

RMD
OpenRheumatic &
Musculoskeletal
Diseases

EXTENDED REPORT

Computerised versus conventional methodology of radiographic joint destruction assessment in early rheumatoid arthritis

Yinghe Huo,^{1,2} Maria J H De Hair,¹ Yasmin O Shaib,¹ Désirée van der Heijde,³ Natalia O Kuchuk,¹ Max A Viergever,² Jacob M van Laar,¹ Koen L Vincken,² Floris P Lafeber,¹ On behalf of the Society for Rheumatology Research Utrecht investigators

To cite: Huo Y, De Hair MJH, Shaib YO, *et al*. Computerised versus conventional methodology of radiographic joint destruction assessment in early rheumatoid arthritis. *RMD Open* 2015;1:e000148. doi:10.1136/rmdopen-2015-000148

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/rmdopen-2015-000148>).

Received 3 July 2015
Revised 3 October 2015
Accepted 1 November 2015



CrossMark

¹Department of Rheumatology & Clinical Immunology, UMC Utrecht, Utrecht, The Netherlands

²Image Sciences Institute, UMC Utrecht, Utrecht, The Netherlands

³Department of Rheumatology, LUMC Leiden, Leiden, The Netherlands

Correspondence to

Maria JH De Hair;
M.J.H.deHair@umcutrecht.nl

ABSTRACT

Objectives: To compare computerised and conventional methodology of radiographic joint destruction assessment in early rheumatoid arthritis (RA).

Methods: We investigated the contribution of the 3rd-to-5th carpometacarpal joints (CMC3-5, which are excluded in computerised assessment so far owing to bone overlapping) to total joint space narrowing (JSN) scores in two cohorts of patients with early RA (n=392). Next, we investigated agreement between JSN scoring using single time point individual joint-based method (individual joint of a single time point (IJSTP), reflecting computerised reading) and conventional JSN scoring using the Sharp-van der Heijde (SvdH) method in a cohort of patients with early RA (n=59). We used intraclass correlation coefficients (ICCs), Bland and Altman plots, and linear mixed modelling to analyse differences in progression between two methods. Radiographs were available at baseline, and at 1 and 2 years of follow-up.

Results: Of all joints affected by JSN at baseline or JSN progression during 2 years of follow-up, 3.9% and 6.6% concerned CMC3-5. Exclusion of CMC3-5 resulted in a decrease of 1.9–4.6% in JSN progression scores during 2 years of follow-up. The ICCs for JSN progression scores using IJSTP with or without CMC3-5 compared with SvdH were 0.71–0.81 and 0.69–0.78 at 1 and 2 years of follow-up. Signal-to-noise ratios for IJSTP-based and SvdH scoring were 0.51 and 0.58, respectively. The progression rate for each year was not statistically significantly different between two scoring methods (p=0.59 and 0.89).

Conclusions: This study showed that excluding CMC3-5 has limited influence on JSN (progression) scores and showed the feasibility of using IJSTP-based reading for computerised scoring of JSN (progression) in RA.

INTRODUCTION

Rheumatoid arthritis (RA) is a chronic inflammatory disease that affects the joints of

Key messages

What is already known about this subject?

- Computerised methods have the potential to improve the discriminative power of radiographic progression assessment in particular in the early stage of rheumatoid arthritis.
- Conventional and computerised joint space narrowing (JSN) assessment procedures of hand joints take into account different joints and make use of different scoring strategies.

What does this study add?

- The joints not included in the computerised JSN assessment have a limited contribution to JSN (progression) scores.
- The JSN (progression) scores determined by an individual joint-based method (reflecting computerised reading) and the conventional Sharp-van der Heijde (SvdH) method are not significantly different on the group level.

How might this impact on clinical practice?

