Article:

Finzel, S, Manske, SL, Barnabe, C et al. (10 more authors) (2020) Reliability and Change in Erosion Measurements by High-Resolution peripheral Quantitative Computed Tomography in a Longitudinal Dataset of Rheumatoid Arthritis Patients. The Journal of Rheumatology. jrheum.191391. jrheum.191391-jrheum.191391. ISSN 0315-162X

https://doi.org/10.3899/jrheum.191391

Reliability and Change in Erosion Measurements by High-Resolution peripheral Quantitative Computed Tomography in a Longitudinal Dataset of Rheumatoid Arthritis Patients

Running Head: Responsiveness in HR-pQCT Imaging

Stephanie Finzel, MD[1, 2], Sarah L. Manske[3,4], PhD, Cheryl Barnabe, MsC, MD[3], Andrew J. Burghardt, BS[5], Hubert Marotte, MD, PhD[6,7,8], Andrea Scharmga, PhD[9], Ellen-Margrethe Hauge, MD, PhD[10], Roland Chapurlat, MD, PhD[11], Klaus Engelke, PhD[2], Xiaojuan Li, PhD[5,12], Bente van Teeffelen[13], Philip G. Conaghan, MD, PhD[14], Kathryn S. Stok, PhD[13,15]

Affiiations:
[1]Department of Rheumatology and Clinical Immunology, Medical Center – University of Freiburg, Medical Faculty, University of Freiburg, Freiburg, Germany; [2]Department of Medicine 3, FAU University Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany; [3]Departments of Medicine and Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, Canada; [4]Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, Canada; [5]Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, United States of America; [6]INSERM 1059/SAINBIOSE, Université Jean Monet, Université de Lyon, Saint-Etienne, France, [6]Department of Rheumatology, CHU Saint-Etienne, Saint-Etienne, France, [8]INSERM CIE3 1408, *Université de Lyon, Saint-Etienne, France; [9]Maastricht University, Maastricht, The Netherlands; [10]Department of Rheumatology and Department of Clinical Medicine, Aarhus University Hospital and Aarhus University, Aarhus, Denmark; [11]INSERM 1033, Hôpital Edouard Herriot, 69003 Lyon Cedex, Lyon, France; [12]Department of Biomedical Engineering, Cleveland Clinic, Cleveland, United States of America, [13]Department of Biomedical Engineering, Melbourne School of Engineering, The University of Melbourne, Victoria 3010 Australia, [14]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Biomedical Research Centre, Leeds, UK; [15]Institute for Biomechanics, ETH Zurich, Zurich, Switzerland.*

ORCID IDs:

Stephanie Finzel: https://orcid.org/0000-0002-7625-9612

Sarah L. Manske: https://orcid.org/0000-0003-0888-9556

Cheryl Carmelle Marie Barnabe: https://orcid.org/0000-0003-3761-237X

Andrew J. Burghardt: https://orcid.org/0000-0002-6343-4944

Hubert Marotte: https://orcid.org/0000-0003-1177-9497

Andrea Scharmga: https://orcid.org/0000-0001-8940-8827

Ellen-Margrethe Hauge: https://orcid.org/0000-0003-2562-9174

Roland Chapurlat: https://orcid.org/0000-0001-8214-6385

Klaus Engelke: https://orcid.org/0000-0001-9875-4123

Xiaojuan Li: https://orcid.org/0000-0002-0567-9935

Bente van Teeffelen: https://orcid.org/0000-0002-6765-870X

Philip G. Conaghan: https://orcid.org/0000-0002-3478-5665

Kathryn Stok: https://orcid.org/0000-0002-0522-4180

Submitted to the Journal of Rheumatology as Brief Communication.

MeSH KEYWORDS: rheumatoid arthritis, computed tomography, metacarpophalangeal joint, outcomes

Funding and conflicts of interest: RELEX-2 was hosted by the University of San Francisco, California, United States of America. The meeting was sponsored in part by Scanco Medical AG. PGC is supported in part by the UK National Institute for Health Research Leeds Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Disclosures: The authors state that they have no conflicts of interest.

