CT-like images in MRI improve specificity of erosion detection in patients with hand arthritis: a diagnostic accuracy study with CT as standard of reference

Sevtap Tugce Ulas, Katharina Ziegeler, Sophia-Theresa Richter, Sarah Ohrndorf, Denis Poddubnyy, Marcus R Makowski, Torsten Diekhoff

ABSTRACT

Objective To compare the diagnostic accuracy of susceptibility-weighted imaging (SWI), standard T1-weighted (T1w) images and high-resolution 3D-gradient echo sequences (volumetric interpolated breath-hold examination (VIBE)) for detection of erosions in patients with peripheral arthritis using CT as standard of reference.

Materials and methods A total of 36 patients were included in the study. All patients underwent CT and MRI, including SWI, VIBE and T1w sequences. Two trained readers scored all imaging datasets separately for erosions in a blinded fashion. Specificity, sensitivity and diagnostic accuracy of MRI sequences were calculated on a per-patient level.

Results CT was positive for erosion in 16 patients and 77 bones (Rheumatoid Arthritis MRI Score >0), T1w in 28 patients, VIBE in 25 patients and SWI in 17 patients. All MRI sequences performed with comparably high sensitivities (T1w 100%, VIBE 94% and SWI 94%), although SWI had the highest specificity of 90%, followed by VIBE (50%) and T1w (40%). Both T1w and VIBE produced significantly higher sum scores than CT (341 and 331 vs 148, p<0.0001), while the sum score for SWI did not differ from CT (119 vs 148; p=0.411).

Conclusion Specificity for erosion detection remains a challenge for MRI when conventional and high-resolution sequences are used but can be improved by direct bone depiction with SWI. Both T1w and VIBE tend to overestimate erosions, when CT is used as the standard of reference.

INTRODUCTION

In imaging, erosions are one of the most specific signs of inflammatory arthritis and usually indicate a chronically progressive course or occur in the context of inadequately adjusted therapy. The initial imaging examination in clinical routine is X-ray. However, especially in early disease, sensitivity of X-ray is low. MRI is a powerful tool for depicting active inflammation and structural damage, but its value has recently been questioned by rheumatological societies due to its low specificity for erosions and synovitis. Therefore, it has been degraded from a modality of first choice in early arthritis to a tool for therapy monitoring, for example, in the context of studies.

Erosion detection with standard MRI techniques such as T1-weighted (T1w) sequences relies on an indirect depiction of cortical bone (through high signal of the surrounding bone marrow and soft tissue fat), which can lead to overestimation of imaging findings when non-erosive bone marrow changes are present. This is also true for the common

Key messages

What is already known about this subject?
► Direct bone imaging can improve erosion detection by MRI.

What does this study add?
► Susceptibility-weighted imaging (SWI) improves the specificity of MRI for erosion detection and is superior to standard T1-weighted (T1w) imaging and 3D-gradient echo sequences.

How might this impact on clinical practice or future developments?
► Accurate detection of erosions is an emerging field in arthritis imaging, and standard MRI techniques such as T1w and 3D-gradient echo sequences tend to overestimate these bony changes.
► SWI provides additional information based on direct depiction of bone structures such as calcifications, which is limited with conventional MRI sequences.
high-resolution, 3D-gradient echo sequences (eg, volumetric interpolated breath-hold examination (VIBE) sequence), which have been successfully used to image the sacroiliac joints but still lack specificity in peripheral arthritis. Evaluation of new innovative MRI sequences is therefore of crucial importance to improve diagnostic accuracy in erosion detection. Recently, some approaches have been proposed to improve direct bone imaging, for example, by using artificial intelligence algorithms to synthesise CT-like images or by directly exploiting the magnetic properties of bone structure for their detection. The latter is achieved with susceptibility-weighted imaging (SWI), which was first applied in neurovascular imaging. Calcium as the main component of bone has diamagnetic properties, which leads to a negative phase shift and is demarked by SWI. A change in bone structure can thus be directly detected with SWI.

The aim of this study was to directly compare the diagnostic accuracy of T1w, VIBE and SWI MRI sequences for the detection of erosions in patients with hand arthritis. CT was used as standard of reference.

**MATERIALS AND METHODS**

**Subjects**

In this study, we prospectively included 36 consecutive patients presenting to our hospital’s rheumatology outpatient centre with proven or suspected inflammatory disease of the wrist and finger joints from October 2018 to 2019. Exclusion criteria were age under 18 years and contraindications to MRI (eg, pacemaker, cochlear implants or claustrophobia) or CT (eg, pregnancy). The final diagnosis was made by experienced rheumatologists of the local rheumatology department based on clinical, laboratory and imaging findings.