- The methodological differences between conventional and computerised JSN assessment of hand joints have limited influence on JSN scores. This increases the feasibility to use computerised methods in clinical trials and maybe even in clinical practice.

predominantly hands and feet. The Sharp-van der Heijde (SvdH) scoring method is currently the gold standard for assessment of radiographic progression in RA clinical trials, consisting of assessments of both erosions and joint space narrowing (JSN).^{1 2} However, over the past years, treatment of RA has been significantly improved by 'tight control' and 'treat-to-target' strategies,^{3–5} as well as the introduction of

biological treatment regimens.^{6 7} As a result radiographic progression has significantly decreased and it is important that scoring methods are sensitive to be able to detect small changes, particularly in the early stage.^{7 8}

In order to try to improve the discriminative power of radiographic progression assessment, researchers are focusing on computerised analysis. Computerised analysis uses continuous rather than ordinal scales for radiographic joint destruction quantification, an inherent advantage over conventional scoring methods. On the contrary, computerised analysis is less rater dependent and possibly more time efficient. So far, the computerised joint destruction assessment mainly focuses on the early stage of RA, with subtle to mild joint destruction. Predominant attention has been paid to JSN via joint space width (JSW) quantification. Research on erosions, however, is still limited since reconstructing an accurate original bone contour is difficult for computers owing to large shape variation of hand bones; estimating erosion volumes from the projected images is even more challenging. As for computerised quantification of JSW, several promising results have been achieved. For instance, fully automatic JSW quantification can be conducted on finger joints and semiautomatic JSW quantification (manual involvement required) can be performed on foot joints and some wrist joints.^{9–11} In spite of these encouraging achievements, there are some issues to be addressed, which are related to methodological differences with the conventional scoring methods. In the present paper, we have focused on two issues.

The first issue concerns the difference in the joints to be assessed between the computerised and the conventional scoring methods. So far, no computerised method is able to assess JSW of the third through the fifth carpometacarpal joints (CMC3-5) accurately, owing to bone overlapping in the two-dimensional (2D) plane (see online supplementary figure S1 for an image of the wrist joints used for JSN assessment). Excluding the CMC3-5 joints in the computerised assessment could be a compromise, but the influence on total JSN needs to be established. Several studies support low JSN incidence rates of the CMC3-5 joints.^{12 13} However, since with the modern effective treatments the progression per year in total Sharp score is small, Landewé *et al*¹³ concluded that every joint that shows progression contributes to the individual total Sharp progression score. Still, the percentage of patients having JSN or progression of JSN in CMC3-5 joints was below 4% and 2%, respectively. We aim to explore the contribution of the CMC3-5 joints to JSN in a cohort of patients with early RA, to examine to what extent excluding the CMC3-5 joints from the conventional SvdH scores influences joint destruction assessment.

The second issue concerns the difference between the reading strategies of the computerised method and the conventional SvdH scoring method. To assess joint radiographic progression in clinical trials with the

conventional SvdH method, surrounding joints of the same hand (or foot) as well as radiographs of both the previous and the follow-up time points are taken into account.^{14 15} In contrast, computerised assessment of joint destruction is purely mathematical and so far based on an individual joint at a single time point only, making it a different assessing procedure. This may result in different scores between the two methods. An advantage of the conventional SvdH scoring method is that image acquisition errors due to possible hand rotation can be reduced by observing the surrounding joints. However, this advantage may become less pronounced now because hospitals increasingly standardise the image acquisition procedure so as to obtain a reproducible assessment. Therefore, in this study, we investigated the agreement between individual joint-based scoring making use of a single time point only (to simulate the computerised assessment method) and the conventional SvdH scoring method under the current radiograph acquisition conditions, as to examine the feasibility of using individual joint-based computerised methods for assessment of radiographic progression. In addition, we investigated the influence of excluding the CMC3-5 joints from the individual joint-based scoring method.

MATERIALS AND METHODS

Patients

For the first analysis radiographs of all patients included in the 2-year double-blind randomised placebo-controlled Computer-Assisted Management of Early RA (CAMERA) I and II trials were used. These studies have been extensively described previously.^{4 16} In short, patients fulfilled the 1987 revised American College of Rheumatology (ACR) criteria for RA,¹⁷ were disease-modifying antirheumatic drug (DMARD) naïve and had disease duration of <1 year. Radiographs of baseline, 1 and 2 years of follow-up were available of 392 patients (219 of CAMERA I and 173 of CAMERA II).

