

## **MRI-based synthetic CT in the detection of knee osteoarthritis Comparison with CT**

Arbabi, Saeed; Foppen, Wouter; Gielis, Willem Paul; van Stralen, Marijn; Jansen, Mylène; Arbabi, Vahid; de Jong, Pim A.; Weinans, Harrie; Seevinck, Peter

**DOI**

[10.1002/jor.25557](https://doi.org/10.1002/jor.25557)

**Publication date**

2023

**Document Version**

Final published version

**Published in**

Journal of Orthopaedic Research

**Citation (APA)**

Arbabi, S., Foppen, W., Gielis, W. P., van Stralen, M., Jansen, M., Arbabi, V., de Jong, P. A., Weinans, H., & Seevinck, P. (2023). MRI-based synthetic CT in the detection of knee osteoarthritis: Comparison with CT. *Journal of Orthopaedic Research*, 41(11), 2530-2539. <https://doi.org/10.1002/jor.25557>

**Important note**

To cite this publication, please use the final published version (if applicable).  
Please check the document version above.

**Copyright**




Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights.  
We will remove access to the work immediately and investigate your claim.

## RESEARCH ARTICLE

# MRI-based synthetic CT in the detection of knee osteoarthritis: Comparison with CT

Saeed Arbabi<sup>1,2</sup>  | Wouter Foppen<sup>3</sup>  | Willem Paul Gielis<sup>2,3</sup> |  
Marijn van Stralen<sup>4</sup> | Mylène Jansen<sup>5</sup>  | Vahid Arbabi<sup>2,6</sup> | Pim A. de Jong<sup>3</sup> |  
Harrie Weinans<sup>2,7</sup> | Peter Seevinck<sup>1,4</sup>

<sup>1</sup>Image Sciences Institute, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>2</sup>Department of Orthopedics, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>3</sup>Department of Radiology, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>4</sup>MRIGuidance B.V., Utrecht, The Netherlands

<sup>5</sup>Rheumatology & Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands

<sup>6</sup>Department of Mechanical Engineering, Faculty of Engineering, Orthopaedic-Biomechanics Research Group, Birjand, Iran

<sup>7</sup>Department of Biomechanical Engineering, Delft University of Technology (TU Delft), Delft, The Netherlands

## Correspondence

Saeed Arbabi, Image Sciences Institute and Department of Orthopedics, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, Utrecht 3508 GA, The Netherlands.

Email: [s.arbabi@umcutrecht.nl](mailto:s.arbabi@umcutrecht.nl)

## Funding information

Innovative Medicines Initiative Joint Undertaking, Grant/Award Number: 115770; H2020 Marie Skłodowska-Curie Actions, Grant/Award Number: 801540

## Abstract

Magnetic resonance Imaging is the gold standard for assessment of soft tissues; however, X-ray-based techniques are required for evaluating bone-related pathologies. This study evaluated the performance of synthetic computed tomography (sCT), a novel MRI-based bone visualization technique, compared with CT, for the scoring of knee osteoarthritis. sCT images were generated from the 3T T1-weighted gradient-echo MR images using a trained machine learning algorithm. Two readers scored the severity of osteoarthritis in tibiofemoral and patellofemoral joints according to OACT, which enables the evaluation of osteoarthritis, from its characteristics of joint space narrowing, osteophytes, cysts and sclerosis in CT (and sCT) images. Cohen's  $\kappa$  was used to assess the interreader agreement for each modality, and intermodality agreement of CT- and sCT-based scores for each reader. We also compared the confidence level of readers for grading CT and sCT images using confidence scores collected during grading. Inter-reader agreement for tibiofemoral and patellofemoral joints were almost-perfect for both modalities ( $\kappa = 0.83$ – $0.88$ ). The intermodality agreement of osteoarthritis scores between CT and sCT was substantial to almost-perfect for tibiofemoral ( $\kappa = 0.63$  and  $0.84$  for the two readers) and patellofemoral joints ( $\kappa = 0.78$  and  $0.81$  for the two readers). The analysis of diagnosis confidence scores showed comparable visual quality of the two modalities, where both are showing acceptable confidence levels for scoring OA. In conclusion, in this single-center study, sCT and CT were comparable for the scoring of knee OA.

## KEYWORDS

CT, MRI, neural networks, osteoarthritis, synthetic CT

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Journal of Orthopaedic Research*® published by Wiley Periodicals LLC on behalf of Orthopaedic Research Society.

## 1 | INTRODUCTION

Medical imaging modalities have helped improve our understanding of osteoarthritis (OA) from traditionally defined as a passive cartilage degeneration disorder to be nowadays characterized as an active whole-joint disease.<sup>1,2</sup>

Recently published clinical guidelines by the European League Against Rheumatism (EULAR) and American College of Radiology (ACR) support the use of 3D imaging to confirm the diagnosis of knee OA.<sup>3,4</sup> This includes modalities such as magnetic resonance Imaging (MRI) for soft tissue assessment and computed tomography (CT) for bone assessment as the second-line modalities. In addition to clinical applications, imaging plays a pivotal role in OA research of the knee, as Osteoarthritis Research Society International (OARSI) guidelines recommend the use of MRI and CT in OA clinical trials.<sup>5</sup>

MRI can provide 3D images with superior soft tissue contrast that enables direct evaluation of cartilage,<sup>6</sup> yet the indiscriminate visualization of hard tissues such as bones on common MR images complicates the diagnosis of bone-related pathologies in knee OA.<sup>7</sup> CT, in comparison, can provide 3D images with a radiodensity contrast that highlights the bony tissues. This is useful in OA, though it exposes the patient to ionizing radiation,<sup>8,9</sup> and causes discomfort to patients by increasing patient burden due to the need for multiple imaging examinations and potentially increased costs.<sup>9,10</sup>

To produce a radiation-free solution that can give clear bone visualization, MR Sequences like Ultrashort TE (UTE) and zero-echo time (ZTE) have been developed to acquire CT-like contrast.<sup>9,10</sup> However, these techniques are not generally available and often generate false-positive bone identification at water-fat interfaces.<sup>7</sup>

MR-based synthetic CT (sCT) is an alternative solution for generating quantitative CT-like contrast from MR images, that often uses a generally available clinical T1-weighted MRI sequence.

