0.3). Additionally, visual analysis of the data was performed to identify where each script was giving anomalous results. Ease of use of the algorithm was subjectively assessed by KS.

SPECTRA JSW consensus method and reproducibility testing

Following the analysis of Dataset 1, it was determined that each script possessed pros and cons compared to the others; particularly around precision of the results. These were identified based on anomalies in the statistical and visual analyses, and well as the ease of use of each method. Thus, a SPECTRA JSW consensus method was developed to take advantage of the strengths and reduce the weaknesses of the three published scripts. In essence, the SPECTRA consensus method is based on that of Burghardt *et al.* (13) and Barnabe *et al.* (11) with an updated distance transform algorithm available from the manufacturer to correct numerical instabilities [achieving a similar definition to (12)]. Additionally, a larger number of erosions was used to reduce anomalies at the joint space boundary (see *Table 1*, *Figure 2*). Using the SPECTRA method, Dataset 2 was used to test its positioning-repositioning reproducibility, while Dataset 3 was used to test the reproducibility of JSW from data captured with both HR-pQCT models. Measures of JSW (mm), JSW.SD (mm), JSW.min (mm), JSW.max (mm), and JSV (mm^3) were calculated and evaluated. Statistical analysis was performed using SPSS (22.0, IBM, NY, USA). A two-way mixed model ICC (27) with a 95% confidence interval for all variables was used to test the consistency between either (I) repositioned subjects, or (II) data captured on

