The use of Digital X-Ray Radiogrammetry in the assessment of Joint Damage in Rheumatoid Arthritis

Wajid B. Jawaid 1, David Crosbie 1, Julie Shotton 2, David M. Reid 1,2, Alison Stewart 2.

1. Department of Rheumatology, NHS Grampian, Ward 3, Aberdeen Royal Infirmary, Aberdeen, UK, AB25 2ZN
2. Department of Medicine and Therapeutics, University of Aberdeen, Aberdeen, UK

Key Words: rheumatoid arthritis, joint erosions, hand bone density, radiogrammetry

Correspondence to:
Prof David M Reid
Head of Department of Medicine & Therapeutics
University of Aberdeen,
Medical School,
Foresterhill,
Aberdeen, AB25 2ZD
E-mail: d.m.reid@abdn.ac.uk
Abstract

OBJECTIVE: To examine cross-sectionally the ability of digital x-ray radiogrammetry (DXR) to assess bone loss in RA compared with a manual radiograph scoring technique and the relationship of both these scores with other disease indices. METHODS: Consecutive consenting subjects attending the RA clinic were enrolled. Subjects were sent for x-ray; had their demographic details recorded; completed a self assessment questionnaire, which collected data on early morning stiffness, pain, patient global assessment and included a modified health assessment questionnaire (HAQ); had blood taken for erythrocyte sedimentation rate measurement and were assessed by author JC, a trained nurse, who recorded an assessor’s global assessment (AGA) and 28 point tender and swollen joint counts. All x-ray films were scored manually using the modified Sharp scoring technique by a single observer and 20 films were randomly selected for rescoring by 3 readers. All films were assessed using the Pronosco X-Posure System version 2.0. Analysis was performed in 225 subjects and included chi-squared tests, independent t-tests, multiple linear regression and partial correlations as appropriate. The precision of scoring techniques was estimated by calculating the smallest detectable difference (SDD), the coefficient of variation (CV) and the coefficient of repeatability (CR) from Bland & Altman plots.

RESULTS: The precision of DXR depending on the index assessed varied from SDD: 0.002 to 0.9; CV: 0.09% to 5.9%; CR: 0.002 to 0.792. These figures were better than that of the intra and inter-observer Sharp scores SDD = 73.9; CV = 27.8%; CR: 33.0 to 47.6). The DXR measurements that significantly predict Sharp scores are bone mineral density (BMD, $R^2 = 0.210$), metacarpal index (MCI, $R^2 = 0.222$) and cortical thickness (CT, $R^2 = 0.215$). The other DXR measurements, the derived porosity index and bone width, did not correlate to any aspect of the modified Sharp score. In females, DXR measurements significantly correlated with the modified HAQ scores but not with any of the other disease indices.
indices. In contrast, Sharp scores significantly correlate with AGA, swollen and tender joint counts, pain, HAQ and DAS28. CONCLUSION: In our study DXR measurements are much more precise than Sharp scores, but just like the latter are related to long-term disease activity in RA. DXR is simple to use, does not require intensive training and if used in a clinical setting may have the ability to quickly and cheaply identify subjects not responding to standard therapy or may allow selection of subjects who would benefit from more aggressive therapy with newer, more expensive treatments at an earlier stage of the disease.

Introduction

Many advances have been made in uncovering the pathogenesis of rheumatoid arthritis and this has led to the introduction of new therapies. Cytokines produced by macrophages, in particular TNFα have been found in higher concentrations in rheumatoid joints(1;2). This has led to development of therapies directed against TNFα and its receptors, which have shown to slow radiographic joint damage, but these are expensive and can have serious side effects(3-6). It would be useful if these and other disease modifying drugs could be specifically targeted at an early stage of disease to those subjects who are unresponsive to safer and cheaper drugs and are having ongoing joint damage.

Current techniques of assessing long-term disease progression in rheumatoid arthritis (RA), although useful in comparing drug treatments in clinically controlled trials, are not very effective in assessing RA in individual subjects. Accordingly they are very rarely used as part of standard day-to-day clinical practice by rheumatologists(7).