Initials, surnames, appointments, and highest academic degrees:

S. Finzel, MD, Senior Attending Physician, Department of Rheumatology and Clinical Immunology, Medical Center – University of Freiburg, Medical Faculty, University of Freiburg, Freiburg, Germany

S.L. Manske, PhD, Assistant Professor, Department of Radiology, Cumming School of Medicine, University of Calgary, Calgary, Canada

C.C.M. Barnabe, MD, MSc, Associate Professor, Departments of Medicine and Community Health Sciences, University of Calgary, and McCaig Institute for Bone and Joint Health, University of Calgary, Calgary, Canada

A.J. Burghardt, BS, Research Specialist, Department of Radiology and Biomedical Imaging, University of California, California, United States of America

Hubert Marotte, MD, PhD, Institut national de la santé et de la recherchémédicale, and Rheumatology Department, University Hospital of Saint-Etienne, France

Andrea Scharmga, PhD, Maastricht University, Maastricht, The Netherlands

Ellen-Margrethe Hauge, MD, PhD, Professor, Department of Rheumatology, Aarhus University Hospital and Department of Clinical Medicin, Aarhus University, Aarhus, Denmark

Roland Chapurlat, MD, PhD, Professor, INSERM 1033, Hôpital Edouard Herriot, 69003 Lyon Cedex, Lyon, France

Klaus Engelke, PhD, Professor, Department of Medicine 3, FAU University Erlangen-Nürnberg and Universitätsklinikum Erlangen, Erlangen, Germany

Xiaojuan Li, PhD, Professor, Department of Biomedical Engineering, Cleveland Clinic, Cleveland, United States of America

Bente van Teeffelen, Department of Biomedical Engineering, Melbourne School of Engineering, The University of Melbourne, Victoria 3010 Australia

P.G. Conaghan, MD, PhD, Professor, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds and NIHR Leeds Biomedical Research Centre, Leed, United Kingdom

K.S. Stok, PhD, Senior Lecturer, Institute for Biomechanics, ETH Zurich, and Department of Biomedical Engineering, University of Melbourne, Melbourne, Australia.

Corresponding Author:

Dr. med. Stephanie Finzel, Senior Attending Physician, Head of Clinical Trials Unit Rheumatology, Department of Rheumatology and Clinical Immunology, University Medical Center, Medical Faculty, University of Freiburg, Hugstetter Strasse 55, 79110 Freiburg, Germany; Phone: +49 761 270 74493, Fax: +49 761 270 74491; email: Stephanie.Finzel@uniklinik-freiburg.de

Number of words in abstract: 230

Number of words in manuscript: 1626

Number of tables: 3

Supplementary tables: 1

Supplementary figures: 1

Abstract

Objectives: The aim of this multi-reader exercise was to assess the reliability and change over time of erosion measurements in rheumatoid arthritis (RA) patients using high-resolution peripheral quantitative computed tomography (HR-pQCT).

Methods: HR-pQCT scans of 23 patients with RA were assessed at baseline and 12 months. Four experienced readers examined the dorsal, palmar, radial, and ulnar surfaces of the metacarpal head (MH) and phalangeal base (PB) of the 2nd and 3rd digits, blinded to time order. In total, 368 surfaces (23 patients x16 surfaces) were evaluated per time point to characterize cortical breaks as pathological (erosion) or physiological, and to quantify erosion width and depth. Reliability was evaluated by intraclass correlation coefficients (ICC), percentage agreement, and Light's kappa; change over time was defined by means ± SD of erosion numbers and dimensions.