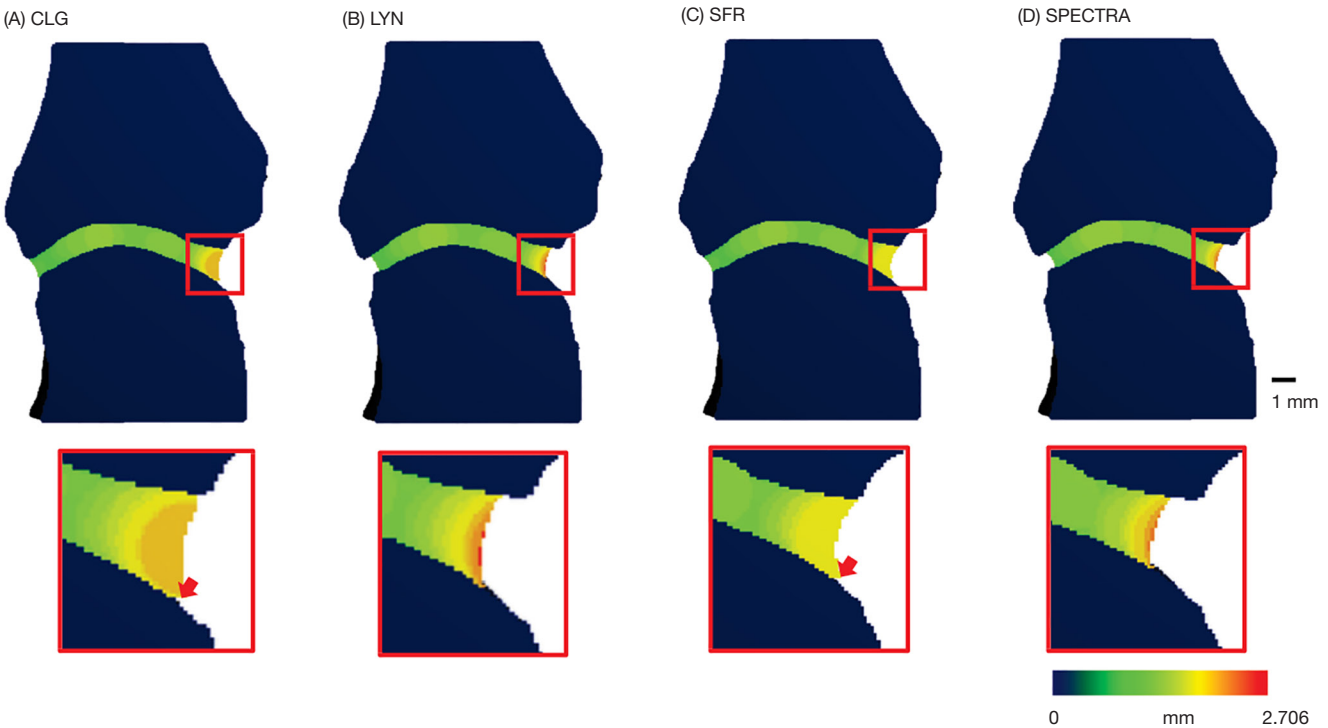

Figure 2 Visual comparison of the four JSW methods, inset highlighting differences (arrows) related to use of the distance transform functions affecting the joint space volume of interest, and the subsequent values for JSW.max. JSW, joint space width; SFR, University of California San Francisco; CLG, University of Calgary; LYN, University of Lyon

the two HR-pQCT models. Additionally, precision errors were calculated, expressed as both an absolute error, PE (SD), and as a percentage of the coefficient of variation, PE (%CV) (28).

Results

Comparing the three published methods for JSW measurement, *Table 2* indicates excellent consistency between each method for each parameter. Between methods there is excellent agreement for mean and minimum JSW, while the ICC measures for JSV (0.801) and JSW.max (0.340) indicate an issue in the definition of the mask volume and/or maximum. A univariate ANOVA of the three methods supports this finding with significant differences for JSV and JSW.max (*Table 3*); specifically LYN consistently gives a smaller volume and larger thickness than CLG and SFR solutions.

There is a significant difference for JSW, JSW.max and JSW.AS for the finger under investigation (that is, MCP2 versus MCP3), indicating that data from different fingers should not be pooled for comparison of means (*Table 3*).

There was no significant interaction for any measure between the individual joints (MCP2 and MCP3) and the applied method indicating the methods are independent from the joint being studied. So even though there is a significant difference in both JSV and JSW.max between MCP2 and 3, this is observed for all three methods (*Table 3*). A closer look at the relationship between the three methods for JSV and JSW.max, indicates that the three methods provide three significantly different JSW.max values, while JSV is significantly different between LYN and each of the other two methods (*Table 3*).

The visual analysis (*Figure 2*) supports a discordance in the definition of the joint space volume of interest between the three algorithms, which also affects the results for JSV and maximum JSW. *Figure 3* shows the general idea that all three approaches take to identify the joint space volume. All three scripts make use of morphometric opening and closing operations to define the volume of interest and then apply a distance transform algorithm (26), which in essence fits spheres in the masked volume of interest (*Figure 3* lower left) to obtain a measure of width. In the CLG and SFR methods, the order in which the distance transform and

Table 2 Consistency, absolute agreement and precision for three different JSW methods from Dataset 1 containing n=60 joints (30 MCP2 and 30 MCP3)

Measure	Consistency			Absolute agreement			Precision	
	ICC	Lower 95%	Upper 95%	ICC	Lower 95%	Upper 95%	PE (SD)	PE (%CV)
JSV (μm³)	0.967	0.950	0.979	0.801	0.161	0.934	13.22	10.60
JSW (mm)	0.991	0.987	0.995	0.973	0.823	0.991	0.04	2.25
Min (mm)	0.994	0.991	0.996	0.994	0.990	0.996	0.03	5.64
Max (mm)	0.843	0.772	0.897	0.340	0.000	0.650	0.30	12.46
AS (–)	0.908	0.863	0.941	0.893	0.825	0.935	1.21	13.34

JSV, joint space volume; JSW, joint space width; Min, minimum joint space width; Max, maximum joint space width; AS, joint asymmetry.