It is necessary to use various imaging techniques to assess structural damage. While there is interest in using MRI(8) and imaging ultrasound(9;10) to assess joint damage including
cartilage loss and erosion counts, the most widely used methods are a variety of plain film x-ray scoring techniques. The Sharp scoring method(11) is such a technique and in its modified forms such as that suggested by van Heijde(12), is commonly used in clinical trials to assess erosions and joint space narrowing for joints of both the hands and feet. Since erosive damage of the peri-articular bones is believed to be largely irreversible, the use of such an end-point clinically might be considered to have limited benefit with treatment always lagging behind disease progression. However, a measurement technique that showed a strong relationship with joint damage, which is in itself related to long-term functional status, might have real clinical value.

Peri-articular bone loss is the earliest radiological feature of RA. Using quantitative assessment of hand bone mass as a surrogate, bone loss can be seen to occur early in the disease and predate erosive damage(13;14). Peripheral bone mass can be quantified by a number of techniques in RA including dual energy X-ray absorptiometry (DXA)(13-15), quantitative ultrasound(9;10;16) and possibly MRI(8). DXA is most widely used for estimating in vivo total hand BMD but may also be used to assess BMD specifically at the metacarpals(14). Recently an improved method of radiogrammetry has been introduced using digitised plain hand radiographs(17) to measure bone density, metacarpal index, cortical thickness and porosity with high precision(18).

In this study we have examined the ability of digital x-ray radiogrammetry (DXR) to assess bone loss in RA, compared the results with a manual plain radiograph scoring technique - the modified Sharp score(19) – and examined the relationship of both methods to other indices of disease activity and progression in RA.
Methods

The study was carried out at the Osteoporosis Research Unit, Aberdeen. The Grampian Research Ethics Committee granted ethical approval.

Subjects attending the RA clinic were provided with information about the study, which included an information leaflet. Consecutive, consenting subjects were enrolled onto the study. Subjects also underwent an x-ray of their hands if such an x-ray had not been performed in the previous 2 years. JS carried out the data collection between May 2001 and July 2002.

Subjects’ demographic details were recorded and a self-assessment questionnaire was completed, which included data of the duration of early morning stiffness (EMS); pain, as assessed on a 10cm horizontal visual analogue scale (VAS) and a patient global assessment (PGA) of disease activity, as assessed on a 10 cm horizontal VAS. A modified health assessment questionnaire (HAQ) was also administered (20). Blood was also taken from subjects and the ESR measured in mm after a 1-hour sample frame. Subject case notes were audited and current medications and dosages, rheumatoid factor positivity, duration of disease and the presence or absence of erosions noted on the database. “Physician” assessments were also carried out by JS, a fully trained research nurse and included an assessor global assessment (AGA) of disease activity as assessed on a 10cm horizontal VAS, and a 28-point tender and/or swollen joint count. If appropriate the subjects were then asked to attend for a plain film x-ray to be taken of both hands. Only those subjects that had a relevant x-ray taken were assessed as part of this study.
Author WBJ scored all the available hand x-rays using the modified Sharp scoring method and the results were recorded. 20 randomly selected subjects were re-scored by WBJ a week after the initial assessment to allow calculation of short-term intra-observer variation. Inter-observer variation was calculated by examining 20 x-rays randomly selected for re-scoring by 3 observers (DC, DMR and WBJ).

All the hand x-rays were then assessed using the Pronosco X-Posure System version 2.0 (Sectra Medical Systems, Sweden) (18;21-24). This system provides the following measurements: DXR-BMD – (Bone mineral density); Porosity (POR); Metacarpal Index (MCI); Cortical Thickness (CT); Bone Width (BW). The Pronosco X-Posure System requires plain x-ray films. The x-rays films of the subjects in this study were of both hands on the same film. To evaluate x-rays using DXR the x-ray of each hand was separately scanned in and then analysed. BMD is calculated using formulae after assessment of the bone volume per area, assuming a cylindrical bone (18) thus negating the need for a phantom. Regions of interest are shown in Legends for Figures

Figure 1.