In terms of OA, CT can complement MRI by accurate depiction of bone deformities and remodelings, including osteophyte generation as well as subchondral bone lesions, facilitating quantification of morphological abnormalities in three dimensions.<sup>11,12</sup> Employing sCT, leads to an MR-only one-stop-shop workflow that can produce CT-like bone visualization images that are intrinsically registered to MR, without exposing the patient to ionizing radiations. This additional information may provide improved comprehension of pathophysiology of knee OA.<sup>13</sup> Prior studies have investigated the use of sCT for radiotherapy planning and positron emission tomography-MRI attenuation correction.<sup>14</sup> More recently, studies have shown promising results in the feasibility of sCT for lower arm orthopedic applications,<sup>15</sup> performance of sCT in the depiction of erosion, sclerosis, and ankyloses of sacroiliac (SI) joints,<sup>16</sup> its capability for spine surgical planning<sup>17</sup> and morphological assessment of the hip joint.<sup>7</sup>

The aim of this study was to evaluate the accuracy of sCT as compared with CT, for detection of knee OA, using a recently published OA scoring method for CT images (OACT).<sup>8</sup> We also compared the diagnostic confidence of the readers while scoring knee OA between sCT and CT using collected confidence scores.

## 2 | METHODS

### 2.1 | Data collection

Data for this study is selected from the IMI-APPROACH,<sup>18</sup> a cohort that recruited knee OA patients fulfilling ACR criteria.<sup>19</sup> Among patients included in IMI-APPROACH from UMC Utrecht ( $n = 153$ ), patients who did not have T1w GRE sequence or had varied repetition time ( $TR \neq 7$  ms) were excluded. This resulted in 69 patients that were used for this study. The study was approved by the Ethics Committee of the UMC Utrecht (METC 17440) and registered under [clinicaltrials.gov](https://clinicaltrials.gov) number NCT03883568. All patients provided written informed consent.

MR images were acquired using a 3T scanner (Achieva; Philips Healthcare) with an RF-spoiled T1weighted 3D multiecho gradient echo (GRE) sequence. Acquisition parameters included echo times of 1.92 and 3.5 ms for echo 1 and echo 2, a repetition time of 7 ms, a total acquisition time of 3 min 21 s, and a flip angle of 10 degrees. Field of view (FOV) was  $164 \times 164 \times 160$  with a pixel bandwidth of 1072 Hz. Images were reconstructed by the scanner at a resolution of  $0.46 \text{ mm} \times 0.46 \text{ mm} \times 0.5 \text{ mm}$ , in a  $352 \times 352 \times 322$  matrix. High-resolution CT scans of the knee (IQon - Spectral CT; Philips Healthcare) were reconstructed with at a resolution of  $0.37 \text{ mm} \times 0.37 \text{ mm} \times 0.34 \text{ mm}$ , using a tube voltage of 120 kVp and X-ray tube current of 115 mA.

### 2.2 | sCT generation

Synthetic CT was generated using a patch-based neural network inspired by the UNet presented by Ronneberger et al.<sup>20</sup> and applied for image synthesis in a way similar to Florkow et al.<sup>21</sup> The neural network was provided as inputs with 3D patches of size  $24 \times 24 \times 24$  with two channels, consisting of echo 1 and echo 2 of the RF-spoiled T1w GRE sequence. The details of architecture and training are described in comprehensive details by Florkow et al.<sup>21</sup> Specifically, the number of convolutional filters at the first depth level started at 32, and was subsequently doubled in four levels to reach the bottleneck of network. The model was trained for 100,000 iterations using the Adam optimizer ( $lr = 0.001$ ), L1 loss function, and a batch size of 32. The two echoes of the RF-spoiled T1w GRE were chosen with echo times such that water and fat signals would be interfering constructively and destructively. This provided important information about the presence of water and fat content, which in turn strongly influenced the CT value in a real CT scan. The benefit of this approach was assessed in more detail by Florkow et al.<sup>21</sup> Pairs of registered MR and CT images of 39 of patients were used for training. CT images were registered to MRI using an in-house implementation of Quaternion Interpolated Registration (QIR),<sup>22</sup> preserving the rigidity of the bones, while applying b-spline-based deformation of the soft tissue. The development of the synthetic CT generation algorithm was done in collaboration with MRI guidance BV.

sCT images were generated fully automatically from the first two echoes of MR images of 30 unseen test patients using the trained algorithm. The generated sCT images had the same resolution, orientation, and matrix size as their corresponding MR images.

### 2.3 | Image analysis

CT and generated sCT images of test patients were anonymized and provided to two specialists for OA scoring. Scoring of OA characteristics in CT and sCT images was performed using the recently published OsteoArthritis Computed Tomography (OACT) scoring system.<sup>8</sup> For the knee, this scoring system proposes tibiofemoral and patellofemoral joints to be scored separately. OACT grades both joints based on the OA characteristics, which include joint space narrowing, osteophytosis, subchondral cysts, and subchondral bone sclerosis. Joint space narrowing was scored between 0 for no joint space narrowing and 3 for obliterated joint space, osteophytes were scored between 0 for no osteophytes and 3 for clear large osteophytes and cysts were scored 0 for absence or 1 for presence of subchondral cyst. The joint score was calculated based on combination of characteristic scores. The patellofemoral joint was scored between 0 for no osteophytes or sclerosis to 3 for large osteophytes, (near) bony contact, and definite sclerosis. The definitions of the four diagnostic confidence levels are as follows<sup>16</sup>: confidence level 1 describes the image as useless for scoring this characteristic, confidence level 2 shows low confidence that may affect the scoring, confidence level 3 represents medium confidence that does not affect the scoring, and level 4 narrates high confidence facilitating a clear scoring.

Each image was scored by an MD researcher with 5 years of experience in the medical imaging of OA (WPG as reader (1)) and a musculoskeletal radiologist (WF as reader (2)). The CT and sCT images were anonymized and provided in random order to readers for the scoring.

### 2.4 | Statistical analysis

The OA characteristics and joint scores were used to calculate the intermodality agreement and interreader agreement. Agreements were assessed using the square-weighted  $K$  statistics for ordinal grades.<sup>23,24</sup> According to prospective power analysis detailed in Supporting Information: material, scoring of 60 paired CT-sCT samples will achieve 80% power with two-sided 5% significance level in discriminating an effect size of 0.3 in Cohen's  $K$  agreement.<sup>25,26</sup> Interpretation of  $\kappa$  values were made according to Landis and Koch<sup>27</sup>: that is, 0–0.20 slight agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 substantial agreement; 0.81–1 almost-perfect agreement. Analyses were carried out using Python version 3.7 and the Scikit-learn library version 0.24.2,<sup>28</sup> R-package KappaSize,<sup>26,29</sup> and IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp. The confidence scores were also used to compare the diagnostic confidence of readers while scoring each of the two modalities (CT or sCT). We have reported the 95% intervals of the confidence scores for CT and sCT by each reader for the purpose of visual confidence analysis. RMSE will tell us the average magnitude of difference between the confidence scores of corresponding CT and sCT images for each reader.