Table 3 Comparison of means to test the effect of JSW method (M), Finger (F), and interaction (MxF) on measured values

	JSV	JSW	Min	Max	AS
Univariate analysis of variance, with interaction[a]					
Method	0.000	n.s.	n.s.	0.000	n.s.
Finger	n.s.	0.015	n.s.	0.000	0.001
MxF	n.s.	n.s.	n.s.	n.s.	n.s.
One-way ANOVA on JSV and JSW comparing three methods[b]					
SFR-CLG	n.s.			0.008	
SFR-LYN	0.004			0.000	
CLG-LYN	0.000			0.000	

[a], comparison of means measured by three different JSW methods (M) from Dataset 1 containing n=60 joints (30 MCP2 and 30 MCP3, F = finger, indicating MCP2 or MCP3), P<0.05. [b], one-way ANOVA to compare means between method-pairs (where the univariate ANOVA identified significant differences in JSV and JSW.max for the applied method). JSV, joint space volume; JSW, joint space width; Min, minimum joint space width; Max, maximum joint space width; AS, joint asymmetry.

application of the boundary mask occurs, results in large spheres (indicating larger widths) not fully fitted within the VOI (e.g., red spheres in *Figure 3*) being excluded from the results. This additionally leads to a small gap appearing between the JS volume of interest and distal metacarpal head. In contrast, in the LYN script the order of DT and masking is reversed, which results in a definition of the maximum width with the size of the largest sphere, which can be considered the more accurate measure. However, in closing the joint space, the outcomes from the LYN method are highly dependent on the initial number of dilations. This can be seen (*Figure 2B*) by the smaller joint space volume of LYN compared to CLG and SFR methods.

The SPECTRA JSW consensus method (*Figure 2D* and *Table 1*) is a hybrid of the three published scripts, taking advantage of recent improvements to the distance transform algorithm, removing dependency on parameters for closing the joint space and volume definition, and ensuring ease of use.

Table 4 shows the results of the SPECTRA method used on Dataset 2 for scan-rescan reproducibility in an RA population. Minimum JSW reliability was sensitive to scan-rescan errors, particularly in joints with substantial narrowing, due to differences in regions with zero or near-zero thickness—resulting in a large PE (%CV). This is also observed in all other original methods (data not shown).

When comparing results for the same samples measured on XtremeCT I or II (Dataset 3), the two machines give highly consistent results with JSW differences in the region of 60 to 130 m; equivalent to 1–2 voxels on the two systems (*Table 5* and *Figure 4*).