Results: ICCs for the mean measurements of width and depth of the pathological breaks ranged between 0.819 - 0.883, and 0.771 - 0.907 respectively. Most physiological cortical breaks were found at the palmar PB, whereas most pathological cortical breaks were located at the radial MH. There was a significant increase in both the numbers and the dimensions of erosions between baseline and follow-up (p=0.0001 for erosion numbers, width, and depth in axial plane, and p=0.001 for depth in perpendicular plane).

Conclusion: This exercise confirmed good reliability of HR-pQCT erosion measurements and their ability to detect change over time.

INTRODUCTION:

High-Resolution peripheral Quantitative Computed Tomography (HR-pQCT) provides accurate detection of periarticular bone changes, which is required for diagnosis and therapeutic monitoring in rheumatoid arthritis (RA) *(1)*. Previously, the Study grouP for xtrEme Computed Tomography in Rheumatoid Arthritis (SPECTRA) collaboration presented a consensus definition for bone erosion, and a common approach for measuring erosion size, with feasibility and preliminary reliability tested in a cross-sectional dataset of RA metacarpophalangeal (MCP) joints (RELEX-1) *(2)*. Good agreement was demonstrated regarding the presence and nature of cortical breaks; however agreement for measuring erosion dimensions needed refinement. We therefore performed this multi-reader HR-pQCT exercise in order to assess the reliability of erosion measurements and to evaluate change over time in RA patients.

METHODS

Images:

Twenty-three seropositive RA patients underwent HR-pQCT imaging of their second and third digit of their dominant hand at baseline (BL, 0 months) and follow-up (FUP, 12 months) at the Universities of Erlangen-Nuremberg, Lyon, San Francisco, and Calgary. Patients were selected according to the presence of bone erosions on X-ray as assessed by the van der Hejde/Sharp-score, and the need to change therapy due to insufficient disease control. Local ethics approval and written informed consent were obtained prior to study entry (IRB numbers: Calgary REB15-0582; San Francisco 12-10418; Lyon CPP: 13/083;

Erlangen 3839). All participants were scanned using a first generation HR-pQCT scanner (Scanco Medical AG, Bruettisellen, Switzerland) and standard acquisition settings were applied with an 82 micrometers (µm) isotropic voxel size resolution *(3)*. Image data sets were viewed using Osirix (version 5.8). Readers were blinded to clinical status and time-sequence of images.

Prior to image evaluation, the four readers calibrated measurements using a test set of images. All of the readers participating had at least 3 years' experience in reading HR-pQCT data sets.

Joint image evaluation:

The image evaluation algorithm involves assessing eight surfaces at each of the second and third MCP joints, specifically the palmar, dorsal, radial, and ulnar surfaces of each of the proximal phalangeal base (PB) and the metacarpal head (MH) *(1)*. Only images of sufficient quality were evaluated *(4)*. Individual surfaces were analyzed for the presence of cortical breaks (present or absent) according to the SPECTRA definition: the cortical break should be present in 2 consecutive slices and 2 perpendicular planes, and should show a loss of underlying trabecular bone. The cortical break is characterized as being pathological (erosion) or physiological, with the former described as a non-linear appearance typical of erosions, and the latter as a parallel/linear break typical of vessel channels *(3)*; ***supplemental figure 1*** gives an example for typical pathological or physiological cortical breaks. For erosions, readers quantified the size of the break by measuring the maximum width and corresponding depth in both axial and perpendicular planes to each surface. The depth of the cortical break was recorded on the same slice where the maximal width was obtained. All measures were quantified in millimeters (mm). Readers noted whether multiple cortical breaks were present on the same surface but only recorded measurements for the largest cortical break.

Statistical analysis:

The inter-reader reliability of the detection of cortical breaks was evaluated using the percentage of agreement and Light's kappa for the chance corrected agreement. *(6)* The intra-class correlation (ICC) was calculated as an indicator of variability in cortical break depth and width measurements between readers 1, 2, 3, and 4. Paired-sample T-test was used to evaluate the longitudinal changes between BL and FUP scans of each subject. Analyses were performed with SPSS (version 23).