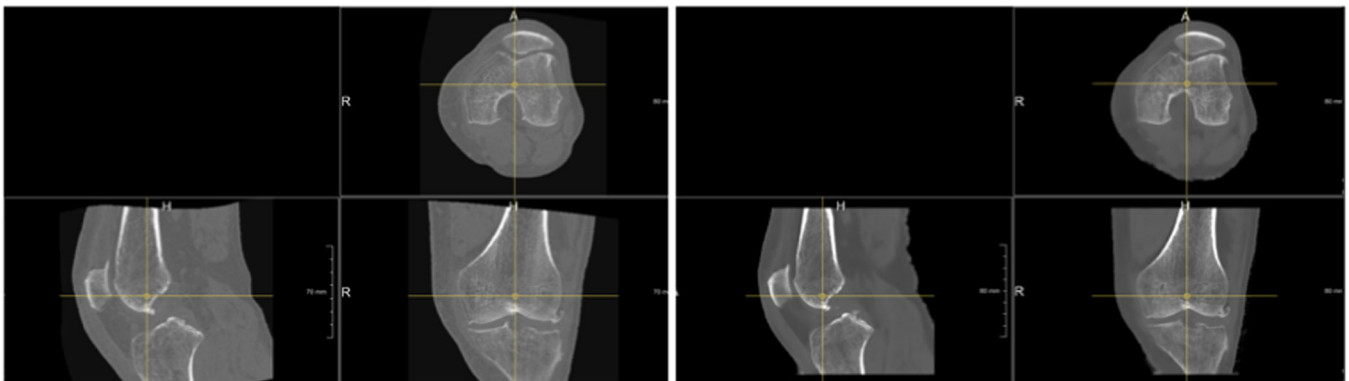
## 3 | RESULTS

sCT images for test patients were generated successfully (Figure 1).

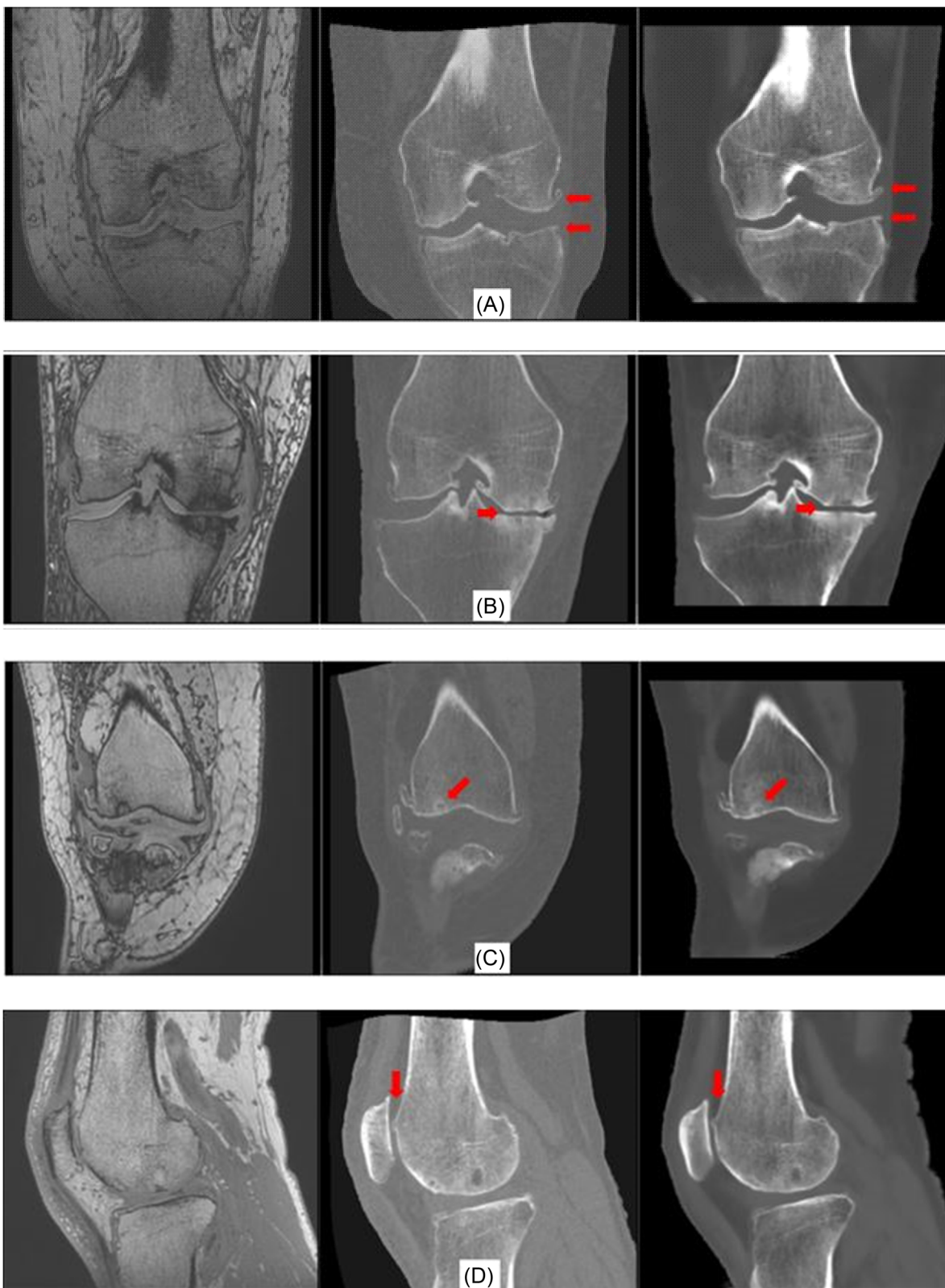
### 3.1 | Scoring results

CT and sCT samples were scored with OACT and confidence scores were attributed. Figure 2 visualizes a gallery of four OA characteristics on corresponding MR, CT, and sCT images illustrated side by side.

The prevalence of each pathology in the data set over all readers, modalities, and compartments are shown in Table 1.



