Automated three-dimensional registration of high-resolution peripheral quantitative computed tomography data to quantify size and shape changes of arthritic bone erosions

Dominique Töpfer1, Bastian Gerner1, Stephanie Finzel2, Sebastian Kraus2, Oleg Museyko1, Georg Schett2 and Klaus Engelke1

Abstract

Objective. To monitor size and shape changes of bone erosions and changes in BMD in the vicinity of the erosion and in the periarticular trabecular compartment of patients with RA using high-resolution peripheral quantitative CT (HR-pQCT) imaging and to compare an automated three-dimensional (3D) image processing technique with manual measurements of erosion width and depth.

Methods. The shape of 40 bone erosions and composition of bone around the erosions were analysed in the MCP joints of 22 RA patients both manually and by semi-automated 3D image processing at two different time points. Periosteal segmentation was performed using volume growing and morphological operations. Image registration was applied for transfer of baseline segmentations to follow-up datasets.

Results. Eight erosions decreased in size, 6 increased and 28 remained stable. Increasing erosions were more spherical and smaller at baseline compared with decreasing or stable erosions. BMD in the vicinity of shrinking erosions increased, while it decreased next to expanding erosions. There was moderate agreement in the determination of erosion volume between semi-automated and manual measurements, but agreement was poor when assessing changes in volume over time.

Conclusion. Longitudinal changes in erosion size and shape and of BMD in the vicinity of an erosion can be measured. BMD changes are associated with progression and regression of erosions. However, the semi-automated and manual approaches did not classify longitudinal changes of erosion volume in the same way. Further research is necessary to define the nature of these differences.

Key words: high-resolution peripheral computed tomography, bone erosions, rheumatoid arthritis, image processing.

Introduction

RA is a chronic inflammatory disease that most often affects peripheral joints and results in periarticular bone loss. Bone erosions are caused by inflammation of the synovial membrane and are an early sign of RA. Thus they are an important parameter for monitoring disease activity and treatment efficacy [1]. The aim of
anti-rheumatic drug therapy is the reduction of inflammation to prevent the development or progression of bone destruction as early as possible. Most often, DMARDs such as MTX or TNF-α inhibitors are used as either monotherapy or combination therapy [2].

Several studies suggest that not only retardation but also repair occurs in RA, even though it is far less frequent. Erosion repair in individual patients was first observed from conventional radiographs showing recor- tication of erosions, filling of erosions with new bone and secondary OA with bone sclerosis and osteophyte formation [3, 4]. The principal occurrence of erosion regression was confirmed by the OMERACT Subcommittee on Healing of Erosions [5], although agreement with the presence of morphological features of bone repair was poor. Newer conventional radiograph-based studies have reported repair in 10.7% and 7.2% of patients, respectively [6, 7]. In contrast, in a 12-month treatment study with adalimumab (ADA), decreased erosion scores were reported in only 1.6% of sites assessed by CT and 1.8% of sites assessed by radiography [8]. They also reported repair of individual erosions using MRI in patients during ADA/MTX treatment [9]. Finally, our own study using high-resolution peripheral quantitative CT (HR-pQCT) indicated that TNF inhibitors may be more effective in promoting repair than MTX [10].

Currently the assessment of bone erosion is based largely on semi-quantitative scoring of radiographs [11]. However, these scores may not be sufficient for the assessment of erosion repair, since an erosion has to close completely to achieve a lower score [3]. Thus the 3D erosion structure has been analysed using US, MRI and CT imaging methods, although most of these studies also applied semi-quantitative scoring systems such as the Rheumatoid Arthritis MRI Score (RAMRIS) [12]. The use of quantitative image analysis techniques such as contouring an erosion slice by slice [8] or other automated image processing techniques [13] is still uncommon. US and MRI can be used to visualize inflammatory processes directly, whereas CT is superior in depicting morphological details of bony structures. In particular, HR-pQCT, with its high spatial resolution of ~130 μm, may be a promising technique for the early detection and quantification of bone erosions.