For the second analysis 59 of 70 preselected early DMARD naïve patients with RA were included, with annual radiographs available at three consecutive time points. The exclusion reasons for the 11 patients were wrist incompleteness and unreadable low-quality images.

All patients included have given informed consent according to Dutch medical ethical regulations.

Scoring of radiographic joint damage

In the first group of patients radiographs were scored according to the conventional SvdH method by two blinded rheumatologists, of which the consensus score was used as described in the original publications.^{4 16}

In the second group of patients two raters scored using two different reading methods with the assistance of a newly developed software package (Imaging Sciences Institute, UMC Utrecht). Scorings using two reading methods were conducted with a 4-week interval.

In the first reading method, a prelocated (by the operator) individual joint of a single time point (IJSTP) could be observed. In the second reading method, all joints of one region (eg, metacarpophalangeal (MCP) joints of the right hand) and the follow-up images displaying the same region were visible. Scale rulers were attached to the interface of both reading methods to aid scoring (see online supplementary figure S2). For clarity, only manual scoring was performed, but we simulated computerised scoring (purely mathematical based on an individual joint at a single time point only) by scoring IJSTP, which is an important characteristic of computerised JSN assessment. We compared this with conventional manual scoring, by scoring joints while taking into account surrounding joints and multiple time points, which is an important characteristic of conventional manual scoring according to the SvdH method.

Contribution of the CMC3-5 joints to the incidences of JSN and JSN progression

The first aim is to investigate to what extent excluding the CMC3-5 joints influences joint destruction assessment. The incidence of JSN at baseline and the incidence of JSN progression during the 2-year trial period for each type of joint were investigated. JSN was defined as a score ≥ 1 and JSN progression was defined by a change in JSN score. The incidence was calculated by dividing the number of each type of joint affected by JSN (or JSN progression) by the total number of joints with the same type in the cohort (twice the number of patients because both left-sided and right-sided joints were included). Additionally, to provide the relative contribution of each joint type to the total incidence of JSN (or JSN progression), the percentage of each type of joint affected by JSN (or JSN progression) relative to the total number of joints affected by JSN (or JSN progression) was calculated. Since measuring JSN of the CMC3-5 joints is a specific challenge for computerised JSN assessment in RA, we also calculated the score percentages of the affected CMC3-5 joints relative to the JSN and the total SvdH scores and progression scores.

Agreement between IJSTP scoring and conventional SvdH scoring methods for assessment of progression

Since progression scores rather than status scores are most critical for monitoring RA activity and assessment of effectiveness of drug treatment, the agreement between the progression scores of the IJSTP and the conventional SvdH scoring method were investigated. The influence of the exclusion of the CMC3-5 joints from the IJSTP method on radiographic progression was evaluated via agreement analysis with the conventional SvdH scoring method. In the study, the average of the two independent raters was used. Bland and Altman plots were created for both JSN and erosion progression scores to visualise the agreement between the two scoring methods for individual participants. In addition,

two-way random absolute agreement intraclass correlation coefficient (ICC) analyses were determined, along with 95% CIs. Analysis of variance were used to estimate the variance components corresponding to all sources of variation. For this simple design in which patients were crossed with time, variance components can be estimated by the random effects model for the three sources of variations: patient (σ_p^2), time (σ_t^2) and residual ($\sigma_{pt,e}^2$). To examine the efficiency of the two reading methods on progression assessment, signal-to-noise ratios of the progression scores were calculated. This was done by dividing the square root of signal variance σ_t^2 by the square root of noise variance $\sigma_{pt,e}^2$.^{18 19}

To determine statistically significant differences in progression between the two scoring methods a linear mixed model was fitted to the data, making use of a random intercept. Scoring method (using conventional SvdH scoring as reference), follow-up time (linear) and the interaction between scoring method and follow-up time were included in the model as independent variables. A *p* value < 0.05 was considered statistically significant. IBM SPSS Statistics, V.21, was used for all statistical analyses.