Discussion

In order to standardise JSW measurements from HR-

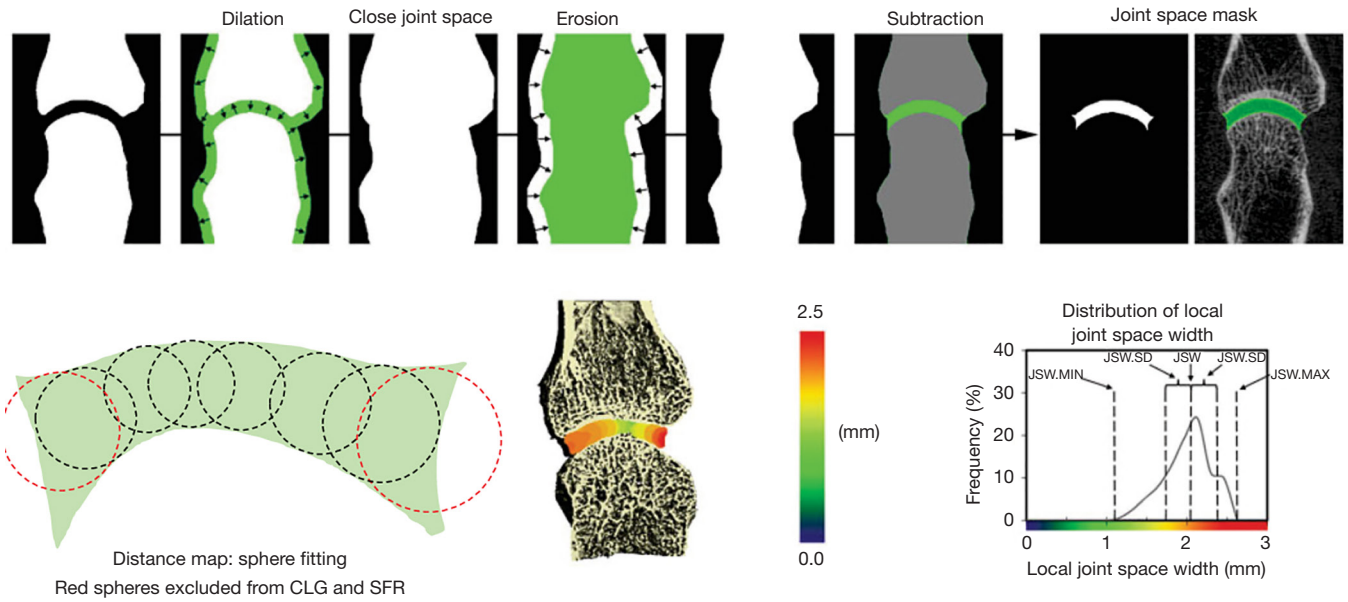

Figure 3 Schematic overview of the Burghardt *et al.* JSW method (13) illustrating the dilation and erosion functions used to close the joint space and subtraction with the original volume to create the joint space mask. Schematic illustration of distance map algorithm with sphere fitting. Red spheres are excluded in SFR and CLG methods. JSW, joint space width; SFR, University of California San Francisco; CLG, University of Calgary.

Table 4 Reproducibility of SPECTRA consensus method when measuring with re-positioning from Dataset 2 containing n=42 joints (21 subjects measured twice)

Measure	ICC	Lower 95%	Upper 95%	PE (SD)	PE (%CV)
JSV (μm³)	0.986	0.975	0.993	3.03	3.2
JSW (mm)	0.952	0.911	0.974	0.05	2.72
Min (mm)	0.656	0.356	0.816	0.18	23.48
Max (mm)	0.731	0.499	0.856	0.06	2.27
AS (–)	0.751	0.54	0.865	1.52	23.02

JSV, joint space volume; JSW, joint space width; Min, minimum joint space width; Max, maximum joint space width; AS, joint asymmetry.

Table 5 Reproducibility of SPECTRA consensus method when measuring with XtremeCT I and XtremeCT II for Dataset 3 containing n=10 joints

Measure	ICC	Lower 95%	Upper 95%	PE (SD)	PE (%CV)
JSV (μm³)	0.85	0.407	0.963	8.06	8.03
JSW (mm)	0.959	0.833	0.99	0.08	3.97
Min (mm)	0.937	0.757	0.984	0.13	13.2
Max (mm)	0.636	-0.332	0.907	0.06	2.19
AS (–)	0.896	0.585	0.974	0.42	12.59

JSV, joint space volume; JSW, joint space width; Min, minimum joint space width; Max, maximum joint space width; AS, joint asymmetry.