Earlier we established an automated 3D technique to segment and quantify erosion size and shape in HR-pQCT datasets of MCP joints [14]. Here, the method was extended to quantify longitudinal changes in erosion size and shape and of parameters characterizing the surrounding trabecular bone architecture.

Materials and methods

Patients

Twenty-two patients [3 males, 19 females, mean age 51.9 years (s.d. 14.8)] of the Rheumatology Outpatient Clinic of the University of Erlangen were included in the study. All patients were treated with DMARDs (n = 12) or TNF inhibitors (n = 10) and fulfilled the 2010 ACR/EULAR RA classification criteria [15]. The study was performed in accordance with the Declaration of Helsinki. Approval from the Ethics Committee of the Friedrich-Alexander-University of Erlangen-Nuremberg and national radiation safety agency (Bundesamt für Strahlenschutz) as well as written informed consent from each patient was obtained for the study.

HR-pQCT acquisition

HR-pQCT scans (XtremeCT, Scanco Medical AG, Brütisellen, Switzerland) were performed of the second through fourth MCP joints of the dominant hand with an isotropic voxel size of 82 μm³ and a nominal resolution of about ~130 μm³. A maximum number of 322 slices was acquired per dataset—80 distally and 242 proximally relative to the top of the third metacarpal head. Two routine measurements obtained in a time interval of 1.2 years (s.d. 0.6; minimum 0.5, maximum 2.9) were used. The hand was positioned in stretched posture and padded. The scan time was about 8 min [1].

Image analysis

As described earlier [16], erosions were defined as juxta-articular breaks within the cortical shell. Erosions were differentiated from physiological breaks indicating entry of blood vessels by the linear shape and occurrence on predilection sites. Pseudo-erosions, structures similar to cortical breaks presented by osteophytes, were excluded [16].

The 3D analysis technique has been described earlier [14]. In short, the periosteal segmentation is based on volume growing followed by morphological operations. After placing a seed point in each erosion, the level-set method is used for inflating a small spherical structure. Inflation stops at surrounding bone while the periosteal segmentation serves as the boundary at the cortical break. In erosions not completely bounded by sclerotic bone, the level-set segmentation may leak into bone marrow. Therefore, in a second step, voxels outside the erosion are removed by morphological operations. The border of the segmented erosion is dilated to obtain four sclerotic bone volumes of interest (VOIs) bordering each erosion (Fig. 1A). The user can always correct the segmentation process if necessary.

In addition, a trabecular VOI is obtained by fitting a half-sphere to the distal end of the metacarpal head. The resulting volume is peeled isotropically to eliminate the cortical shell. Trabecular VOIs are not obtained in the phalanges because of little anatomical coverage in the HR-pQCT scans.

Image registration

In baseline and follow-up datasets of a subject, each bone containing erosions was registered separately. First, the periosteal surface was segmented independently in both datasets. Then a rigid 3D registration was applied to the resulting binary VOIs. In a third step, the periosteal surface of the baseline dataset was transferred to the follow-up
dataset, which ensured that the virtual bone surface in the vicinity of a cortical break was as similar as possible in both datasets (Fig. 1B and C). Finally, erosions were segmented independently in both datasets. If the registration failed, the periosteal segmentation from the follow-up dataset was used without registration after correcting obvious differences manually. Erosions for which this was not possible were excluded.

(A) Two orthogonal multiplanar reformations of a 3D HR-pQCT dataset showing segmented periosteal surface, segmented erosion, surrounding sclerotic bone VOIs 1–4 and the trabecular VOI. (B) Segmented periosteal surface in the baseline dataset (left) and the same slice from the corresponding follow-up dataset with transferred periosteal surface of the baseline dataset (right). (C) Quality of 3D registration: follow-up dataset is shown with both the follow-up periosteal segmentation (green) and the transferred periosteal surface of the baseline dataset (red). HR-pQCT: high-resolution peripheral quantitative CT; 3D: three-dimensional; VOIs: volumes of interest.
Assessments

For each segmented erosion volume, surface area and sphericity were determined. Sphericity measures how closely a shape resembles a sphere [17]. Its value varies between 0 and 1, with 1 being the result for a perfect sphere. BMD was measured in the four sclerotic bone VOIs bordering an erosion.