RESULTS

Contribution of the CMC3-5 joints to the incidences of JSN and JSN progression

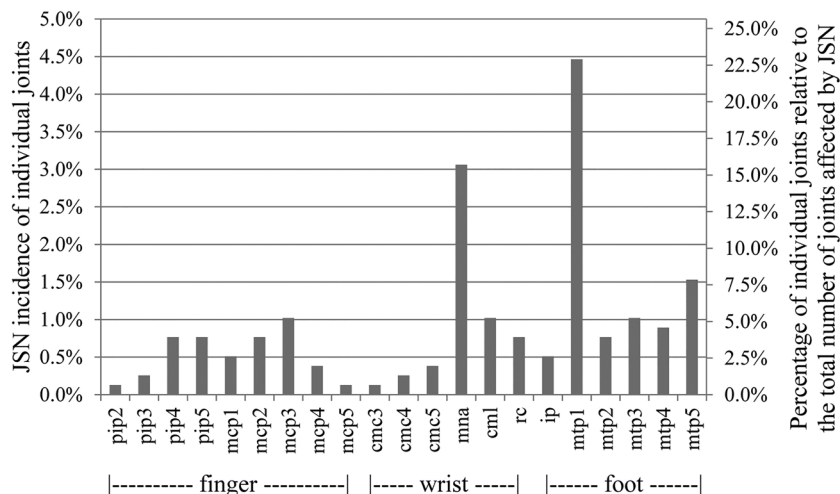
Of the 392 patients, 72 patients (18%) had at least one joint with a JSN score ≥ 1 at baseline and 103 (26%) patients had progression of JSN during the 2-year trial period.

Figure 1 shows the JSN incidence of each type of joint at baseline (left y axis). At the right y axis the percentage of individual joints relative to the total number of joints affected by JSN (relative contribution to JSN damage) is shown. The most frequently affected joint was MTP-1, with a JSN incidence of 4.5%. The most frequently affected hand joint was present in the wrist (multangular-navicular joint, MNA, 3.1%). The JSN incidence of the finger joints was not as high as the MTP-1 or MNA joints, but the total contribution of the finger joints to the number of joints affected by JSN was still considerable (24.2%).

Figure 2 provides the percentage of the joints that showed JSN progression in the first 2 years. The MTP-1 and MNA joints were most frequently affected by JSN progression and contributed the most to the total number of joints affected by JSN progression.

Next, we focused on the CMC3-5 joints because of the difficulty in obtaining an automatic JSN assessment of these joints. Overall, the CMC3-5 contributions to JSN assessment were limited. The incidence of the CMC3-5 joints at baseline was relatively low, with a sum of $< 1\%$. The percentage of the CMC3-5 joints affected by JSN was 3.9% relative to all the joints affected by JSN at baseline, and 6.6% of all the joints showed JSN progression during the 2-year period.

Figure 1 JSN incidence of individual joints (left axis) and percentage of individual joints relative to the total number of joints affected by JSN (right axis) at baseline. JNS, joint space narrowing; PIP2-5, second to fifth proximal interphalangeal joints; MCP1-5, first to fifth metacarpophalangeal joints; CMC3-5, third to fifth carpometacarpal joints; MNA, multangular-navicular joint; CNL, capitate-navicular-lunate joint; RC, radiocarpal joint; IP, first interphalangeal joint of the foot; MTP1-5, first to fifth metatarsophalangeal joints.



Exclusion of the CMC3-5 joints resulted in a decrease of 2.8–4.2% and 1.3–2.0% compared with the original scores, respectively, for the JSN and the total SvdH score at the three time points (figure 3A,B). The JSN and total SvdH progression scores for the first year and the 2-year period decreased with 1.9–4.6% and 0.8–2.1%, respectively (figure 3C,D). The raw data are shown in online supplementary table S1.

Agreement between the IJSTP and conventional SvdH scoring methods for assessment of progression

Next we investigated the agreement between the IJSTP and conventional SvdH scoring methods for JSN progression scores. Bland and Altman plots are shown in figure 4. The score differences concern the scores of the IJSTP scoring method (with and without the CMC3-5 joints) minus the scores of the conventional SvdH scoring method. Compared with the conventional scoring method, IJSTP scoring with and without the CMC3-5 joints both systematically underscored JSN progression (difference -0.3 to -0.2 and -0.5 to -0.4 ,

respectively). However, compared with the total progression scores, these systematic differences were relatively small (see online supplementary table S2 for the mean status and progression scores).