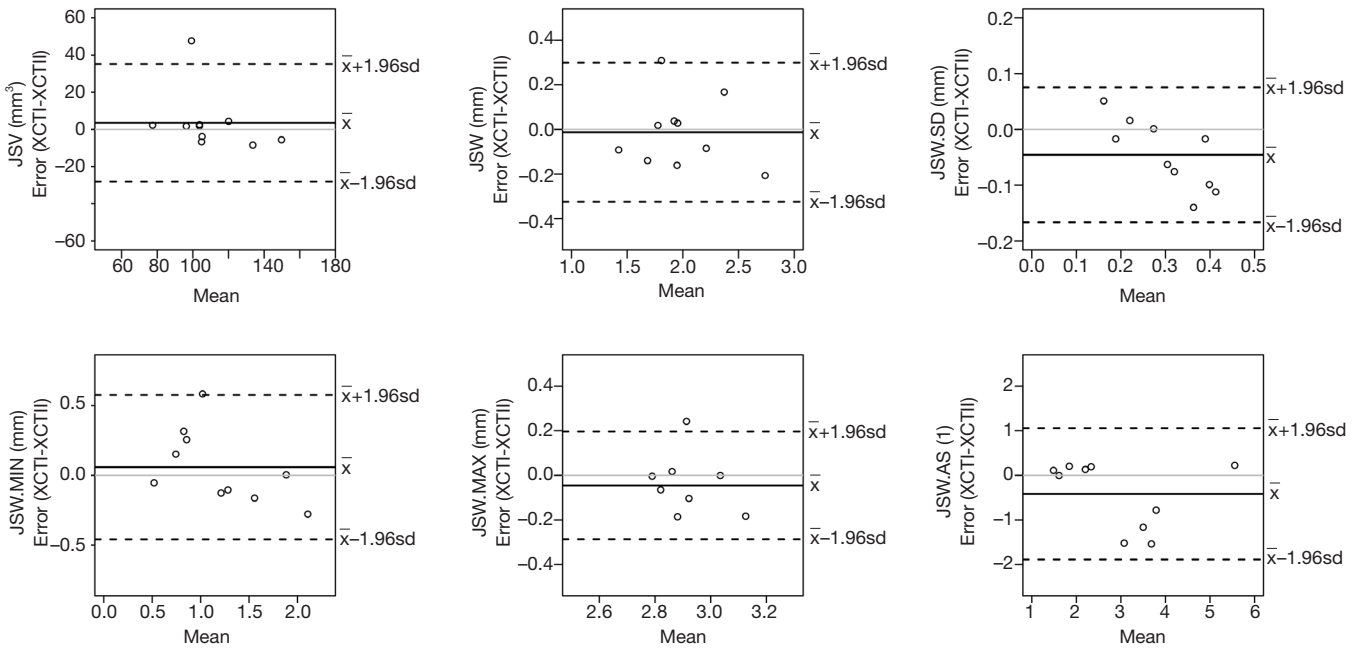

Figure 4 Bland-Altman plots comparing both models of HR-pQCT scanner (XtremeCT I and XtremeCT II), showing relatively even scatter at both low and high values of mean measures for each parameter, and indicating no obvious trend of increasing or decreasing scatter, except for standard deviation (JSW.SD). JSW, joint space width; HR-pQCT, high-resolution peripheral quantitative computed tomography.

pQCT images, we tested three existing JSW measurement methods using images of patients with RA, with a spectrum of MCPs from normal through to extensive arthritis damage. The results showed excellent consistency within each method, where all three methods were able to discriminate similar JSW trends across a spectrum of disease, and matching the results from the individual publications for each method (11-13). Additionally, for the absolute agreement between the three methods, there was no significant difference in the results for mean and minimum JSW, or the asymmetry measure (ratio of JSW.max to JSW.min). Therefore, results from different studies and using different calculation methods could be feasibly compared for all parameters, except for JSV and JSW.max.

We did identify limitations in the comparability of JSV and JSW.max however, with the three methods giving significantly different values. Examination of the three methods highlighted a smaller joint space VOI defined by the LYN method compared to SFR and CLG (*Figure 2*), where the definition of the joint space VOI leads to very different measurements (>10% $PE_{\%CV}$) for JSV and JSW.max depending on how many dilations are applied initially. The LYN method was able to accurately define the maximum JSW as the size of the largest sphere fitting between the two joint parts, where the CLG and SFR underestimate this value due to early exclusion in the algorithm of large widths near the boundary of the joint space volume. This is illustrated conceptually in *Figure 1*, where by using a larger VOI for the joint space, SFR and CLG are more sensitive to small variations in positioning when joint boundaries are more rounded than square, while LYN is less sensitive. These findings highlight the impact of apparently unimportant nuances, that eventually impact JSW.max. In essence we highlighted the need for a solution to handle the "border effects" at the edges of joint space volume, an issue that had previously been identified as a key point to address by SPECTRA before recommending JSW assessment method. From these findings, a SPECTRA consensus method based on the three published scripts was developed and tested, ensuring large spheres near the border of the joint space volume mask are not excluded and that there is no dependency on dilations used in the closing operation. As with the three original methods, the consistency within this method is excellent, and the method shows excellent agreement with the LYN results specifically for JSV and JSW.max.