In the trabecular VOI of the metacarpal head, BMD and texture parameters were determined. These characterize the trabecular bone architecture, but in contrast to standard histomorphometric parameters such as trabecular thickness (Tb.Th) or trabecular spacing (Tb.Sp) do not require segmentation of the individual trabeculae. Five different texture parameters were calculated directly from the grey value images: global and local inhomogeneity, local anisotropy, variogram slope and entropy. Their dependence on trabecular architecture was recently described [18]. In short, global inhomogeneity is the S.D. of the grey values, whereas local inhomogeneity describes the average of standard deviations calculated in a small neighbourhood around each voxel. Local anisotropy is similar to local inhomogeneity but measures the directedness of the bone structure. Variogram slope, also a measure of local variation, describes grey value differences of voxels that are separated by a certain distance. Finally, entropy refers to the Shannon entropy from information theory and measures the information content [19].

Erosion volume was also estimated from manual measurements of erosion depth and width used as a standard evaluation of HR-pQCT images [1]. The two distances are used to construct a half-ellipsoid, the volume of which is defined as erosion volume [14]. Manual measurements performed by S.F. or S.K. were available for all erosions. In the following, the two techniques will be denoted as automated 3D (Auto3D) and manual (MAN).

Analysis strategy

Erosions were stratified according to the percentage change in volume using the criterion of least significant change (LSC) [20]. LSC was calculated from the interreader precision error, which for the Auto3D analysis was 8.3% for erosions <10 mm³ and 6.0% for erosions >10 mm³, resulting in LSC values of 23% and 17%, respectively [14]. Erosions that did not change in volume by at least the LSC were considered stable. A preliminary analysis indicated higher precision errors for the MAN analysis [14]. However, interreader precision for the MAN method can probably be improved by further training, thus the LSC values above were also used for the MAN method. An analysis of variance (ANOVA) with post hoc Tukey’s (for equal variance) or Games-Howell test was used for comparing parameters measured at baseline and of changes over time between the subgroups of increasing, decreasing and stable erosions.

Cohen’s κ scores and Bland-Altman plots were used to compare the ability of Auto3D and MAN to separate subsets of erosions that either increased, decreased or did not change in volume. Pearson correlation coefficients were calculated for correlations between change in erosion volume and change in the other assessments.

Results

Longitudinal dynamics of erosion size

Forty erosions were assessed in 22 RA patients constituting group 1 (G1). Subgroup G2 contained 30 erosions that at baseline were <10 mm³. Baseline assessments are shown in Table 1. Based on the LSC, in G1 8 of the 40 erosions (20%) decreased in volume and 6 erosions (15%) increased. Twelve patients had only one erosion, while 10 patients had multiple erosions (supplementary Table S1, available at Rheumatology Online). There was no patient who had increasing and decreasing erosions, but combinations of decreasing/stable and increasing/stable were frequently found. In G2, six erosions (20%) decreased and five erosions (17%) increased in volume. The trabecular VOI could not be analysed in three metacarpals (containing four erosions) due to limited scan range.

At baseline (Table 1), erosions in G2 that increased in volume had a statistically significant larger sphericity, that is, they were rounder than erosions that either were stable or had decreased. Also increasing erosions were statistically significantly smaller at baseline and had a statistically significantly smaller surface area than stable erosions. At baseline, BMD in the sclerotic bone VOIs was numerically higher for increasing erosions than for erosions that were either stable or decreased in size. As shown in Table 1, in VOI2 and VOI3 some differences attained statistical significance. Finally, BMD in the immediate vicinity of an erosion (VOI1) was statistically significantly higher (P < 0.001; ANOVA and Tukey’s post hoc test) than further away from the erosion, and this gradient between sclerotic bone VOI1 and the trabecular VOI was numerically higher for erosions that increased in size.