**FIGURE 1** Computed tomography (left) and the generated synthetic computed tomography (right).

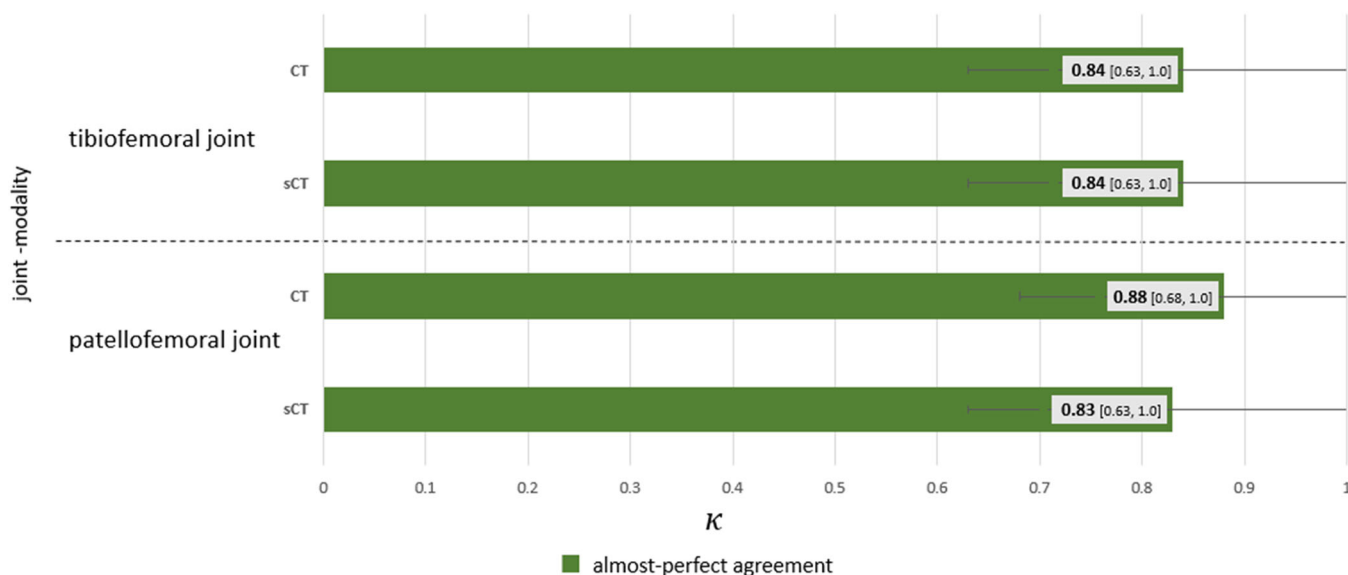


**FIGURE 2** left: MR 1st echo, middle: CT, right: sCT (A) patient with osteophyte grade 2 in the lateral tibiofemoral joint, (B) patient with joint space narrowing grade 3 in the medial tibiofemoral joint, (C) patient with joint cyst in medial tibiofemoral joint, and (D) patient with OA in patellofemoral joint scored as 2. Red arrows point out the pathologies. CT, computed tomography; OA, osteoarthritis; sCT, synthetic computed tomography.

**TABLE 1** Frequency of different grades for pathologies found on CT images in each joint and OA characteristic.

Joint space narrowing	Osteophyte	Cyst	Tibiofemoral joint	Patellofemoral joint
Level 0: 18%	Level 0: 20%	Level 0: 63%	Level 0: 32%	Level 0: 27%
Level 1: 30%	Level 1: 28%	Level 1: 37%	Level 1: 28%	Level 1: 28%
Level 2: 38%	Level 2: 20%		Level 2: 25%	Level 2: 28%
Level 3: 14%	Level 3: 32%		Level 3: 15%	Level 3: 17%

Abbreviations: CT, computed tomography; OA, osteoarthritis.

**FIGURE 3** Interreader agreement for scoring whole-joint OA on CT or sCT images in terms of  $\kappa$  score with 95% confidence interval bounds shown in brackets and error bars. CT, computed tomography; OA, osteoarthritis; sCT, synthetic computed tomography.

### 3.2 | Interreader agreement

The interreader agreement between readers scoring CT or sCT images of the whole joint was almost-perfect for both the tibiofemoral joint ( $\kappa = 0.84$  for both CT and sCT) and patellofemoral joint ( $\kappa = 0.88$  for CT and  $0.83$  for sCT) as detailed in Figure 3.

Collective interreader agreement irrespective of modalities was  $0.85$  [ $0.69, 1.0$ ] for tibiofemoral joint and  $0.85$  [ $0.7, 1.0$ ] for patellofemoral joint.

Interreader reliability for scoring individual characteristics of tibiofemoral OA in CT and sCT images ranged from Cohen's  $\kappa$  of  $0.41$  for cysts (lowest score) on CT to  $0.88$  for osteophytes on sCT images (highest score), see Figure 4.

To make a comparison with the original OACT paper,<sup>8</sup> the results of inter and intrareader agreement in that study are shown along with our interreader agreement results in Supporting Information: material and discussed in more details in Section 4.

### 3.3 | Intermodality agreement of CT and sCT

To assess the performance of sCT compared with CT for OA scoring, the results of Cohen's  $\kappa$  analysis for each joint are determined as well.

The results of the agreement analysis show substantial to almost-perfect agreement between tibiofemoral scores on CT and sCT ( $\kappa = 0.63$  and  $0.84$  for two readers) and substantial to almost-perfect agreement between OA scores on CT and sCT for patellofemoral joint ( $\kappa = 0.78$  and  $0.81$  for the two readers) as shown in Figure 5.

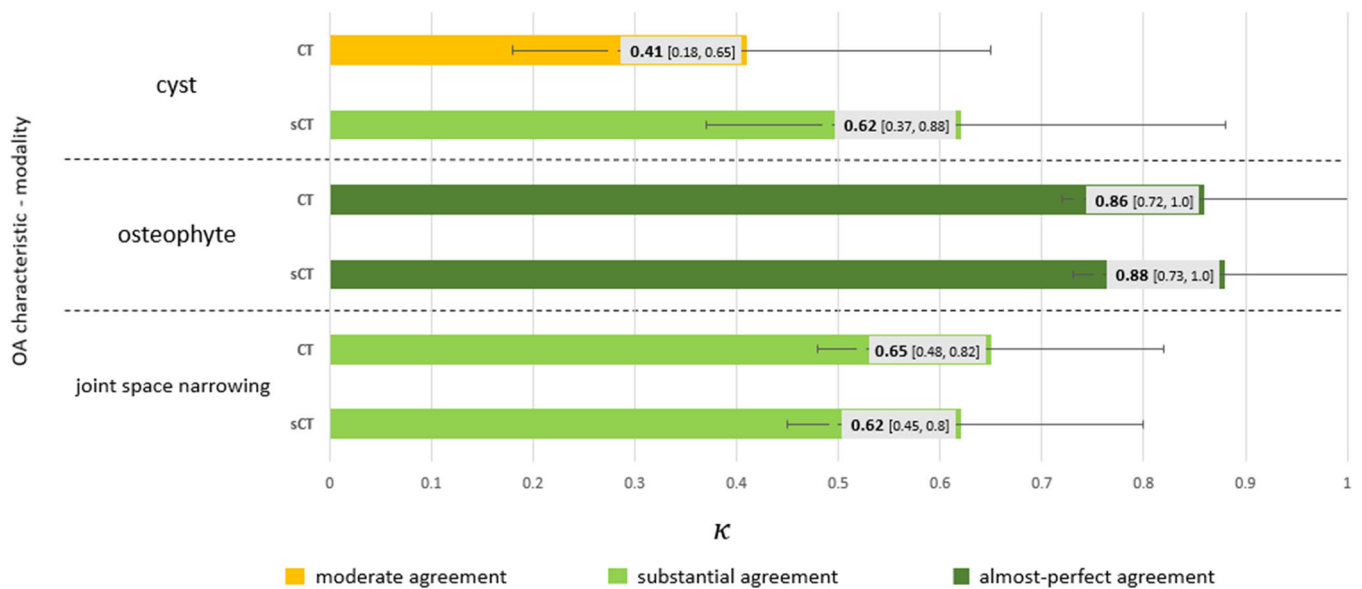
Collective intermodality agreement irrespective of reader was  $0.75$  [ $0.59, 0.91$ ] for tibiofemoral joint and  $0.8$  [ $0.64, 0.96$ ] for patellofemoral joint.