**Imaging procedure**

All patients underwent CT and MRI including T1w, VIBE and SWI sequences of the clinically more affected hand on...
the same day. The MRI examination was performed on a 1.5 T scanner (Siemens Magnetom Avanto, Siemens Healthcare) using a four-channel flex coil (Siemens Healthcare). MRI pulse sequences and parameters were as follows: T1w in coronal orientation (3 mm slice thickness, repetition time (TR) of 401 ms, echo time (TE) of 21 ms, 512×512 resolution matrix and 90° flip angle), coronal SWI (0.5 mm slice thickness, TR of 34 ms, TE of 15 ms, 384×252 resolution matrix and 15° flip angle), coronal contrast-enhanced VIBE (0.5 mm slice thickness, TR of 34 ms, TE of 15 ms, 384×252 resolution matrix and 15° flip angle). Contrast agent was administered at a dose adjusted to body weight (0.2 mL/kg gadolinium-DOTA (Dotarem) or 0.1 mL/kg gadolinium-BTDO3A (Gadovist)).

CT was performed on a 320-row single-source scanner (Canon Aquilion ONE Vision, Canon Medical Systems) using a sequential volume technique (135 and 80 kVp, 150 mAs) with 16 cm z-axis coverage without table movement and a rotation time of 0.275 s for acquisition of dual-energy datasets. For evaluation in the present study, a sharp bone kernel reconstruction from the 80 kVp source data and coronal reformation with 0.5 mm slice thickness were used.

Both MRI and CT were performed in prone position with the hand stretched over the head (superman position).

Image reading
All datasets were separately anonymised before reading. The SWI magnitude image was inverted to simulate the impression of a CT scan. Two well-trained readers

Figure 2 Imaging examples. (A) 51-year-old male patient with seronegative RA. No erosions are apparent in CT or in any of the three MRI sequences. (B) 52-year-old female patient with seronegative RA during therapy with methotrexate. An erosion at the head of metacarpal bone II and a pseudoerosion at the capitate bone are clearly visible in all modalities (arrowheads). However, T1 and VIBE show a false-positive detection of another erosion in the head of the metacarpale bone II (arrow). (C) 64-year-old male patient with seronegative RA treated with corticosteroids. VIBE and SWI show a cystic bone lesion in the lunate bone (arrow), which is misinterpreted as an erosion in T1w imaging. A small erosion is detected in the head of metacarpal bone V in CT and SWI (arrowhead), whereby T1w and VIBE overestimated this lesion. SWI, susceptibility-weighted imaging; T1w, T1-weighted; VIBE, volumetric interpolated breath-hold examination.
The total sum score was 148 for CT (mean 4.11, SD 9.02, range 0–38), 341 for T1w (mean 9.47, SD 11.35, range 0–46), 331 for VIBE (mean 9.19, SD 11.48, range 0–41) and 119 for SWI (mean 3.31, SD 5.98, range 0–25). Only SWI did not differ significantly from CT using the Mann-Whitney U test (p=0.90, p<0.05 for T1w and for VIBE). All modalities showed moderate agreement with CT (ICC VIBE 0.72, ICC SWI 0.71 and ICC T1w 0.66). Inclusion is presented in figure 1. The patients had a mean C reactive protein of 16.9 mg/L (SD 29.2) and mean duration of symptoms of 1.7 years (SD 3.7). The final clinical diagnosis was rheumatoid arthritis in 25 (six seropositive and 19 seronegative), psoriatic arthritis/peripheral spondyloarthritis in four patients, osteoarthritis (OA) in three patients, calcium pyrophosphate deposition disease (CPPD) in two patients, limited systemic sclerosis in one patient and chronic pain syndrome also in one patient.

**Image reading and statistical analysis**

CT was positive for erosions (RAMRIS >0) in 16 patients, T1w in 28 patients, VIBE in 25 patients and SWI in 17 patients. On the per-patient level, all three pulse sequences yielded high sensitivities: T1w 100%, VIBE 93.8% and SWI 93.8%. SWI had the highest specificity of 90%, followed by 50% for VIBE and 40% for T1w. The highest diagnostic accuracy was achieved with SWI (91.7%), followed by VIBE with a diagnostic accuracy of 69.4% and T1w with 65.7%. On the per-bone level, all 828 eligible bone regions were assessed by CT, 827 by T1w, 826 by VIBE and 823 by SWI. The results of contingency table analysis on the patient level are summarised in table 1. The results on the region level are summarised in online supplemental tables 1 and 2.

The Bland-Altman plots were presented for T1w, VIBE and SWI. The plots show that T1w and VIBE tended to overestimate the erosion extent compared with CT.