Differences between measurements at the two time points are shown for all assessments in Table 2. For G2, absolute differences are also shown as box plots in Fig. 2. Although not statistically significant, there was a consistent pattern in the change of sclerotic bone BMD: for erosions decreasing in size, BMD increased numerically, whereas the opposite was observed for increasing erosions. Results of percentage changes and results for G1 showed similar patterns.

Fig. 3 shows correlations between the change in erosion volume and changes of all other parameters assessed for G2. Not surprisingly, there was a high correlation between Δvolume and Δsurface area (Δ = change in). Interestingly the results indicate an inverse relationship between Δvolume and Δsphericity, although between groups, differences in Δsphericity were not statistically significant (Table 2). An increase in volume correlated statistically significantly with a decrease in BMD in the sclerotic bone VOIs. Regression coefficients were low to moderate (r < 0.5).
Values are given as mean (S.D.). Values in the three subgroups of decreasing, stable and increasing erosions were compared by ANOVA. Significant difference (ANOVA and Tukey’s post hoc test) between increasing and stable erosions: *P < 0.05, **P < 0.001. Significant difference between increasing and decreasing erosions: *P < 0.05, **P < 0.01. ANOVA: analysis of variance; VOI: volume of interest.

**Table 2** Percentage and absolute differences between baseline and follow-up measurements

<table>
<thead>
<tr>
<th>Group 1</th>
<th>Sclerotic bone VOIs 1-4</th>
<th>Trabecular VOI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta )Volume</td>
<td>( \Delta )Surface area</td>
</tr>
<tr>
<td>Decrease</td>
<td>Abs: 1.3 (2.2)</td>
<td>−7.9 (6.8)</td>
</tr>
<tr>
<td>%</td>
<td>−39 (17)</td>
<td>−30 (17)</td>
</tr>
<tr>
<td>Stable</td>
<td>Abs: 0.07 (2.3)</td>
<td>−1.6 (6.0)</td>
</tr>
<tr>
<td>%</td>
<td>−5.2 (11)</td>
<td>−3.6 (12)</td>
</tr>
<tr>
<td>Increase</td>
<td>Abs: 1.8 (2.2)</td>
<td>8.1 (8.9)</td>
</tr>
<tr>
<td>%</td>
<td>101 (12)</td>
<td>77 (89)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2</th>
<th>Sclerotic bone VOIs 1-4</th>
<th>Trabecular VOI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \Delta )Volume</td>
<td>( \Delta )Surface area</td>
</tr>
<tr>
<td>Decrease</td>
<td>Abs: 1.9 (2.2)</td>
<td>−8 (7.9)</td>
</tr>
<tr>
<td>%</td>
<td>−44 (19)</td>
<td>−36 (15)</td>
</tr>
<tr>
<td>Stable</td>
<td>Abs: 0.2 (0.6)</td>
<td>−1.1 (3.8)</td>
</tr>
<tr>
<td>%</td>
<td>−6 (12)</td>
<td>−4 (12)</td>
</tr>
<tr>
<td>Increase</td>
<td>Abs: 1 (1.3)</td>
<td>5.4 (6.6)</td>
</tr>
<tr>
<td>%</td>
<td>117 (118)</td>
<td>88 (95)</td>
</tr>
</tbody>
</table>

Values are given as mean (S.D.). Values in the three subgroups of decreasing, stable and increasing erosions were compared by ANOVA. Significant difference (ANOVA and Tukey’s or Games–Howell post hoc test) between increasing and stable erosions: *P < 0.01. Significant difference between increasing and decreasing erosions: *P < 0.05, **P < 0.01. ANOVA: analysis of variance; \( \Delta \): change in; VOI: volume of interest.