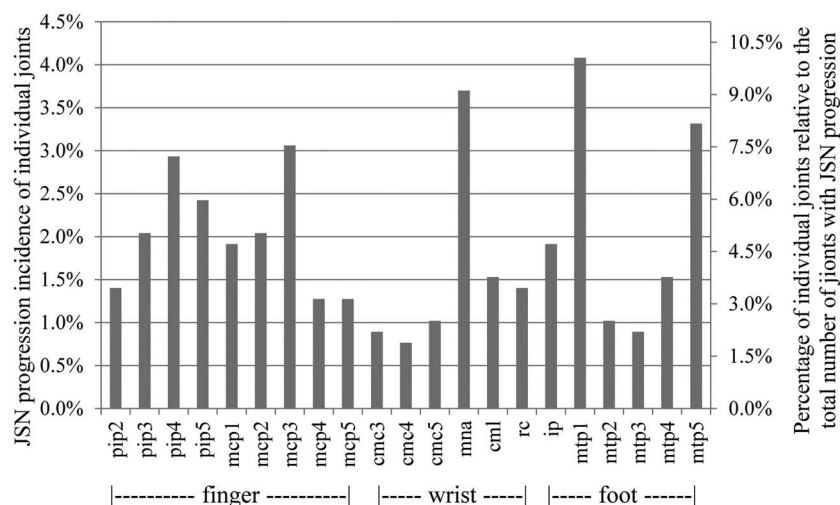
The two-way random absolute agreement ICCs for the JSN progression scores were high for both follow-up time points using the IJSTP method with and without the CMC3-5 joints, with values 0.71 to 0.81 and 0.69 to 0.78, respectively.

The signal-to-noise ratios for JSN progression as derived from the conventional SvdH scoring method and the IJSTP scoring methods with and without the CMC3-5 joints were 0.58, 0.51 and 0.57, respectively.

Compared with the conventional SvdH scoring method, IJSTP with and without the CMC3-5 joints both detected somewhat less JSN progression. However, CIs were largely overlapping and these differences are therefore not statistically significant (see table 1).

For all data regarding erosion results see online supplementary results section.

Figure 2 JSN progression incidence of individual joints and the percentage of individual joints relative to the total number of joints with JSN progression in the first 2 years. JNS, joint space narrowing; PIP2-5, second to fifth proximal interphalangeal joints; MCP1-5, first to fifth metacarpophalangeal joints; CMC3-5, third to fifth carpometacarpal joints; MNA, multangular-navicular joint; CNL, capitate-navicular-lunate joint; RC, radiocarpal joint; IP, first interphalangeal joint of the foot; MTP1-5, first to fifth metatarsophalangeal joints.