## RESULTS

### Subjects

A total of 36 patients (24 women; mean age 54, 23–75 years) were included in the study. A flowchart of patient inclusion is presented in figure 1. The patients had a mean C reactive protein of 16.9 mg/L (SD 29.2) and mean duration of symptoms of 1.7 years (SD 3.7). The final clinical diagnosis was rheumatoid arthritis in 25 (six seropositive and 19 seronegative), psoriatic arthritis/peripheral spondyloarthritis in four patients, osteoarthritis (OA) in three patients, calcium pyrophosphate deposition disease (CPPD) in two patients, limited systemic sclerosis in one patient and chronic pain syndrome also in one patient.

### Imaging procedure

The total dose-length product was 46.6 mGy·cm with an estimated effective dose of 0.075 mSv. All patients tolerated the examinations well and there were no artefacts degrading interpretation of MRI or CT. Imaging examples are presented in figure 2.

### Statistical analysis

Statistical analysis was performed using GraphPad Prism V.9.2.0. (for MacOS; GraphPad Software, La Jolla, California, USA). Scoring results were dichotomised into positive (RAMRIS >0) versus negative (RAMRIS=0) for erosions. Erosion detection by CT served as the standard of reference. Contingency tables were created separately for the wrist (including the metacarpal head and phalangeal bases II–V, were scored separately. The readers were blinded to clinical data and other imaging findings.

(KZ with 5 years and STU with 2 years of experience in musculoskeletal imaging) scored the MRI and CT datasets in consensus using the Rheumatoid Arthritis MRI Score (RAMRIS) criteria for erosion with scores of 0–10. According to the RAMRIS erosion scale, the metacarpal bases, carpal bones, distal radius and ulna, as well as the metacarpal head and phalangeal bases II–V, were scored separately. Sensitivity (including the metacarpal head and the phalangeal bases), specificity, diagnostic accuracy, positive and negative predictive value of T1w, VIBE and SWI were calculated on the patient level and region level using the Wilson/Brown method and the contingency table analysis. Mean sum scores were calculated to determine correlation of the different MRI pulse sequences. The Kolmogorov-Smirnov test was performed to test for normal distribution. Since the assumption of a normal distribution was hereby refuted, the Mann-Whitney-U test was conducted to test the differences between the different MRI sequences and CT. Intraclass correlation coefficient (ICC) analysis (CT vs the different MRI modalities) was performed using the sum scores. In addition, Bland-Altman plots were calculated to evaluate the agreement.

A p-value smaller than 0.05 was considered statistically significant.

### Table 1

Results of contingency table analysis on a per-patient level

<table>
<thead>
<tr>
<th></th>
<th>CT+</th>
<th>CT−</th>
<th>Total</th>
<th>SE</th>
<th>DA</th>
<th>SP</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1w/CT</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1w+</td>
<td>16</td>
<td>12</td>
<td>28</td>
<td>0.40</td>
<td>0.66</td>
<td>0.48</td>
<td>0.19</td>
<td>0.64</td>
</tr>
<tr>
<td>T1w−</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>0.66</td>
<td>0.66</td>
<td>0.66</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>20</td>
<td>36</td>
<td>0.56</td>
<td>0.91</td>
<td>0.91</td>
<td>0.47</td>
<td>0.64</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CT+</th>
<th>CT−</th>
<th>Total</th>
<th>SE</th>
<th>DA</th>
<th>SP</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIBE/CT</td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
<td></td>
<td></td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>VIBE+</td>
<td>15</td>
<td>10</td>
<td>25</td>
<td>0.50</td>
<td>0.69</td>
<td>0.69</td>
<td>0.27</td>
<td>0.73</td>
</tr>
<tr>
<td>VIBE−</td>
<td>1</td>
<td>11</td>
<td>12</td>
<td>0.69</td>
<td>0.69</td>
<td>0.69</td>
<td>0.52</td>
<td>0.84</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>20</td>
<td>36</td>
<td>0.60</td>
<td>0.91</td>
<td>0.91</td>
<td>0.49</td>
<td>0.70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CT+</th>
<th>CT−</th>
<th>Total</th>
<th>SE</th>
<th>DA</th>
<th>SP</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWI/CT</td>
<td></td>
<td></td>
<td></td>
<td>0.94</td>
<td></td>
<td></td>
<td>0.70</td>
<td>1.00</td>
</tr>
<tr>
<td>SWI+</td>
<td>15</td>
<td>2</td>
<td>17</td>
<td>0.90</td>
<td>0.92</td>
<td>0.92</td>
<td>0.68</td>
<td>0.99</td>
</tr>
<tr>
<td>SWI−</td>
<td>1</td>
<td>18</td>
<td>19</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.78</td>
<td>0.98</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>20</td>
<td>36</td>
<td>0.88</td>
<td>0.95</td>
<td>0.95</td>
<td>0.67</td>
<td>0.97</td>
</tr>
</tbody>
</table>