RESULTS

Patient's Characteristics:

Mean age ± standard deviation (SD) was 46 (13) years, 60% women, mean disease duration 2.3 (±2.8) years, and mean DAS28 at baseline was 3.51 ± 1.03. There was no significant change detectable over time in van der Heijde/Sharp-score. All patients received methotrexate; 18 patients were also treated with or started on a tumor necrosis factor alpha-inhibitor.

Images from two time points (BL-FUP) for 23 subjects were evaluated, resulting in 46 individual joints with 368 unique surfaces that were evaluated per time point. Thus in total 736 surfaces were evaluated.

Evaluability of images

The percentage agreement for evaluability of all the images between all readers was 80% (589/736). The chance-corrected agreement was fair (Cohen's kappa=0.218; ranges for all individual reader pairs 0.005-0.519). Only the surfaces in which all four readers agreed that the image was evaluable were included beyond this step (n=585). Evaluability was affected by the presence of motion artifacts and/or technical artifacts such as stack artifacts.

Presence of cortical breaks

The percentage agreement for the presence or absence of cortical breaks on all evaluable images between all readers was 57% (334/585). The chance-corrected agreement resulted in a moderate k value of 0.493. Cohen's kappa for all individual reader pairs (reader 1 vs reader 2, etc.) ranged between 0.405-0.551.

Characterization of cortical breaks

In total, 99 cortical breaks were identified on baseline and follow-up images. The percentage agreement for the appearance of a cortical break as pathological or physiological between all readers was 81% (80/99). The chance-corrected agreement resulted in a substantial k value of 0.796. Cohen's kappa for all individual reader pairs ranged between 0.765-0.838.

Numbers and localizations of erosions and physiological cortical breaks

Table 1 shows the number of breaks (total and erosions) for the eight individual surfaces in which all readers agreed on the presence of a cortical break. The distribution of cortical breaks confirmed findings from previous publications. *(3;7,8)*

Widths and Depths of Erosions

There were 41 cortical breaks detected as erosions by all readers. ***Table 1*** shows the mean dimensions with SD measured by all readers of these 41 erosions for the respective surfaces the erosions were detected in.

Inter-reader agreement regarding measurements of the sizes of cortical breaks

Numbers and dimensions of cortical breaks were determined on surfaces where all readers agreed that an erosion was present (n=41). ICC was calculated; for all four measures the ICC was high: mean values ±SD and ICC for erosion numbers, axial width and depth, as well as perpendicular width and depth were 1.39±0.62, ICC 0.803; 2.31±1.39, ICC 0.883; 1.85±0.86, ICC 0.907; 1.99±0.87, ICC 0.819, and 1.89±0.91, ICC 0.771, respectively (see ***table 2*** for details; for further measures of precision see ***supplemental table 1***).

Longitudinal change of cortical breaks over time

All pairs of measures (BL and FUP) were evaluated and the mean BL and FUP values were compared to test for significant differences over time. In total, all pairs from all readers gave 285 pairs. Mean values (±SD) of erosion numbers, widths and depths are shown in ***table 3***. There was a significant increase in both the numbers of erosions and the dimensions of the cortical breaks between BL and FUP scans (all $p<0.01$).

DISCUSSION

In this multi-reader responsiveness exercise, we applied HR-pQCT imaging to assess reliability and change over time of erosion measurements in a dataset of patients with RA. We applied our consensus definition of bone erosion as well as a previously agreed evaluation algorithm. *(3)* The exercise yielded good reliability for HR-pQCT measurements (ICC > 0.771) and a significant increase was observed in both number and dimensions of erosions between baseline and follow up ($p<0.01$). Furthermore, most physiological cortical breaks were found at the palmar PB, whereas most erosions were located at the radial MH; the distribution of erosions and physiological cortical breaks confirmed the findings from earlier studies (3;7,8).