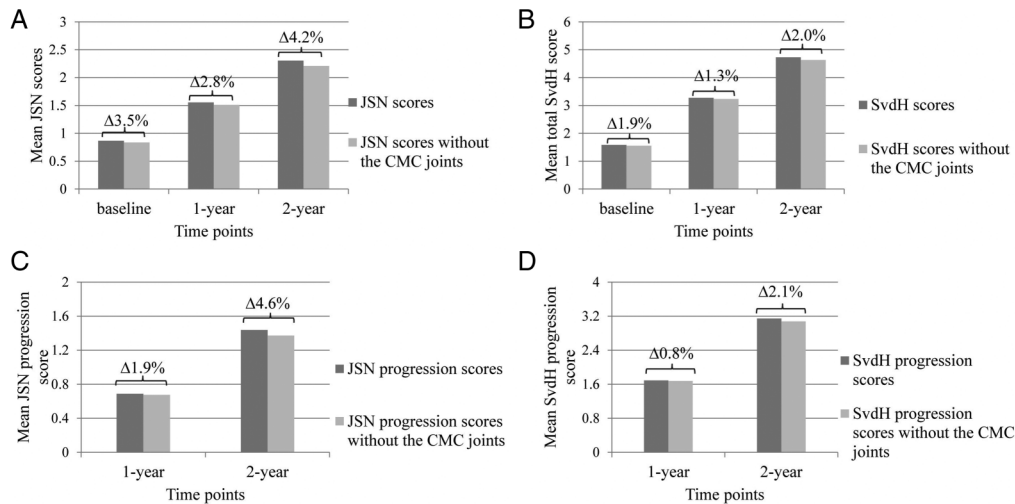


Figure 3 JSN and SvdH damage and progression scores with and without the CMC3-5 joints. The mean score of all patients with and without the CMC3-5 joints at each time point included: (A) JSN score; (B) total SvdH score. The mean progression score of all patients with and without the CMC3-5 joints included at each time point from baseline: (C) JSN, (D) total SvdH. The percentage in the figure depicts the extent of decrease in the score when excluding the CMC joints (compared with the original score). JNS, joint space narrowing; CMC3-5, third to fifth carpometacarpal joints; SvdH, Sharp-van der Heijde scoring.

DISCUSSION

In this study, we evaluated two methodological differences between computerised scoring and conventional SvdH scoring of JSN (progression): (1) so far no computer-based method is able to evaluate CMC3-5 joints (owing to bone overlapping) which are included in the SvdH scoring method; (2) computer-based assessment of JSN (progression) is individual joint and single time point based, while the conventional SvdH scoring method takes into account the surrounding joints and may take the previous and the follow-up time points into

account. The aim of the study was to explore to what extent excluding the CMC3-5 joints influenced JSN (progression) scores as well as the feasibility of using an IJSTP-based method for assessment of JSN (progression) in the early stage of RA.

We showed low incidence rates of affected CMC3-5 joints, and small contributions of the CMC3-5 joints to the JSN and the total SvdH (progression) scores. Excluding the CMC3-5 joints has a limited influence on JSN assessment on group level, with a progression score decrease of 1.9–4.6% for the first year and the 2-year

Figure 4 Bland and Altman plots showing the JSN progression score difference (y axis) against the mean JSN progression scores (x axis) of the two reading methods. (A) 1-Year JSN progression score (IJSTP vs SvdH) and (B) 2-year JSN progression score (IJSTP vs SvdH). (C): 1-Year JSN progression score (IJSTP without CMC3-5 vs SvdH) and (D) 2-year JSN progression score (IJSTP without CMC3-5 vs SvdH). JNS, joint space narrowing; CMC3-5, third to fifth carpometacarpal joints; SvdH, Sharp-van der Heijde scoring; IJSTP, individual joint of a single time point.