**Comparison between automated and manual measurements**

Correlations of erosion volume between the MAN and Auto3D techniques were \( r = 0.58 \) at baseline and \( r = 0.49 \) at follow-up, similar to results published earlier [14]. Correlations of \( \Delta \)volume between the two techniques were \( r = 0.34 \) \( (P = 0.031) \) for absolute and \( r = 0.04 \) (NS) for relative changes in G1 and \( r = −0.05 \) (NS) for absolute and \( r = −0.01 \) (NS) for relative changes in G2. At baseline, correlation between volume measured by Auto3D and maximum erosion depth was moderate \( (r = 0.67) \), correlation of changes was poor (absolute changes \( r = 0.001 \); relative changes \( r = −0.08 \)). The Bland–Altman plot (Fig. 4) shows that for most erosions the two methods differed by only \( \pm 2 mm^3 \). However, 7 of the 10 erosions \( > 10 mm^3 \) had data points outside the central cluster. Cohen’s \( k \) for the ability to classify erosions into groups of decreasing, increasing or stable erosions was negative for all erosions \( (k = −0.14) \) and for erosions \( < 10 mm^3 \) \( (k = −0.05) \), indicating that the methods agreed less than...
would be expected just by chance [21]. Thus we found no effective agreement between the methods when assessing longitudinal changes.

**Bone structure assessment**

Changes in texture of the trabecular VOIs were compared separately for the second and third fingers. Fourth fingers were not assessed because of the limited scan range. Also, the second fingers of two patients were excluded because of limited scan coverage. At baseline there were no significant differences for any of the texture parameters when comparing decreasing, increasing and stable erosions. The same result was found for trabecular BMD. Also, correlations between absolute and relative changes in volume and texture parameters were low ($r < 0.3$) and not statistically significant.
Discussion

We applied a 3D image technique for the analysis of bone erosions in HR-pQCT images to quantify longitudinal changes of erosion volume, shape and BMD in the vicinity of the erosion. Specifically, a 3D registration was applied to obtain a consistent segmentation of the periosteal surface in the vicinity of the cortical break, which is highly important for accurate quantification of bone damage.

Correlations between Auto3D and MAN in measuring erosion volume were moderate ($r = 0.5-0.6$) but correlations for measuring changes in erosion volume were poor. The Bland–Altman plot (Fig. 4) shows that in G2 most erosions clustered in the centre around ±2 mm$^3$, however, as the maximum volume change was <6 mm$^3$, a difference of 2 mm$^3$ between the two methods still implies a considerable deviation. The $\kappa$ results reflect this discrepancy, showing that the methods did not agree in classifying erosions into groups of decreasing, stable and increasing erosions, despite the fact that the same LSC values were applied for the two methods. These LSC values were based on an interoperator precision analysis of Auto3D [14] and favoured the MAN technique, for which a preliminary analysis had shown higher precision errors. Nevertheless, even the LSC for Auto3D indicates a high variability of the measurements, which is a cause of the difficulty in classifying longitudinal changes and a potential reason why the majority of the erosions were classified as stable. Further effort is needed to more comprehensively compare Auto3D and MAN. For example, the impact of repositioning probably has a larger accuracy impact on MAN, in which distances are measured on multiplanar reformations. In contrast, Auto3D is a true 3D image processing technique, largely independent of the angle of flexion of the fingers over the joint.

With increasing volume, erosions often become less sphere-like, that is, sphericity decreases. Based on our limited data, the assumption of MAN that erosion...
volume can be approximated by a half-ellipsoid may therefore be less accurate for larger erosions. However, sphericity depends not only on erosion shape, but also on the roughness of the surface between the erosion and the surrounding bone. Changes in roughness may therefore impact on the relationship between volume and sphericity.

In this study, two routine measurements were analysed retrospectively from RA patients randomly selected from the Rheumatology Outpatient Clinic. The baseline measurement did not mark a time when patient care or treatment was changed, but no information was obtained on the length and type of medication prior to the measurements analysed here. Also, the number of subjects was relatively small and obviously not powered to obtain medication-specific results. Thus the results discussed herein may not be representative of the general RA population. The characterization of change of erosion volume over time is difficult because the variations of erosion volume and its changes are large. Tiny erosions may double or triple in size, thus the relative volume increase is large, but the absolute volume increase remains small when compared with volume changes of large erosions. For large erosions, the opposite is true. As a consequence, average values may be biased one way or the other and may largely depend on a very small number of erosions. Thus both absolute and relative changes are reported. Still, pooling results from erosions without considering size is problematic. In particular, increasing erosions were rather small at baseline (<2.4 mm³, with one exception). Therefore a subgroup (G2) of small erosions was analysed separately. The maximum erosion size of 10 mm³ for G2 was chosen because a recent publication [22] showed that erosions <10 mm³ are often missed by MRI. For these small erosions, HR-pQCT seems the imaging method of choice. Also, published precision errors for erosion volume that determined the LSC were reported for erosions >10 and <10 mm³ [14]. Finally, the limit of 10 mm³ coincides approximately with the mean of erosion volume measured for the study patients investigated here (Table 1).