Agreement between CT and sCT scores is fair for cyst, moderate for joint space narrowing, and excellent for osteophyte for both the readers. In certain instances, disparities were observed in the visualization of cysts between CT and sCT (Figure 6).

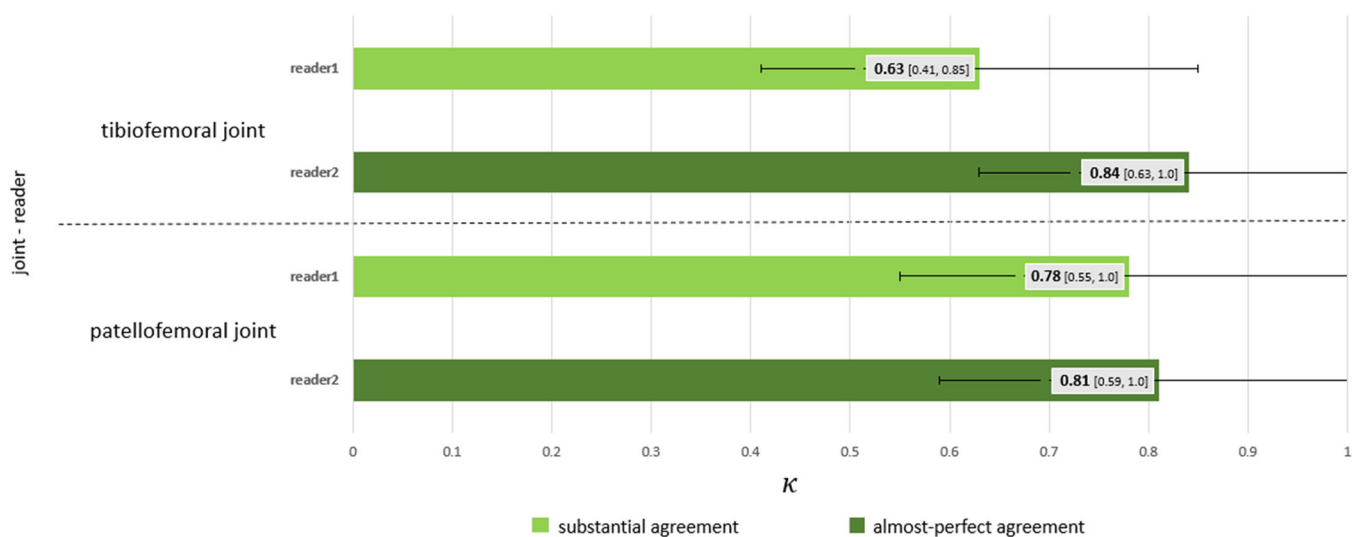
The inter-modality agreement between CT and sCT images was calculated by utilizing the individual OA characteristic scores (Figure 7).

### 3.4 | Confidence scores

The reported confidence scores show that the readers feel very confident in scoring the sCT and CT images. Ninety-five percent of readings on sCT images and 98% of readings on CT were scored with high or medium level diagnostic confidence. We presented the confidence level scores assigned to CT and sCT images by each



**FIGURE 4** Interreader agreement for scoring individual OA characteristics on CT or sCT images in terms of  $\kappa$  score with 95% confidence interval bounds shown in brackets and error bars. CT, computed tomography; OA, osteoarthritis; sCT, synthetic computed tomography.



**FIGURE 5** Agreement between whole joint OA scores on sCT and CT in terms of  $\kappa$  score with 95% confidence interval bounds shown in brackets and error bars. CT, computed tomography; OA, osteoarthritis; sCT, synthetic computed tomography.