We then used the SPECTRA JSW consensus method to evaluate robustness with repositioning, which gave

comparable results to the previously published methods. As with these, minimum JSW reliability was sensitive to scan/rescan errors, particularly in joints with substantial narrowing, due to differences in handling zones with zero or near-zero thickness where one to two voxels can have a significant impact on the result (*Figure 4*). This is not uncommon in measures of %CV when the mean value is close to zero, since as CV approaches infinity, PE (%CV) is sensitive to small changes in the mean. Finally, the SPECTRA JSW consensus method was also tested to demonstrate feasibility using the newer model HR-pQCT, the XtremeCT II, with a nominal isotropic resolution of 61 m (compared to 82 m in XtremeCT I). No difference in JSW measurements were found when measuring the same joints using the two HR-pQCT models. The relative scatter of JSW.min varied by as much as 0.5 mm for a 1 mm measurement (*Figure 4*) however the absolute errors are not larger than for other metrics indicating the sensitivity of JSW.min around the minimum. These differences can also be attributed in part to the dilation/erosion values used to offset the difference in voxel resolution between the two scanners.

It should be noted that cadaveric donors were used in Dataset 3 since it is rare to have both generations of scanner available at a single site. Furthermore, access to human material or ethics approval for multiple scans on single patients is difficult. Patient scanning may introduce motion artefact and affect JSW precision between scanner generations.

The SPECTRA JSW consensus method thus represents a novel tool in the assessment of joint space in RA. Not only will this allow more precise determination of longitudinal changes to joint space during the disease trajectory, and increase power to detect change in clinical trials, it will also allow us to study with high specificity hypotheses related to variations in and acute changes in state of MCP joints in RA. For example, it could be determined if the presence of joint effusion can increase the joint space, or what impact administration of anti-inflammatory medication may have on the joint. This method will be made widely available through the HR-pQCT manufacturer to ensure global uptake of a validated method. The next step to determine the most appropriate JSW parameters (if any) is dissemination and adoption of the method in the community for use in clinical trials and observational studies.

One caveat to this is that the data must all be filtered and segmented prior to evaluation, and there is no standardised segmentation approach for HR-pQCT finger scans. Further considerations include determining the impact of osteophytes (present in other pathologies affecting the MCPs such as psoriatic arthritis and osteoarthritis) on joint space measurements.

Future developments of the JSW algorithm involves investigating its value with other rheumatological burdens affecting the MCP joints, as well as extending it to other joints affected by arthritis conditions such as the knee. There is also work underway to extend the script to analyse joint space in animal models of arthritis.

In conclusion the high correlation between the 3D methods suggest that they all calculate similar JSW metrics, where volume definition has a significant effect on the JSV. The key differences between the three methods are related to distance transformation methods used, and a consensus method removing this dependency is proposed for common use in ongoing and future clinical trials for rheumatology.

Acknowledgments

None.

Footnote

Conflicts of Interest: KS Stok and V Stadelmann are former employees of Scanco Medical AG. N Vilayphiou is a current employee of Scanco Medical AG. The other authors have no conflicts of interest to declare.

Ethical Statement: Ethical approval was obtained from the respective institutional review boards; specifically, the Conjoint Health Research Ethics Board, University of Calgary (REB 15-0582) along with patient consent, the Human Research Protection Program at UCSF, and CPP Sud-Est in Lyon, France.