In agreement with other studies, a decrease in erosion volume was observed in some patients, although none of the erosions disappeared completely. As pointed out by Matzelle et al. [23], remission is found more often than repair, because localized, subclinical inflammation may persist even when the disease is controlled. Our study is one of the first to quantify the change in erosion volume based on 3D segmentation techniques independently of an assumption of erosion shape. Another novel approach is the quantification of BMD in the sclerotic bone VOIs adjacent to the erosion. Interestingly, at baseline, BMD in these VOIs was higher for erosions increasing in size compared with stable or decreasing erosions. However, BMD increased over time at sites with shrinking bone erosions and decreased at sites with expanding erosions (Fig. 2 and Table 2). This change over time is also confirmed in Fig. 3. Moreover, the change in BMD became smaller with increasing distance from the erosion and was not significant for trabecular BMD. Apparently erosion growth or shrinkage is a local process. The same dependence was observed for sphericity: over time, sphericity decreased with expanding bone erosions and increased with shrinking erosions. As changes in sphericity were moderate, this finding may be explained by the above-mentioned dependence of the inner surface on roughness. The surface of a shrinking erosion may become smoother through apposition of sclerotic bone, while the roughness of a surface of an expanding erosion may increase through depletion. This is consistent with the findings of Matzelle et al. [23], who describe roughened, irregular bone surfaces in micro-CT images of mice after the application of an arthritogenic serum. In reality, a single parameter such as sphericity may be insufficient to fully characterize erosion shape.

A limitation of this study is the overall rather small number of erosions analysed, preventing further differentiation into subgroups with erosions of similar size and dynamics of erosion repair. The question, whether already sclerosed erosions further decrease in size, could not be systematically addressed. The small number of erosions did not allow for differentiation of results according to specific medication. Also, the erosion analysis was carried out by a single reader instead of using a consensus reading from several experts.

Finally, motion artefacts that frequently occur in HR-pQCT have an influence on precision. Engkelke et al. [24] reported that 25% of scans contained severe motion artefacts for a standard 3 min scan of the distal radius. Due to larger coverage, scan times for the MCP joints are about three times as long. Despite the fact that for many patients more than two scans were available and the ones with fewer artefacts were selected, 43% of the patients of the initial cohort had to be excluded due to severe
motion artefacts. In the future, a major reduction of movement artefacts can be expected due to a new recently introduced holder, providing better fixation of the limb, and the introduction of the XtremeCT II, reducing scan time [25].

In summary, these data show that semi-automated 3D registration allows assessment of the size and form of arthritic bone erosions. Regression of erosion volume and associated changes of the underlying bone can be documented. Semi-automated 3D registration permits definition of the changes in periarticular bone in a site-specific manner dependent on the respective distance from the bone erosion. Currently correlations of this technique with the manual assessment of erosion size are still limited, warranting more detailed analyses to further improve the assessment of erosion volume. The future direction for research is to determine which method is best capable of detecting longitudinal changes accurately.

Acknowledgements

Preliminary results were presented at EULAR 2013 [26] and at ACR 2013 [27]. The present work was performed in partial fulfilment of the requirements for obtaining the degree Dr rer. biol. hum. at the University of Erlangen-Nuremberg (D.T.).

Funding: This work was supported by the Bundesministerium für Bildung und Forschung (project ANCYLOSS).

Disclosure statement: The authors have declared no conflicts of interest.

Supplementary data

Supplementary data are available at Rheumatology Online.

References