Agreement (ICCs) for erosion numbers, width and depth of cortical breaks were high, kappas for appearance of cortical breaks were good. The reliability measures in this study revealed better results than in the RELEX-1 exercise. *(3)* For the current exercise, we used only 4, not 11 readers as in the first exercise with pre-study calibration. *(3)* It should be noted that we used four readers, unlike the two readers typically used in a clinical trial, and that images were read in unknown time order, which may also reduce responsiveness. A limitation of the study might be that only those surfaces were analyzed further, in which all readers agreed that the image quality was sufficient, and a cortical break was present, which reduced the number of analyzable surfaces. This emphasizes the need for adequate training before reading HR-pQCT images. On the other hand, this could be overcome by developing semi-automated algorithms, allowing for volumetric assessment of pathological cortical breaks.

The analysis of change over time yielded highly significant values for mean ± SD of number, width and depth of cortical breaks. Our findings showed responsiveness over time despite small sample sizes and achieving disease control.

Ongoing work has evaluated the nature of small cortical breaks. Boutroy et al *(9)* performed a perfusion study on a cadaveric hand using contrast perfusion, confirming the location of vascular foramen and their comparative frequency in periarticular bone. This provides construct validity for the SPECTRA erosion definition. Scharmga et al compared vascular foramen in histology and in HR-pQCT. *(10)* Perhaps not surprisingly due to differences in spatial resolution, there was a substantially higher number of vessel channels found in histology than in HR-pQCT. It needs to further be assessed, however, how uniquely identified HR-pQCT small cortical breaks are of added value in RA monitoring.

While the assessment of radiographic joint space width in HR-pQCT may be semi-automated, *(11-13)* the evaluation algorithm of cortical breaks still requires training and time *(14,15)*. Therefore, our collaboration is pursuing the investigation of a common technical algorithm for semi- or fully automated erosion detection and measurement allowing for volumetric erosion assessment. *(16)*

In conclusion, HR-pQCT evaluation using trained readers allows for highly reliable and precise detection of cortical breaks and facilitates differentiation of pathological from physiological cortical breaks. Reading by less experienced readers results in fair kappa values with regards to evaluability and break detection. Moreover, our results suggest that HR-pQCT responsiveness of erosion measures over time.

Acknowledgements: The paper was written on behalf of all SPECTRA collaboration members.

SPECTRA Collaboration Members:

Cheryl Barnabe, University of Calgary; Anne-Birgitte Blavnsfeldt, University of Aarhus; Stephanie Boutroy, Université de Lyon; Steven K Boyd, University of Calgary; Andrew Burghardt, University of California San Francisco; Roland Chapurlat, Université de Lyon; Angela Cheung, University of Toronto; Ko Chiba, University of Nagasaki; Joost de Jong, Maastricht University Medical Centre; Klaus Engelke, University of Erlangen; Stephanie Finzel, University of Freiburg; Ursula Heilmeier, University of California San Francisco; Harry Genant, University of California, San Francisco; Piet Geusens, Maastricht University; Ellen-Margrethe Hauge, Aarhus University Hospital; Joost de Jong, Maastricht University; Denis Julien, Université de Lyon; Rashid Kapadia, Scanco, USA; Kresten Keller, Aarhus University Hospital; Roland Kocijan, University of Vienna; Sebastian Kraus, Kantonsspital Baden; Eric Lespessailles, University of Orleans; Xiaojuan Li, University of Cleveland; Sarah Manske, University of Calgary; Hubert Marotte, Saint-Etienne, Université de Lyon; Liam Martin, University of Calgary; Michiel Peters, Maastricht University; Valentina Pedoia, University of California San Francisco; Andrea Scharmga, Maastricht University; Georg Schett, University of Erlangen; Kathryn S. Stok, The University of Melbourne; Nikolay Tzaribachev, Bad Bramstedt; Joop van den Bergh, Maastricht University; Bert van Rietbergen, Eindhoven University of Technology; Tomohiro Shimuzu, University of California San Francisco; Lai-Shan Tam, University of Hong Kong; Karen Troy, University of Worchester; Mira van Veenendaal, University of

Toronto; Nicolas Vilayphiou, Scanco, Switzerland; Paul Willems, Maastricht University Medical Centre; Rae Yeung, University of Toronto.