DA, diagnostic accuracy; NPV, negative predictive value; PPV, positive predictive value; SE, sensitivity; SP, specificity; SWI, susceptibility-weighted imaging; T1w, T1-weighted imaging sequence; VIBE, volumetric interpolated breath-hold examination.
In this study, we investigated the diagnostic potential of SWI and VIBE in the detection of erosions in patients with suspected inflammatory arthritis of the hand. CT was used as standard of reference. SWI performed with very high sensitivity (94%) and specificity (90%), whereas VIBE and T1w had much poorer specificities (50% and 40%). Furthermore, SWI had the highest diagnostic accuracy—92% on the per-patient level and 94% on the per-bone level. While almost excellent correlations of the sum scores with CT were shown for all three MRI sequences, only SWI showed no statistical difference from CT, while VIBE and T1w overrated the amount of erosions.

The early and accurate detection of erosion is crucial, as it has a considerable impact on differential diagnostic considerations and therapeutic decision making. Moreover, it has prognostic implications and progression may ultimately lead to a loss of joint function. Our findings suggest that SWI has higher specificity for erosions than VIBE, and the latter improves erosion depiction compared with standard thick-sliced T1w imaging. SWI may also identify additional abnormalities such as calcifications and may improve osteophyte detection, which is of great importance in the differentiation of arthritis from other conditions such as CPPD or OA. Overall, SWI has great potential to improve the specificity of MRI and its ability to identify differential diagnostics, where MRI is currently limited and which is the main reason why MRI is not recommended as a first-line imaging method in early arthritis.

Initial results indicate that direct bone imaging can also be improved by synthetic CT images reconstructed from MRI datasets using various new deep learning algorithms. Another promising method is offered by new pulse sequences with ultrashort TEs or zero ZTE, which allow direct visualisation of bone by more precise T2* effects, our results suggest that it is less well suited for direct bone depiction than previously reported.26

Our study has some limitations. SWI parameters were successfully optimised to improve spatial resolution, which comes at the cost of a slightly longer examination time and a higher susceptibility to motion artefacts. Further technical advances and tuning of sequence parameters are needed to shorten examination time. Despite optimisation of scan parameters and careful placing of patients and coils in the scanner, MRI could not detect all regions in all patients due to joint deformity or pain-related difficulties in positioning of the hand. The number of swollen joints was not assessed in this study design. Furthermore, the modus of reading did not allow for calculation of inter-rater reliability. Moreover, we did not compare the MRI techniques investigated here with arthrosonography, which is the current clinical reference standard.

In conclusion, SWI improves the specificity for erosion detection and is thereby superior to standard T1w and 3D-gradient echo (VIBE) sequences. Therefore, it has the potential to increase the diagnostic accuracy and differential diagnostic power of MRI compared with arthrosonography. Our results underline the importance of direct bone depiction and its potential for improving in arthritis imaging.

Author affiliations
1Department of Radiology, Charité Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany
2Department of Gastroenterology, Infectiology and Rheumatology, Charité Universitätsmedizin Berlin, Campus Mitte, Berlin, Germany
3Department of Gastroenterology, Infectiology and Rheumatology, Charité Universitätsmedizin Berlin, Campus Benjamin Franklin, Berlin, Germany
4Department of Radiology, Klinikum rechts der Isar der Technischen Universität München, München, Germany

Acknowledgements The authors thank Ms Bettina Herwig for language editing. The authors thank Ms Bettina Herwig for language editing.

Contributors STU: design of scoring system, image scoring, data evaluation, statistical calculations, article draft and critical revision of the manuscript for important intellectual content. KZ: conception and design of the study, imaging scoring, data evaluation, statistical calculations, critical revision of the manuscript for important intellectual content. S-TR: patient acquisition, data collection and critical revision of the manuscript for important intellectual content. S0: patient acquisition, image scoring and critical revision of the manuscript for important intellectual content. DP: patient acquisition and critical revision of the manuscript for important intellectual content. MRM: conception and design of the study and critical revision of the manuscript for important intellectual content. TD: conception and design of the study and critical revision of the manuscript for important intellectual content.
and design of the study, design of scoring system, image scoring, data evaluation and critical revision of the manuscript for important intellectual content, acts as guarantor.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants and was approved by the local ethics committee (EA4_005_18) and the Federal Office for Radiation Protection Z5-22482/2-2019-039. Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Results of the image scoring.

**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, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Svetlapugie Ulas http://orcid.org/0000-0003-3871-3746
Katharina Ziegeler http://orcid.org/0000-0002-9763-2420
Sarah Ohndorf http://orcid.org/0000-0001-5943-4688
Denis Podubnyy http://orcid.org/0000-0002-4537-6015
Torsten Diekhoff http://orcid.org/0000-0003-3593-1449

**REFERENCES**