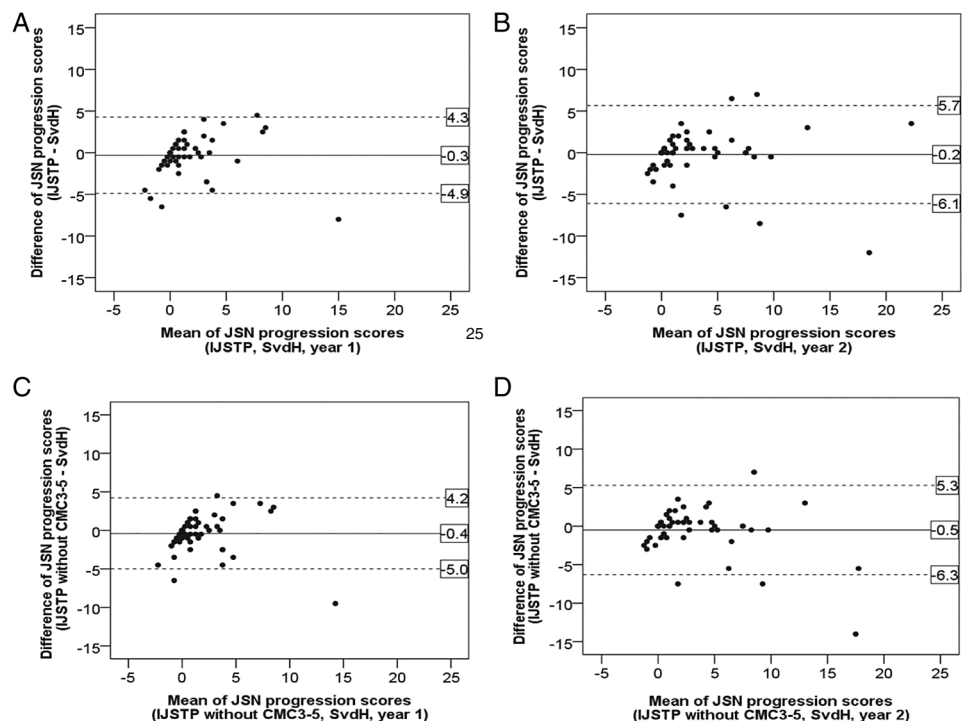


Table 1 JSN progression rate for years 1 and 2, based on linear mixed modelling, using conventional SvdH scoring and IJSTP-based scoring with and without CMC3-5

Assessment method of JSN	Progression rate year 1 (95% CI)	p Value	Progression rate year 2 (95% CI)	p Value
Conventional SvdH scoring (reference)	1.70 (0.92 to 2.49)	–	1.47 (0.65 to 2.29)	–
IJSTP	1.40 (0.58 to 2.22)	0.59	1.56 (0.55 to 2.56)	0.89
IJSTP without CMC3-5	1.33 (0.55 to 2.11)	0.50	1.36 (0.63 to 2.10)	0.85

CMC3-5, third to fifth carpometacarpal joints; IJSTP, individual joint of a single time point; JSN, joint space narrowing; SvdH, Sharp-van der Heijde scoring.

period. Our observation was in line with that of Hulsmans *et al.*¹² but differed slightly from what Landewé *et al.*¹³ have shown, in terms of individual proximal interphalangeal (PIP) and MCP joint incidences (a lower incidence of the affected PIP joints and a higher incidence of the affected MCP joints were shown in their study). The most likely explanation is the difference in the study cohorts because Landewé *et al.* only included early methotrexate (MTX) naïve patients with RA and established patients with RA failing successful MTX treatment and our cohort consisted of early MTX naïve patients with RA with either prednisone or placebo. Although excluding the CMC3-5 joints has a limited influence on JSN assessment on group level, we admit that it has a potential risk of JSN misdetection for a single patient with predominant CMC3-5 joint destructions. This is the inevitable downside of the joint exclusion, but on a group level this would have minor relevance.

The feasibility of using individual joint-based computerised methods for RA assessment in the early stage was investigated by evaluating the agreement in progression score between the IJSTP scoring method (simulating computerised assessment methods) and the conventional SvdH scoring method. To the best of our knowledge, there is no literature available to compare our results with. Our study shows that the mean score difference between the two scoring methods was relatively small and there was no statistically significant difference between the progression scores of the two reading methods on the group level. These results suggest that individual joint-based computerised methods might assess joint damage progression in a reliable manner in the early stage of RA. As expected, exclusion of the CMC3-5 joints from IJSTP scoring resulted in only a slight decrease in progression scores. These small differences may not relevantly affect the comparison of treatment arms in clinical trials. Despite good statistical agreement of the results using the two scoring methods, we should notice the limited power of the study due to the small number of patients involved. While both IJSTP and conventional SvdH scoring detected yearly joint damage progression, the mean JSN progression score and signal-to-noise ratio were somewhat lower for the IJSTP scoring method than for the conventional SvdH scoring method. The assessment procedure of the

current computerised method therefore still has room for improvement, for example, using the radiograph longitudinal information for abnormal JSW change detection and rectification. In addition, computerised radiographic progression assessment may benefit from standardised procedures in acquisition conditions, such as using a hand mold.

Acknowledgements The authors thank Maal van Everdingen and Marja Lafeber for scoring of the radiographs, and for thoughtful discussions. The authors gratefully acknowledge the use of MeVisLab (MeVis Medical Solutions AG, Bremen, Germany) for software development.²⁰

Contributors FPL, DvdH, KLV, MAV and YH initialised the conception and design of the study. All authors were involved in data acquisition. FPL, DvdH, KLV, MAV, YH and MJHDH were involved in data analyses. YH and MJHDH drafted the manuscript. All authors contributed in editing and reviewing the manuscript, and approved the final manuscript for submission.

Funding This work was financially supported by the Center for Translational Molecular Medicine (CTMM, grant no. 04I-202).