All data was entered into a spreadsheet and analysed using SPSS 11.0.1 (SPSS Inc, USA) and Microsoft Excel (Microsoft Corp, USA). Tests used included chi-squared tests, independent t-tests, multiple linear regression analysis (MLR) and partial correlation as appropriate. The precision was estimated by calculating the smallest detectable difference (SDD) and the standardised coefficient of variation (CV). The reproducibility of the repeated measurements was assessed using Bland & Altman plots and coefficients of repeatability were calculated. Correlation coefficients were compared using Fisher’s z transformation.
Results

Subject flow and comparison

The flow of subjects through the study is shown in Figure 2. 537 subjects were on the original database but only 225 met the entry requirements. Demographics and disease activity indices for the current study group were compared with the total study population in Table 1. We found the HAQ score was significantly lower (p=0.026) in the study group than in the whole population and disease duration was significantly shorter (p=0.023) but there were no other significant differences between the groups.

Precision

Sharp scores

The mean (+/-1SD) Sharp Scores for the study population was 79.38 (+/-53.45). The intra-observer variation calculations for the modified Sharp scoring techniques were CV = 7.2%. The intra-observer Bland & Altman plot (Figure 3) shows that variation at the extremes of the modified Sharp scores is smaller than the variation at intermediate Sharp scores. The coefficient of repeatability is 14.8.

For the inter-observer variation the CV was 27.8% and the SDD was calculated at 73.9 units for a mean sharp score of 94. The co-efficients of repeatability were as follows; between DC and DMR = 33.6 sharp score units, between WBJ and DMR = 33.0 units and between DC and WBJ = 47 units. An example of the inter-observer Bland and Altman plot is shown in Figure 4.

DXR measurements
The intra-observer precision values for the different DXR parameters scored on the same 20 radiographs assessed on two separate occasions 2 months apart are summarised in Table 2. Of the DXR measurements porosity appears to be the least precise measurement while bone width and cortical thickness are the most precise. The co-efficient of repeatability ranged from 0.002 to 0.792 for the DXR measurements.

**Correlations between Sharp scores and DXR measurements**

Table 3 shows the Partial correlations between Sharp scores and DXR measurements. In male subjects correlation between DXR-MCI and the modified Sharp score and its component scores appears to be weaker than that in female subjects although the differences are not significant (p > 0.05 for all coefficients). The derived porosity index and bone width did not relate to any aspect of the Sharp index.

**Correlation of radiological measures of disease severity with clinical disease activity measurements**

Table 4 show the correlation between DXR measurements, Sharp scores and clinical measures of disease activity. In female subjects for DXR measurements the only significant correlations seen were with HAQ scores (\(r=-0.218, P=0.008\)). In contrast Sharp scores significantly correlate with AGA, swollen and tender joint counts, VAS pain, HAQ and DAS28 but with low \(r\)-values of around 0.17 to 0.30 (\(P=0.043-0.01\)). Sharp scores correlated slightly better with the HAQ than did DXR (\(r=+0.298, P<0.005\)). In contrast the smaller number of male patients DXR measurements show better correlation with certain disease measures while the Sharp scores did not show significant correlation with any of the clinical measures.
Discussion

The precision results clearly indicate that DXR measurements are much more reproducible than Sharp scores which vary greatly not only between observers but also when re-scored by the same observer even a week later. The precision of DXR measurements in this study is better than that reported by Jorgensen et al. (18). This discrepancy is most likely to be due to the fact that we are simply measuring noise in the system, since the same radiograph is used each time. Perhaps a truer estimate would be given if 2 x-rays were taken at the same time point and used for calculation of precision. The precision of the modified Sharp score was lower than that reported by Sharp et al. (25), who found an inter-observer CV of 17%. However this may be due to the limited experience of two of the scorers or may be due to the long duration of disease making reading of the x-ray more difficult. However a recent comparison of intra- and inter-observer variation by Sharp et al. has shown figures for the SDD which are very similar to those we were able to achieve (26).