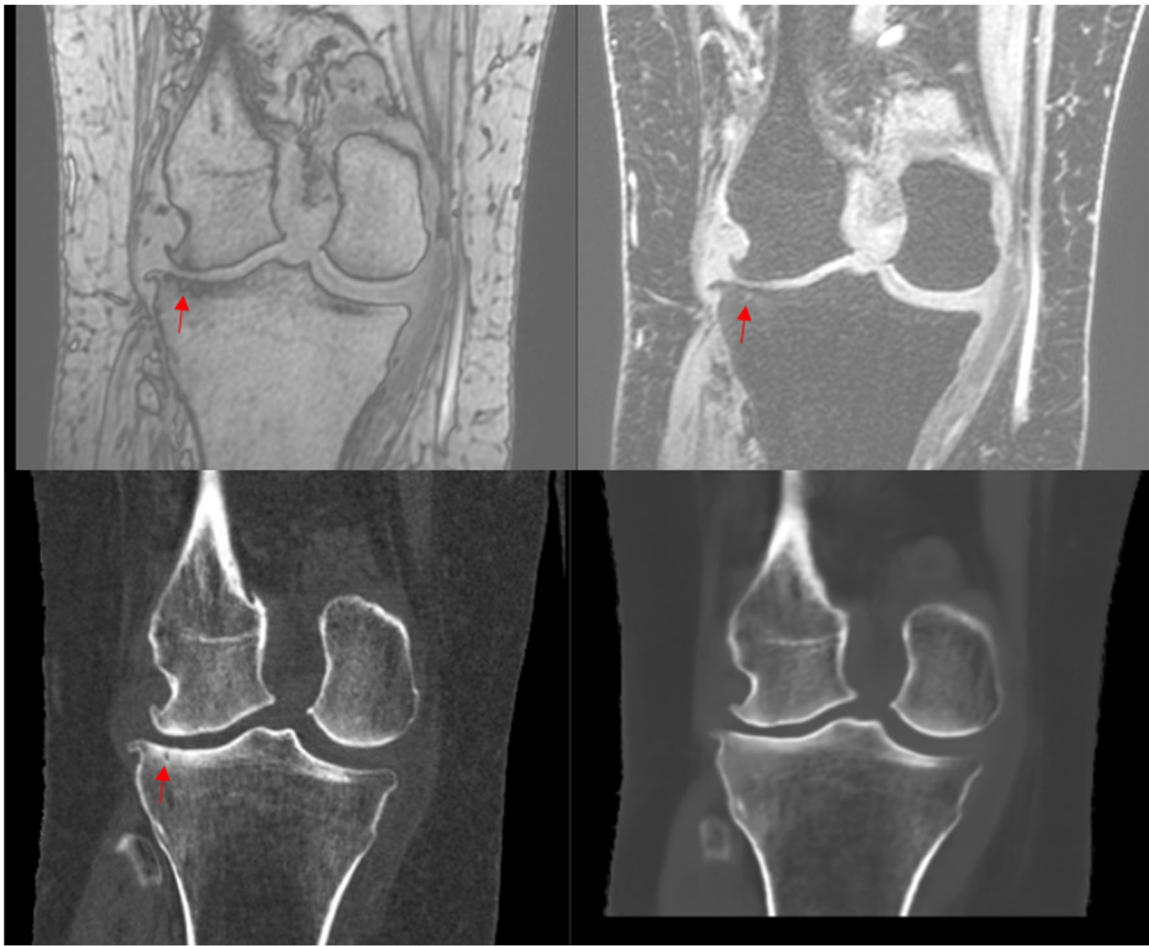
reader. The results presented in Table 2 indicate that on average, CT slightly outperforms sCT in terms of confidence scores, but both are showing acceptable confidence levels for scoring OA.

## 4 | DISCUSSION

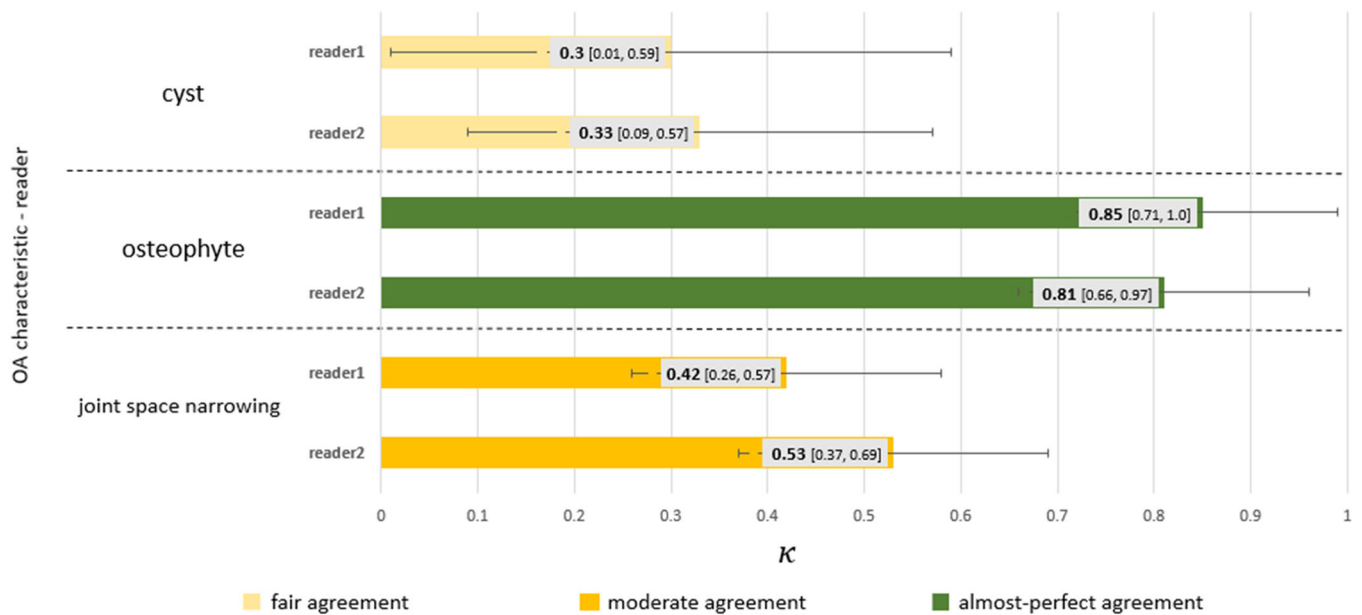
In this study, CT and sCT images were scored for OA features. Two readers conducted the scorings based on an earlier established OACT scoring method.<sup>8</sup> Interreader and intermodality agreement between scorings was calculated. We organized our presentation in Section 3 to follow a systematic approach that allowed us to first present our

findings at the joint and OA characteristic levels for both interreader and intermodality agreement. In this section, we will discuss the findings for each type of agreement in a more granular fashion, starting with joints and then moving to OA characteristics.

At the level of joints, interreader agreement for scoring OA in both tibiofemoral and patellofemoral joints, was almost-perfect for CT and sCT images. This is in agreement with interreader agreements in the OACT paper,<sup>8</sup> where agreements were substantial to almost-perfect (Supporting Information: material Table 1). Additionally, intermodality (CT-sCT) agreement in our study was substantial to almost-perfect for tibiofemoral and patellofemoral joints for the two readers. This is congruent with intrareader agreement in the OACT



**FIGURE 6** A case where synthetic CT fails to depict very small cyst. Top left: MRI first echo, top right: MRI water only, bottom left: CT, bottom right: synthetic CT. CT, computed tomography.



**FIGURE 7** Agreement between the scores of each OA characteristic on CT and sCT in terms of  $\kappa$  score with 95% confidence interval bounds shown in brackets and error bars. CT, computed tomography; OA, osteoarthritis; sCT, synthetic computed tomography.