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

REFERENCES

1. Van der Heijde DM. Plain X-rays in rheumatoid arthritis: overview of scoring methods, their reliability and applicability. *Baillieres Clin Rheumatol* 1996;10:435–53.
2. Bruynesteyn K, Van Der Heijde D, Boers M, *et al.* The Sharp/van der Heijde method out-performed the Larsen/Scott method on the individual patient level in assessing radiographs in early rheumatoid arthritis. *J Clin Epidemiol* 2004;57:502–12.
3. Smolen JS, Aletaha D, Bijlsma JWJ, *et al.* Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.
4. Bakker MF, Jacobs JWG, Welsing PMJ, *et al.* Low-dose prednisone inclusion in a methotrexate-based, tight control strategy for early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2012;156:329–39.
5. Bakker MF, Jacobs JWG, Verstappen SMM, *et al.* Tight control in the treatment of rheumatoid arthritis: efficacy and feasibility. *Ann Rheum Dis* 2007;66(Suppl 3):iii56–60.
6. Breedveld FC, Weisman MH, Kavanaugh AF, *et al.* The PREMIER study: a multicenter, randomized, double-blind clinical trial of combination therapy with adalimumab plus methotrexate versus methotrexate alone or adalimumab alone in patients with early, aggressive rheumatoid arthritis who had not had previous methotrexate treatment. *Arthritis Rheum* 2006;54:26–37.

7. St Clair EW, van der Heijde DMFM, Smolen JS, *et al.* Combination of infliximab and methotrexate therapy for early rheumatoid arthritis: a randomized, controlled trial. *Arthritis Rheum* 2004;50:3432–43.
8. Lipsky PE, van der Heijde DM, St Clair EW, *et al.* Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-Tumor Necrosis Factor Trial in Rheumatoid Arthritis with Concomitant Therapy Study Group. *N Engl J Med* 2000;343:1594–602.
9. Huo Y, Vincken KL, Viergever MA, *et al.* Automatic joint detection in rheumatoid arthritis hand radiographs. *10th IEEE International Symposium on Biomedical Imaging (ISBI)*; 2013:125–8.
10. Sharp JT, Angwin J, Boers M, *et al.* Computer based methods for measurement of joint space width: update of an ongoing OMERACT project. *J Rheumatol* 2007;34:874–83.
11. Lukas C, Sharp JTJT, Angwin J, *et al.* Automated measurement of joint space width in small joints of patients with rheumatoid arthritis automated measurement of joint space width in small joints of patients with rheumatoid arthritis. *J Rheumatol* 2008;35:1288–93.
12. Hulsmans HMJ, Jacobs JWG, Van Der Heijde DMFM, *et al.* The course of radiologic damage during the first six years of rheumatoid arthritis. *Arthritis Rheum* 2000;43:1927–40.
13. Landewé RBM, Strand V, Conaghan PG, *et al.* Damage and progression on radiographs in individual joints: data from pivotal randomized controlled trials. *J Rheumatol* 2011;38:2018–22.
14. Van der Heijde D. How to read radiographs according to the Sharp/van der Heijde method. *J Rheumatol* 1999;26:743–5.
15. Van der Heijde DM. Assessment of radiographs in longitudinal observational studies. *J Rheumatol* 2004;31(Suppl 69):46–7.
16. Verstappen SMM, Jacobs JWG, van der Veen MJ, *et al.* Intensive treatment with methotrexate in early rheumatoid arthritis: aiming for remission. Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA, an open-label strategy trial). *Ann Rheum Dis* 2007;66:1443–9.
17. Arnett FC, Edworthy SM, Bloch DA, *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
18. Van der Heijde DM, Boonen A, Boers M, *et al.* Reading radiographs in chronological order, in pairs or as single films has important implications for the discriminative power of rheumatoid arthritis clinical trials. *Rheumatology (Oxford)* 1999;38:1213–20.
19. Brennan RL. *Generalizability theory*. New York: Springer-Verlag, 2001.
20. McInerney T, Terzopoulos D. Deformable models in medical image analysis: a survey. *Med Image Anal* 1996;1:91–108.