Bone loss is an early feature of RA and precedes the permanent and irreversible erosive changes that occur (13;14). The Pronosco device can measure bone loss at the metacarpals, which are in the region most commonly affected by RA and therefore may be able to identify individuals requiring a change in therapy before any irreversible joint damage occurs. It has been well established that BMD is reduced in RA subjects and the degree of loss is related to disease activity. Daragon et al found significantly reduced BMD of the whole hand using DEXA at 6 months in RA subjects compared to subjects with other rheumatic diseases (27).
This study has demonstrated that both DXR and Sharp score measurements are related more to disease severity than to current disease activity. This has also been shown previously (28). DXR measurements are as good as Sharp scores in predicting HAQ and may even be better in male subjects. Their predictive value may not be so good in females due to other confounding factors which influence the measurement of bone mass including menopausal state and pre-disease bone mass status. Of the various DXR parameters available, BMD seems numerically to correlate best with modified Sharp scores but the differences between the correlation coefficients and the other DXR parameters are not statistically significant.

At present rheumatologists do not routinely use all of the RA measures mentioned in this study in their clinical review of subjects. Less than 10% of rheumatologists use pain scales, fewer of them use functional status questionnaires and although joint counts are done few are recorded in medical records (7). Wolfe et al showed that currently clinicians base their decisions on treatment change mainly on pain and joint count, which are immediately obvious during the subject interview and physical examination (29).

In our study DXR measurements are related to long-term disease activity in RA and therefore serial measurements in individual subjects may have the ability to assess response to therapy and this may be possible over short enough time periods to improve patient care, due to the high precision of the technique. Indeed a recent study by Jensen et al (30) has demonstrated that during a two year period measurements of DXR BMD showed more rapid changes in those with active compared to inactive disease and was also able to distinguish those with erosive as compared to non-erosive RA. A recent pilot longitudinal study from our own centre in subjects with early RA has identified how a measurement of
the rate of change of DXR BMD in the first year of follow-up may be able to identify those who become erosive by 4 years of disease with high specificity and reasonable sensitivity (31). DXR is simple to use, does not require intensive training and if used in a clinical setting may have the ability to quickly and cheaply identify subjects not responding to standard therapy or may allow selection of subjects who would benefit from more aggressive therapy with newer, more expensive treatments at an earlier stage of disease. Such a hypothesis will require testing in a formal prospective clinical trial.

In summary we describe a method of assessing peripheral bone mass in RA that could have use in routine clinical care unlike the time consuming and rather imprecise measurement of joint space narrowing and erosion scores such as the Sharp score. Longitudinal studies will be required to determine the value of the technique in assessing suitability of subjects for expensive biological agents which can effectively limit joint destruction and bone loss.

Acknowledgments
We are grateful to the Arthritis Research Campaign (ARC) for supporting this work in the form of a studentship for WBJ and for infrastructure support to AS and DMR.

Competing Interests
The authors have no competing interest to declare

Ethics Approval
The Grampian Research Ethics Committee granted the study full ethical approval.
Reference List


(11) Sharp JT, Lidsky MD, Collins LC, Moreland J. Methods of scoring the progression of radiologic


(21) Bouxsein ML, Palermo L, Yeung C, Black DM. Digital x-ray radiogrammetry predicts hip, wrist and
vertebral fracture risk in elderly women: a prospective analysis from the study of osteoporotic fractures.
Osteoporosis International 13[5], 358-65. 2002 May.

(22) Malich A, Freesmeyer MG, Mentzel HJ, Sauner D, Boettcher J, Petrovitch A et al. Normative values of
bone parameters of children and adolescents using digital computer-assisted radiogrammetry (dxr).

(23) Shepherd JA, Cheng XG, Lu Y, Njeh C, Toschke J, Engelke K et al. Universal standardization of


(25) Sharp JT, Bluhm GB, Brook A, Brower AC, Corbett M, Decker JL et al. Reproducibility of multiple-
observer scoring of radiologic abnormalities in the hands and wrists of patients with rheumatoid

(26) Sharp JT, Wolfe F, Lassere M, Boers M, Van Der Heijde D, Larsen A et al. Variability of precision in
scoring radiographic abnormalities in rheumatoid arthritis by experienced readers. Journal of
Rheumatology 31[6], 1062-72. 2004 June.

densitometry and ultrasonography study of the hand bones of patients with rheumatoid arthritis of

mass and its association with radiographic joint damage and fractures in 50-70 year old female patients
with rheumatoid arthritis: cross sectional Oslo-Truro-Amsterdam (OSTRA) collaborative study. Ann