**TABLE 2** The mean and 95% interval of confidence scores assigned to sCT and CT by each reader.

		Joint space narrowing	Osteophyte	Cyst	Tibiofemoral joint	Patellofemoral joint
Reader1	sCT	3.85 [3.75, 3.95]	3.67 [3.52, 3.81]	3.72 [3.56, 3.87]	3.74 [3.65, 3.84]	3.07 [2.8, 3.33]
	CT	3.8 [3.7, 3.9]	3.75 [3.63, 3.87]	3.92 [3.83, 4.0]	3.82 [3.75, 3.9]	3.57 [3.39, 3.75]
Reader2	sCT	3.63 [3.51, 3.76]	3.22 [3.06, 3.37]	3.47 [3.3, 3.64]	3.44 [3.34, 3.53]	3.3 [3.09, 3.51]
	CT	3.72 [3.6, 3.83]	3.93 [3.87, 4.0]	3.73 [3.6, 3.86]	3.79 [3.71, 3.88]	3.53 [3.35, 3.71]

Abbreviations: CT, computed tomography; sCT, synthetic computed tomography.

paper<sup>8</sup> being almost-perfect, and is in the same range as interreader agreement of the OACT paper<sup>8</sup> being substantial to almost-perfect (Supporting Information: material Table 1).

In terms of individual OA characteristics, we reported interreader and intermodality agreement for scoring osteophytes, joint space narrowing, and cysts.

Scoring osteophytes in our data set resulted in almost-perfect, interreader, and intermodality agreement. This is the same as interreader and intrareader analysis of the OACT study<sup>8</sup> for scoring osteophytes, showing almost-perfect agreement as well (Supporting Information: material Table 1).

Scoring joint space narrowing resulted in substantial interreader agreement for both CT and sCT datasets. This is akin to interreader agreement in the OACT paper<sup>8</sup> for scoring joint space narrowing reported between moderate to substantial (Supporting Information: material Table 1). Intermodality agreement of scoring joint space narrowing for both readers were moderate. This is in the same range as interreader agreement of OACT paper,<sup>8</sup> yet slightly lower than their reported intrareader agreement, which was substantial. As explained in the methods section, the sCT images are intrinsically registered to MR images, so the knee flexion angle and positioning of the patient can be slightly different than on CT. This can explain why intermodality agreement for scoring joint space narrowing is slightly lower than intrareader agreement in OACT paper.<sup>8</sup>

Scoring cyst in our study resulted in moderate interreader agreement in CT images. This is congruent with the inter-reader agreement of OACT paper<sup>8</sup> which was in the range of fair to moderate. Scoring cyst on sCT data set, resulted in substantial interreader agreement, which is slightly higher than interreader agreement on CT. We also measured intermodality agreement of scoring cysts as fair. This is also in the same range as interreader agreement of OACT data set as fair to moderate, but is lower than OACT intrareader agreement reported as almost-perfect. Visual analysis of images shows that some very small cysts are not reconstructed in sCT images. This can be effect of difference in resolutions of MR and CT or training settings, among others. As a result, some very small cysts could be scored as present on CT scans while are not visible on sCT scans. The fact that some very small cysts are not reconstructed on sCT can also be an explanation for slightly higher inter-reader agreement in sCT compared with CT, since the small cysts, which are probably the cause of lower inter-reader agreement on CT are not visible on some of our sCT images. Investigation of this issue can be topic of a future study. Also, cysts

are generally harder to score for readers, as shown by lower interreader agreement in this study and the previous OACT study.<sup>8</sup>

Limitations of the study warrant discussion. Mainly, although we have done randomization and shuffling of data, which restrains the readers from knowing which patient they are scoring, still due to texture and resolution differences between sCT and CT images, a trained reader could discern whether they were scoring sCT or CT. This should not cause bias in pathology scoring nor in the agreement analyses reported, though it could influence the objectivity of the confidence scores. As another limitation, the comparison is solely based on scorings of CT and sCT imaging data, which means that we consider CT scores as gold standard since it is known as the best available imaging technique for bone evaluation, but this may not always be perfect with respect to the actual joint structure. Although MR might underestimate some pathologies like knee osteophytes compared with CT,<sup>30</sup> it is considered the gold standard for some other pathologies like cystic lesions in the knee.<sup>31</sup> It could be an interesting future study topic to discuss the effect of image enhancement techniques on pathology classification in MR images. Another limitation of the study was that we have only considered scoring values in our study and not determined continuous measurements like joint space width or bone density. This can be the topic of a future study. Another limitation lies in the fact that this has been a single-center study. In this study, the sCT generation model has been trained on 39 patients from the IMI-APPROACH (UMC Utrecht) data set and tested on the data from another 30 patients that came from the same cohort. Generalizing the sCT generation model to any other data set and data acquired from more centers needs further investigation and might raise the need for training on new or extended datasets. Although the variations of OA features in the data set shown in the results section seem to be diverse, it might not be representative of features in the general population, which can only be addressed in a larger study that includes subjects from multiple OA subgroups and controls.

## 5 | CONCLUSION

In this study, the result of OA scoring of CT and sCT images by two readers in a single-center data set shows that sCT provides comparable scoring accuracy to CT for knee OA scoring. This is the first step for assessment of sCT usage in clinical and research purposes of knee OA, creating an MR-only noninvasive workflow for both bone and soft tissue visualization.

## AUTHOR CONTRIBUTIONS

Peter Seevinck, Wouter Foppen, Willem Paul Gielis, Harrie Weinans, and Saeed Arbabi contributed to the study design. Peter Seevinck, Marijn van Stralen, Pim A. de Jong, Vahid Arbabi, and Saeed Arbabi contributed to the data acquisition. Wouter Foppen and Willem Paul Gielis scored the images for osteoarthritis. Saeed Arbabi and Mylène Jansen performed data processing, statistical analysis, and visualizations. Saeed Arbabi drafted the article. Wouter Foppen, Willem Paul Gielis, Marijn van Stralen, Mylène Jansen, Vahid Arbabi, Pim A. de Jong, Harrie Weinans, and Peter Seevinck contributed to critical revision of the manuscript. All authors have read and approved the final submitted manuscript.

## ACKNOWLEDGMENTS

This research received funding from the EU's H2020 research, innovation program under Marie S. Curie co-fund RESCUE grant agreement No 801540 and APPROACH, which has received support from the Innovative Medicines Initiative Joint Undertaking under Grant Agreement no 115770, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. See [www.imi.europa.eu](http://www.imi.europa.eu) and [www.approachproject.eu](http://www.approachproject.eu).

## CONFLICT OF INTEREST STATEMENT

Marijn van Stralen and Peter R. Seevinck are minority shareholders at MRGuidance B.V. Other authors declare that there is no conflict of interest.