REFERENCES

1. Stach CM, Bauerle M, Englbrecht M, Kronke G, Engelke K, Manger B, et al. Periarticular bone structure in rheumatoid arthritis patients and healthy individuals assessed by high-resolution computed tomography. Arthritis Rheum 2010;62:330-9.
2. Barnabe C, Toepfer D, Marotte H, Hauge EM, Scharmga A, Kocijan R, et al. Definition for Rheumatoid Arthritis Erosions Imaged with High Resolution Peripheral Quantitative Computed Tomography and Interreader Reliability for Detection and Measurement. J Rheumato. 2016;43:1935-1940.
3. Barnabe C, Feehan L. High-resolution peripheral quantitative computed tomography imaging protocol for metacarpophalangeal joints in inflammatory arthritis: the SPECTRA collaboration. J Rheumatol 2012;39:1494-5.
4. Pialat JB, Burghardt AJ, Sode M, Link TM, Majumdar S. Visual grading of motion induced image degradation in high resolution peripheral computed tomography: impact of image quality on measures of bone density and micro-architecture. Bone 2012;50:111–18.
5. Light RJ. Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychol Bull 1971:76:365.
6. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 2005;85:257-68. Review.
7. Finzel S, Ohrndorf S, Englbrecht M, Stach C, Messerschmidt J, Schett G, et al. A detailed comparative study of high-resolution ultrasound and micro-computed tomography for detection of arthritic bone erosions. Arthritis Rheum 2011;63:1231-6.
8. Finzel S, Rech J, Schmidt S, Engelke K, Englbrecht M, Schett G. Interleukin-6 receptor blockade induces limited repair of bone erosions in rheumatoid arthritis: a micro CT study. Ann Rheum Dis 2013;72:396-400.
9. Boutroy S, Chapurlat R, Vanden-Bossche A, Locrelle H, Thomas T, Marotte H. Erosion or vascular channel? Arthritis Rheumatol. 2015;67:2956.
10. Scharmga A, Keller KK, Peters M, van Tubergen A, van den Bergh JP, et al. Vascular channels in metacarpophalangeal joints: a comparative histologic and high-resolution imaging study. Sci Rep. 2017;7:8966.
11. Barnabe C, Buie H, Kan M, Szabo E, Barr SG, Martin L, et al. 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. Burghardt AJ, Lee CH, Kuo D, Majumdar S, Imboden JB, Link TM, et al. Quantitative in vivo HR-pQCT imaging of 3D wrist and metacarpophalangeal joint space width in rheumatoid arthritis. Ann Biomed Eng 2013;41:2553-64.
13. 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;44:1911-1915.
14. Töpfer 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 (Oxford) 2014;53:65-71.
15. Töpfer D, Gerner B, Finzel S, Kraus S, Museyko O, Schett G,et al. Automated three-dimensional registration of high-resolution peripheral quantitative computed

tomography data to quantify size and shape changes of arthritic bone erosions. Rheumatology (Oxford) 2015;54:2171-80.
16. Peters M, Scharmga A, de Jong J, van Tubergen A, Geusens P, Arts JJ, et al.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.

Figure Legend:

Supplemental Figure 1: The third metacarpophalangeal joint of a 25-year-old female patient with rheumatoid arthritis is shown in axial plane. White arrows indicate a typical pathological cortical break (erosion) at the radial site (image A), and a typical physiological cortical break (vessel channel) at the palmar site (image B).