(29) Wolfe F, Zwillich SH. The long-term outcomes of rheumatoid arthritis: a 23-year prospective,
longitudinal study of total joint replacement and its predictors in 1,600 patients with rheumatoid

unclassified polyarthritis and early rheumatoid arthritis is better detected by digital x ray
radiogrammetry than dual x ray absorptiometry: relationship with disease activity and radiographic

(31) Stewart A, Mackenzie LM, Black AJ, Reid DM. Predicting erosive disease in rheumatoid arthritis: A
longitudinal study of changes in bone density using digital x-ray radiogrammetry: A Pilot Study.
Table 1: Comparison between demographic and disease activity indices for all subject’s on database and those that were selected for the study with numbers for each index shown in paraentheses.

<table>
<thead>
<tr>
<th>Percentage ± 95% CI</th>
<th>Total study population (n=537))</th>
<th>Sub-study population (n=325)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender ratio</td>
<td>76% ± 4% (537)</td>
<td>71% ± 6% (225)</td>
<td>0.160</td>
</tr>
<tr>
<td>RF % seropositive</td>
<td>79% ± 4% (482)</td>
<td>79% ± 6% (217)</td>
<td>0.842</td>
</tr>
<tr>
<td>Mean ± 95% CI</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at 01/01/2000 (years)</td>
<td>55.1 ± 0.96 (534)</td>
<td>54.5 ± 1.54 (225)</td>
<td>0.521</td>
</tr>
<tr>
<td>AGA</td>
<td>34.5 ± 1.9 (520)</td>
<td>32.7 ± 2.7 (220)</td>
<td>0.291</td>
</tr>
<tr>
<td>PGA</td>
<td>40.1 ± 2.1 (520)</td>
<td>39.9 ± 3.1 (220)</td>
<td>0.917</td>
</tr>
<tr>
<td>VAS pain</td>
<td>37.4 ± 2.2 (521)</td>
<td>36.0 ± 3.2 (221)</td>
<td>0.477</td>
</tr>
<tr>
<td># of swollen joints</td>
<td>5.2 ± 0.4 (514)</td>
<td>4.7 ± 0.6 (222)</td>
<td>0.118</td>
</tr>
<tr>
<td># of tender joints</td>
<td>3.4 ± 0.4 (514)</td>
<td>3.3 ± 0.6 (222)</td>
<td>0.640</td>
</tr>
<tr>
<td>EMS (minutes)</td>
<td>50.8 ± 5.7 (531)</td>
<td>50.2 ± 8.6 (165)</td>
<td>0.909</td>
</tr>
<tr>
<td>HAQ</td>
<td>1.33 ± 0.07 (519)</td>
<td>1.19 ± 0.10 (221)</td>
<td>0.026*</td>
</tr>
<tr>
<td>ESR (minutes)</td>
<td>26.7 ± 2.4 (340)</td>
<td>27.3 ± 3.4 (157)</td>
<td>0.262</td>
</tr>
<tr>
<td>DAS28</td>
<td>5.88 ± 0.40 (334)</td>
<td>5.61 ± 0.57 (156)</td>
<td>0.447</td>
</tr>
<tr>
<td>Disease Duration at 2001 (years)</td>
<td>10.5 ± 0.7 (519)</td>
<td>8.9 ± 1.1 (223)</td>
<td>0.023*</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>70.6 ± 1.3 (483)</td>
<td>71.7 ± 2.1 (210)</td>
<td>0.362</td>
</tr>
</tbody>
</table>
Table 2: Precision of DXR measurements

<table>
<thead>
<tr>
<th>CV</th>
<th>MEAN</th>
<th>POR</th>
<th>MCI</th>
<th>CT</th>
<th>BW</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.304%</td>
<td>0.480</td>
<td>0.869%</td>
<td>0.322%</td>
<td>0.412%</td>
<td>0.087%</td>
</tr>
</tbody>
</table>

CV – Coefficient of Variation
Table 3: Partial correlation analysis between Sharp score and its component scores and DXR measurements, where shaded area indicates non-significant results. (Age at x-ray and weight were used as controlling variables)