## ORCID

Saeed Arbabi  <http://orcid.org/0000-0003-0831-4015>

Wouter Foppen  <http://orcid.org/0000-0003-4970-8555>

Mylène Jansen  <http://orcid.org/0000-0003-1929-6350>

## REFERENCES

- Hayashi D, Roemer FW, Jarraya M, Guermazi A. Imaging in osteoarthritis. *Radiol Clin North Am*. 2017;55(5):1085-1102. doi:10.1016/j.rcl.2017.04.012
- Guermazi A, Roemer FW, Crema MD, Englund M, Hayashi D. Imaging of non-osteochondral tissues in osteoarthritis. *Osteoarthritis Cartilage*. 2014;22(10):1590-1605. doi:10.1016/j.joca.2014.05.001
- Sakellariou G, Conaghan PG, Zhang W, et al. EULAR recommendations for the use of imaging in the clinical management of peripheral joint osteoarthritis. *Ann Rheum Dis*. 2017;76(9):1484-1494. doi:10.1136/annrheumdis-2016-210815
- Fox MG, Chang EY, Amini B, et al. ACR appropriateness criteria® chronic knee pain. *J Am Coll Radiol*. 2018;15(11):S302-S312. doi:10.1016/j.jacr.2018.09.016
- Hunter DJ, Altman RD, Cicuttini F, et al. OARSI clinical trials recommendations: knee imaging in clinical trials in osteoarthritis. *Osteoarthr Cartil*. 2015;23(5):698-715. doi:10.1016/j.joca.2015.03.012
- Roemer FW, Demehri S, Omoumi P, et al. State of the art imaging of osteoarthritis. *Radiology*. 2020;47(3):2009.
- Florkow MC, Willemsen K, Zijlstra F, et al. MRI-based synthetic CT shows equivalence to conventional CT for the morphological assessment of the hip joint. *J Orthop Res*. 2021;40(January):954-964. doi:10.1002/jor.25127
- Gielis WP, Weinans H, Nap FJ, Roemer FW, Foppen W. Scoring osteoarthritis reliably in large joints and the spine using whole-body ct: osteoarthritis computed tomography-score (oact-score). *J Pers Med*. 2021;11(1):5. doi:10.3390/jpm11010005
- Florkow MC, Willemsen K, Mascarenhas VV, Oei EHG, van Stralen M, Seevinck PR. Magnetic resonance imaging versus computed tomography for Three-Dimensional bone imaging of musculoskeletal pathologies: a review. *J Magn Reson Imaging*. 2022;56:11-34. doi:10.1002/jmri.28067
- Bharadwaj UU, Coy A, Motamedi D, et al. CT-like MRI: a qualitative assessment of ZTE sequences for knee osseous abnormalities. *Skeletal Radiol*. 2022;51(0123456789):1585-1594. doi:10.1007/s00256-021-03987-2
- Harris MD, Datar M, Whitaker RT, Jurrus ER, Peters CL, Anderson AE. Statistical shape modeling of cam femoroacetabular impingement. *J Orthop Res*. 2013;31(10):1620-1626. doi:10.1002/jor.22389
- Li Q, Amano K, Link TM, Ma CB. Advanced imaging in osteoarthritis. *Sport Heal: A Multidiscip Approach*. 2016;8(5):418-428. doi:10.1177/1941738116663922
- Bousson V, Lowitz T, Laouisset L, Engelke K, Laredo JD. CT imaging for the investigation of subchondral bone in knee osteoarthritis. *Osteoporos Int*. 2012;23(8 suppl):861-865. doi:10.1007/s00198-012-2169-5
- Edmund JM, Nyholm T. A review of substitute CT generation for MRI-only radiation therapy. *Radiat Oncol*. 2017;12(1):28. doi:10.1186/s13014-016-0747-y
- Zijlstra F, Willemsen K, Florkow MC, et al. CT synthesis from MR images for orthopedic applications in the lower arm using a conditional generative adversarial network (Published online 2019: 54). doi:10.1117/12.2512857
- Jans LBO, Chen M, Elewaut D, et al. MRI-based synthetic CT in the detection of structural lesions in patients with suspected sacroiliitis: comparison with MRI. *Radiology*. 2021;298:343-349. doi:10.1148/radiol.2020201537
- Staatjes VE, Seevinck PR, Vandertop WP, van Stralen M, Schröder ML. Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept. *Neurosurg Focus*. 2021;50(1):E13. doi:10.3171/2020.10.focus.20801
- van Helvoort EM, van Spil WE, Jansen MP, et al. Cohort profile: the applied Public-Private research enabling OsteoArthritis clinical headway (IMI-APPROACH) study: a 2-year, European, cohort study to describe, validate and predict phenotypes of osteoarthritis using clinical, imaging and biochemical markers. *BMJ Open*. 2020;10(7):e035101. doi:10.1136/bmjopen-2019-035101
- Altman R, Asch E, Bloch D, et al. Development of criteria for the classification and reporting of osteoarthritis: classification of osteoarthritis of the knee. *Arthritis Rheumatism*. 1986;29(8):1039-1049. doi:10.1002/art.1780290816
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 9351. Springer Verlag; 2015:234-241. doi:10.1007/978-3-319-24574-4\_28
- Florkow MC, Zijlstra F, Willemsen K, et al. Deep learning-based MR-to-CT synthesis: the influence of varying gradient echo-based MR images as input channels. *Magn Reson Med*. 2020;83(4):1429-1441. doi:10.1002/mrm.28008
- Kuiper RJA, Van Stralen M, Sakkars RJB, et al. CT to MR registration of complex deformations in the knee joint through dual quaternion interpolation of rigid transforms. *Phys Med Biol*. 2021;66(17):175024. doi:10.1088/1361-6560/ac1769

23. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* 1973;33:613-619.
24. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *JCE.* 2011;64(1):96-106. doi:10.1016/j.jclinepi.2010.03.002
25. Bujang MA, Baharum N. Guidelines of the minimum sample size requirements for Cohen's Kappa. *Epidemiol Biostat Public Heal.* 2017;14(2):e12267-1-e12267-10. doi:10.2427/12267
26. Rotondi MMA. KappaSize: Sample Size Estimation Functions for Studies of Interobserver Agreement. R Packag Version 12. Published online 2018.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159. doi:10.2307/2529310
28. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res.* 2011;12:2825-2830.
29. Team RC. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing.
30. Roemer F, Engelke K, Laredo J-D, Li L, Guermazi A. Mri underestimates presence and size of knee osteophytes using ct as a reference standard. *Osteoarthr Cartil.* 2022;30(February):S264. doi:10.1016/j.joca.2022.02.360
31. Beaman FD, Peterson JJ. MR imaging of cysts, ganglia, and bursae about the knee. *Radiol Clin North Am.* 2007;45(6):969-982. doi:10.1016/j.rcl.2007.08.005

### SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Arbabi S, Foppen W, Gielis WP, et al. MRI-based synthetic CT in the detection of knee osteoarthritis: comparison with CT. *J Orthop Res.* 2023;41:2530-2539. doi:10.1002/jor.25557