References

1. Schett G, Gravallese E. Bone erosion in rheumatoid arthritis: mechanisms, diagnosis and treatment. Nat Rev Rheumatol 2012;8:656-64.
2. Hayer S, Bauer G, Willburger M, Sinn K, Alasti F, Plasenzotti R, Shvets T, Niederreiter B, Aschauer C, Steiner G, Podesser BK, Smolen JS, Redlich K. Cartilage damage and bone erosion are more prominent determinants of functional impairment in longstanding experimental arthritis than synovial inflammation. Dis

Model Mech 2016;9:1329-38.
3. van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. J Rheumatol 2000;27:261-3.
4. Topfer D, Finzel S, Museyko O, Schett G, Engelke K. Segmentation and quantification of bone erosions in high-resolution peripheral quantitative computed tomography datasets of the metacarpophalangeal joints of patients with rheumatoid arthritis. Rheumatology 2014;53:65-71.
5. Toepfer D, Finzel S, Museyko O, Schett G, Engelke K. Longitudinal Quantification Of Bone Erosions In Patients With Rheumatoid Arthritis Using High-Resolution Computed Tomography. Arthritis Rheumatol 2013;65:S837-8.
6. Topfer D, Gerner B, Finzel S, Kraus S, Museyko O, Schett G, Engelke K. Automated three-dimensional registration of high-resolution peripheral quantitative computed tomography data to quantify size and shape changes of arthritic bone erosions. Rheumatology 2015;54:2171-80.
7. Peterfy C, Strand V, Tian L, Østergaard M, Lu Y, DiCarlo J, Countryman P, Deodhar A, Landewé R, Ranganath VK, Troum O, Conaghan PG. Short-term changes on MRI predict long-term changes on radiography in rheumatoid arthritis: an analysis by an OMERACT Task Force of pooled data from four randomised controlled trials. Ann Rheum Dis 2017;76:992.
8. Peters M, de Jong J, Scharmga A, van Tubergen A, Geusens P, Loeffen D, Weijers R, Boyd SK, Barnabe C, Stok KS, van Rietbergen B, van den Bergh J. An automated algorithm for the detection of cortical interruptions and its underlying loss of trabecular bone; a reproducibility study. BMC Med Imaging 2018;18:13.
9. Peters M, Scharmga A, de Jong J, van Tubergen A, Geusens P, Arts JJ, Loeffen D, Weijers R, van Rietbergen B, van den Bergh J. An automated algorithm for the detection of cortical interruptions on high resolution peripheral quantitative computed tomography images of finger joints. PLoS One 2017;12:e0175829.
10. Peters M, Scharmga A, van Tubergen A, Arts J, Loeffen D, Weijers R, van Rietbergen B, Geusens P, van den Bergh JP. The Reliability of a Semi-automated Algorithm for Detection of Cortical Interruptions in Finger Joints on High Resolution CT Compared to MicroCT. Calcif Tissue Int 2017;101:132-40.
11. Barnabe C, Buie H, Kan M, Szabo E, Barr SG, Martin L, Boyd SK. Reproducible metacarpal joint space width measurements using 3D analysis of images acquired with high-resolution peripheral quantitative computed tomography. Med Eng Phys 2013;35:1540-4.
12. Boutroy S, Hirschenhahn E, Youssof E, Locrelle H, Thomas T, Chapurlat R, Marotte H. Importance of hand positioning in 3D joint space morphology assessment. Arthr Rheum 2013;65:S840.
13. Burghardt AJ, Lee CH, Kuo D, Majumdar S, Imboden JB, Link TM, Li X. Quantitative in vivo HR-pQCT imaging of 3D wrist and metacarpophalangeal joint space width in rheumatoid arthritis. Ann Biomed Eng 2013;41:2553-64.
14. Tom S, Frayne M, Manske SL, Burghardt AJ, Stok KS, Boyd SK, Barnabe C. Determining Metacarpophalangeal Flexion Angle Tolerance for Reliable Volumetric Joint Space Measurements by High-resolution Peripheral Quantitative Computed Tomography. J Rheumatol 2016;43:1941-4.
15. Nagaraj S, Finzel S, Stok KS, Barnabe C. High-resolution Peripheral Quantitative Computed Tomography Imaging in the Assessment of Periarticular Bone of Metacarpophalangeal and Wrist Joints. J Rheumatol 2016;43:1921-34.
16. Barnabe C, Feehan L; SPECTRA (Study GrouP for XTrEme-CT in RA). High-resolution peripheral quantitative computed tomography imaging protocol for metacarpophalangeal joints in inflammatory arthritis: the SPECTRA collaboration. J Rheumatol 2012;39:1494-5.
17. Barnabe C, Toepfer D, Marotte H, Hauge EM, Scharmga A, Kocijan R, Kraus S, Boutroy S, Schett G, Keller KK, de Jong J, Stok KS, Finzel S; SPECTRA Collaboration. Definition for Rheumatoid Arthritis Erosions Imaged with High Resolution Peripheral Quantitative Computed Tomography and Interreader Reliability for Detection and Measurement. J Rheumatol 2016;43:1935-40.
18. Stok KS, Finzel S, Burghardt AJ, Conaghan PG, Barnabe C; SPECTRA Collaboration. The SPECTRA Collaboration OMERACT Special Interest Group: Current Research and Future Directions. J Rheumatol 2017;44:1911-5.
19. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd, Birnbaum NS, Burmester GR, Bykerk VP, Cohen MD, Combe B, Costenbader KH, Dougados M, Emery P, Ferraccioli G, Hazes JM, Hobbs K, Huizinga TW, Kavanaugh A, Kay J, Kvien TK, Laing T, Mease P, Menard HA, Moreland LW, Naden RL, Pincus T, Smolen JS, Stanislawska-Biernat E, Symmons D, Tak PP, Upchurch KS, Vencovsky J, Wolfe F, Hawker G. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum

2010;62:2569-81.
20. Laib A, Ruegsegger P. Comparison of structure extraction methods for in vivo trabecular bone measurements. Comput Med Imaging Graph 1999;23:69-74.
21. Stok KS, Besler BA, Steiner TH, Villarreal Escudero AV, Zulliger MA, Wilke M, Atal K, Quintin A, Koller B, Muller R, Nesic D. Three-Dimensional Quantitative Morphometric Analysis (QMA) for In Situ Joint and Tissue Assessment of Osteoarthritis in a Preclinical Rabbit Disease Model. PLoS One 2016;11:e0147564.
22. Stok KS, Noel D, Apparailly F, Gould D, Chernajovsky Y, Jorgensen C, Muller R. Quantitative imaging of cartilage and bone for functional assessment of gene therapy approaches in experimental arthritis. J Tissue Eng Regen Med 2010;4:387-94.
23. Stok KS, Pelled G, Zilberman Y, Kallai I, Goldhahn J, Gazit D, Muller R. Revealing the interplay of bone and cartilage in osteoarthritis through multimodal imaging of murine joints. Bone 2009;45:414-22.
24. Kohler T, Stauber M, Donahue LR, Muller R. Automated compartmental analysis for high-throughput skeletal phenotyping in femora of genetic mouse models. Bone 2007;41:659-67.
25. Besler BA, Sondergaard RE, Muller R, Stok KS. Reproducibility of compartmental subchondral bone morphometry in the mouse tibiofemoral joint. Bone 2015;81:649-53.
26. Hildebrand T, Rüegsegger P. A new method for the model-independent assessment of thickness in three-dimensional images. J Microsc 1997;185:67-75.
27. Shrout PE. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420-8.
28. Gluer CC, Blake G, Lu Y, Blunt BA, Jergas M, Genant HK. Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int 1995;5:262-70.

Cite this article as: Stok KS, Burghardt AJ, Boutroy S, Peters MPH, Manske SL, Stadelmann V, Vilayphiou N, van den Bergh JP, Geusens P, Li X, Marotte H, van Rietbergen B, Boyd SK, Barnabe C; for the SPECTRA Collaboration. Consensus approach for 3D joint space width of metacarpophalangeal joints of rheumatoid arthritis patients using high-resolution peripheral quantitative computed tomography. Quant Imaging Med Surg 2020;10(2):314-325. doi: 10.21037/qims.2019.12.11