<table>
<thead>
<tr>
<th></th>
<th>SHARP SCORE</th>
<th>EROSION SCORE – Both Hands</th>
<th>JOINT SPACE SCORE – Both Hands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>BMD</td>
<td>-0.481</td>
<td>-0.394</td>
<td>-0.441</td>
</tr>
<tr>
<td></td>
<td>p&lt;.0005</td>
<td>p=.003</td>
<td>P&lt;.0005</td>
</tr>
<tr>
<td>MCI</td>
<td>-0.479</td>
<td>-0.278</td>
<td>-0.417</td>
</tr>
<tr>
<td></td>
<td>p&lt;.0005</td>
<td>p=.040</td>
<td>p&lt;.0005</td>
</tr>
<tr>
<td>CT</td>
<td>-0.478</td>
<td>-0.369</td>
<td>-0.433</td>
</tr>
<tr>
<td></td>
<td>p&lt;.0005</td>
<td>p=.006</td>
<td>p&lt;.0005</td>
</tr>
<tr>
<td>Por</td>
<td>+0.133</td>
<td>-0.155</td>
<td>+0.135</td>
</tr>
<tr>
<td></td>
<td>p=0.102</td>
<td>p=0.259</td>
<td>p=0.099</td>
</tr>
<tr>
<td>BW</td>
<td>+0.037</td>
<td>-0.180</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>p=0.654</td>
<td>p=0.189</td>
<td>0.771</td>
</tr>
</tbody>
</table>
### Table 4: Partial correlation analysis between structural measures and clinical measures, where shaded area indicates non-significant results.

Each cell shows correlation coefficient and p value, in that order. (Age at x-ray and weight were used as controlling variables).
Legends for Figures

Figure 1: Automatic selection of ROIs by Pronosco

Figure 2: Flow of subjects through the study

Figure 3: Bland & Altman plot for the modified Sharp score intra-observer variation as scored by WBJ

Figure 4: Bland & Altman plot for the modified Sharp score inter-observer variation as scored by DC and DMR
537 subjects in database

246 excluded because relevant x-ray had not been carried out

291 had required x-ray

34 folders were unavailable from x-ray department

257 x-ray folders were retrieved

9 x-ray folders did not contain the required x-ray film

248 required x-rays were found

231 were scored by modified Sharp scoring method

233 were scored by the Pronosco X-Posure system

225 x-ray films had both assessment techniques used.
AVERAGE of 1ST_OBSERVATION and 2ND_OBSERVATION

1ST_OBSERVATION - 2ND_OBSERVATION

Mean 1.7
-1.96 SD -13.1
+1.96 SD 16.5

group.bmj.com on October 22, 2017 - Published by group.bmj.com
Downloaded from http://ard.bmj.com/
AVERAGE of DC and DMR

Mean: -33.6
+1.96 SD: 0.0
-1.96 SD: -67.3

DC - DMR

0.0

AVERAGE of DC and DMR

20 40 60 80 100 120 140 160 180
The use of Digital X-Ray Radiogrammetry in the assessment of Joint Damage in Rheumatoid Arthritis

Wajid B Jawaid, David Crosbie, Julie Shotton, David M Reid and Alison Stewart

*Ann Rheum Dis* published online August 26, 2005

Updated information and services can be found at:

**Email alerting service**

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**

Articles on similar topics can be found in the following collections

- Clinical diagnostic tests (1282)
- Connective tissue disease (4253)
- Degenerative joint disease (4641)
- Immunology (including allergy) (5144)
- Musculoskeletal syndromes (4951)
- Pain (neurology) (883)
- Radiology (1113)
- Radiology (diagnostics) (750)
- Rheumatoid arthritis (3258)

**Notes**

To request permissions go to:
[http://group.bmj.com/group/rights-licensing/permissions](http://group.bmj.com/group/rights-licensing/permissions)

To order reprints go to:
[http://journals.bmj.com/cgi/reprintform](http://journals.bmj.com/cgi/reprintform)

To subscribe to BMJ go to:
[http://group.bmj.com/subscribe/](http://group.bmj.com/subscribe/)