scientific data

OPEN

DATA DESCRIPTOR

Multi-scanner and multi-modal lumbar vertebral body and intervertebral disc segmentation database

Yasmina Al Khalil[1], Edoardo A. Becherucci[2], Jan S. Kirschke[2,3], Dimitrios C. Karampinos[5], Marcel Breeuwer[1], Thomas Baum[2] & Nico Sollmann[2,3,4,6]✉

Magnetic resonance imaging (MRI) is widely utilized for diagnosing and monitoring of spinal disorders. For a number of applications, particularly those related to quantitative MRI, an essential step towards achieving reliable and objective measurements is the segmentation of the examined structures. Performed manually, such process is time-consuming and prone to errors, posing a bottleneck to its clinical applicability. A more efficient analysis would be achieved by automating a segmentation process. However, routine spine MRI acquisitions pose several challenges for achieving robust and accurate segmentations, due to varying MRI acquisition characteristics occurring in data acquired from different sites. Moreover, heterogeneous annotated datasets, collected from multiple scanners with different pulse sequence protocols, are limited. Thus, we present a manually segmented lumbar spine MRI database containing a wide range of data obtained from multiple scanners and pulse sequences, with segmentations of lumbar vertebral bodies and intervertebral discs. The database is intended for the use in developing and testing of automated lumbar spine segmentation algorithms in multi-domain scenarios.

Background & Summary

Magnetic resonance imaging (MRI) is the modality of choice for detecting and visualizing almost all spinal disorders, as it provides non-invasive soft tissue visualization with excellent contrast, more detailed compared to other modalities[1–5]. Thus, it is widely utilized in orthopedic and neurosurgical diagnostics, ranging from dedicated imaging in scoliosis, intervertebral disc disease, and osteoporosis to injuries including vertebral fractures as well as to computer-assisted surgical intervention planning and guidance[1–4,6].

Visual image reading represents the standard approach in the clinical setting for evaluation of routine spine MRI. Reporting derived from visual image assessment by the radiologist may be enhanced by using the approach of structured reporting (e.g., use of predefined formats and terms to create radiological reports, using a high level of standardized and organized information in template context[7–9]) and by implementing semi-quantitative grading schemes (e.g., Pfirrmann classification for lumbar disc degeneration[10]). Going one step further, generating quantitative measures for specific anatomical structures along the spine would be welcome to be able to provide meaningful objective markers related to spinal disorders, which could facilitate patient phenotyping, clinical management, and adequate treatment selection. A robust and precise segmentation of vertebral bodies and intervertebral discs is a major step towards a reliable diagnosis of various conditions in automated and computer-assisted systems, as well as for quantitative MRI regarding extraction of image-based markers[11–15].

[1]Biomedical Engineering Department, Eindhoven University of Technology, Eindhoven, The Netherlands. [2]Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [3]TUM-Neuroimaging Center, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [4]Department of Diagnostic and Interventional Radiology, University Hospital Ulm, Ulm, Germany. [5]Department of Diagnostic and Interventional Radiology, School of Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [6]Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, CA, USA. ✉e-mail: nico.sollmann@tum.de

ID	Gender	Age	Clinical indication	Clinical indication code*
02	Female	51.1	FU resection ependymoma	1
04	Male	38.3	Postop resection neurinoma	1
05	Female	30.0	FU resection ependymoma	1
07	Male	77.2	LBP and radiculopathy	2
10	Female	73.6	FU resection neurinoma	1
11	Female	40.1	Mamma-Ca with spinal metastases	3
12	Male	48.1	FU resection ependymoma	1
13	Female	52.3	Screening spinal tumor	3
14	Male	57.2	FU inflammatory lesion	4
15	Female	82.4	LBP	2
16	Male	61.2	Thymus-Ca with spinal metastases	3
17	Female	70.0	Spondylodiscitis	4
18	Female	61.1	LBP and radiculopathy	2
19	Female	71.7	Lung-Ca with spinal metastases	3
20	Male	48.9	Radiculopathy S1 left	2
21	Male	71.0	Radiculopathy L4 (both sides)	2
22	Female	86.5	FU resection meningioma	1
23	Male	45.9	Radiculopathy S1 right	2
24	Female	72.1	Sacral fracture	5
25	Female	52.9	Lung-Ca with spinal metastases	3
26	Female	40.9	LBP and radiculopathy	2
27	Female	49.1	LBP and radiculopathy	2
28	Male	64.0	Spondylodiscitis	4
29	Female	88.1	Mamma-Ca with spinal metastases	3
30	Male	81.7	LBP and radiculopathy	2
31	Male	55.7	LBP and radiculopathy	2
32	Male	39.8	Postop resection meningioma	1
33	Male	62.3	LBP and radiculopathy	2
34	Female	57.2	Radiculopathy L5 left	2
35	Male	69.6	Spondylodiscitis	4
36	Female	56.0	FU resection ependymoma	1
37	Female	79.1	LBP and radiculopathy	2
38	Female	75.8	LBP and radiculopathy	2
39	Female	43.1	Spondylodiscitis	4

Table 1. Patient characteristics with clinical indications. *Clinical indication code legend: 1 Postoperative/follow-up (FU) imaging for tumor after resection, 2 Low back pain (LBP) with or w/o radiculopathy due to degenerative changes, 3 Malignancy with (suspected) spinal metastases, 4 Spondylodiscitis or other inflammation/infection, 5 Trauma/fracture.

However, automated spine segmentation has not yet made the transition to clinical routine and remains a challenging problem due to the variable and complex shape of the spine anatomy, as well as the presence of noise and other artefacts in imaging data[6,16]. Moreover, routine spine MRI acquisitions pose several challenges for achieving robust and accurate segmentations, which is mostly due to some unavoidable MRI acquisition characteristics and pitfalls. Specifically, these include the presence of partial volume effects related to anisotropic spatial resolution, non-homogeneous intensities between central and marginal areas of the spine due to bias field artefacts, and the non-existence of standardized measurement units (unlike Hounsfield Units as used for computed tomography [CT])[6]. This is additionally highlighted in large multi-site studies, where variations in scan parameters often produce images that vary significantly in contrast and quality, especially in cases where multiple diagnostic MRI sequences are used among different clinical practices[17,18].

Up to now, diagnostic and analysis methods of MRI data predominantly rely on manual annotation of anatomical objects, such as vertebrae and intervertebral discs. This process is time-consuming and subjective, and often prone to intra- and inter-annotator variability[17,19]. Of note, the high time expenditure of manual segmentation may hamper direct use, particularly when multi-level segmentations are desired (e.g., segmentation of the entire lumbar spine). Thus, automating the segmentation process would strongly benefit clinicians and scientists, especially for large-scale and multi-site studies targeting quantitative MRI. However, a reliable automated segmentation method, which has the potential to be included in standard clinical routines in the future, has to be able to generalize to the large variety of MRI sequences and parameter settings, while being reasonably fast and not overly complex to operate. This remains a difficult challenge, despite the surge in recent developments of automated segmentation techniques[18,20].

ID	No. of scans	MRI 1 vendor/model	MRI 1 sequence*
02	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
04	2	Siemens Espree	T1-w, nc T1-w, STIR
05	2	Philips Achieva	T1-w, nc T1-w
07	2	Siemens Avanto	nc T1-w, T2-w, STIR
10	2	Siemens Verio	nc T1-w, T1-fs, T2-w
11	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
12	2	Philips Ingenia	T1-w, nc T1-w
13	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
14	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
15	2	Siemens Symphony	nc T1-w, T2-w
16	2	Siemens Avanto	nc T1-w, T2-w, STIR
17	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
18	1	Philips Achieva	nc T1-w, T2-w Dix.
19	2	Siemens Avanto	nc T1-w, T2-w
20	1	Philips Achieva	nc T1-w, T2-w Dix.
21	2	Siemens Amira	nc T1-w, T2-w
22	2	Siemens Verio	T1-w, nc T1-w, T2-w
23	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
24	1	Philips Elition	nc T1-w, T2-w Dix.
25	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
26	1	Philips Achieva	nc T1-w, T2-w Dix.
27	1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
28	2	Siemens Verio	nc T1-w, T2-w, STIR
29	2	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
30	2	Siemens Espree	nc T1-w, T2-w
31	1	Philips Ingenia	nc T1-w, T2-w Dix.
32	2	Philips Elition	T1-w, nc T1-w, T2-w Dix.
33	2	Siemens Aera	nc T1-w, T2-w
34	2	Philips Achieva	nc T1-w, T2-w Dix.
35	2	Philips Elition	nc T1-w
36	2	Siemens Magnetom	T1-w, nc T1-w, T2-w
37	2	Philips Achieva	nc T1-w, T2-w
38	2	GE Signa	nc T1-w, T2-w
39	2	Siemens Avanto	T1-w, nc T1-w, T2-w, STIR

Table 2. Scan characteristics with image sequences per scanner vendor and model type for the first scan. *T1-w, nc T1-w, T2-w, T1-fs, STIR, and T2-w Dix. stand for T1-weighted contrast-enhanced, T1-weighted non-contrast-enhanced, T2-weighted, T1-weighted fat-saturated, short tau inversion recovery, and T2-weighted Dixon sequences.

Most existing methods focus on one particular pulse sequence, which is not sufficient to adequately acknowledge the frequent multi-sequence acquisitions in clinical routine, or they may suffer from considerably long computational time, as well as the requirement for user input and navigation[21,22]. Other methods require extensive prior knowledge, such as shape information, to construct a representative preliminary model of anatomical shape from training data and to approximate optimization to new data[16,23,24]. Such methods are clearly limited by prior knowledge, which – in the case of large-scale and multi-site studies – should contain enough representative information to handle all possible variations in data. Current constraints in data collection and data sharing pose a challenge in acquiring enough of such data, thus producing highly specific models that are often only applicable to a small variety of cases. Automated medical image analysis algorithms should be robust enough to inherent data variability to ensure their successful integration into existing infrastructure, with the long-term perspective of becoming clinically feasible. However, without enough representative data comprising all of the diverse variations in spine MRI, algorithm generalizability is hard to achieve.

Recent advances in deep learning-based segmentation methods hold a lot of promise to alleviate the challenges described above[21,25–28]. This seems especially true for improving the generalization on multi-site and multi-scanner data. However, used algorithms require large datasets annotated by experts for both development and testing. Yet, publicly available datasets of annotated vertebrae and intervertebral discs, collected from multiple scanners and different acquisition protocols, are very limited. Therefore, the purpose of this article and its related dataset is to provide a reference database containing a wide range of MRI data obtained from multiple scanners and involving various pulse sequences, along with the segmentations of lumbar vertebral bodies (L1 to L5) and intervertebral discs.

ID	T*	MRI 2 vendor/model	MRI 2 sequence**
04	2.8	Philips Ingenia	T1-w, nc T1-w, T2-w Dix.
05	15.6	Siemens Verio	T1-w, nc T1-w, T2-w, STIR
07	13.7	Philips Achieva	nc T1-w, T2-w Dix.
10	4.5	Philips Ingenia	T1-w, nc T1-w, T2-w Dix.
12	9.3	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
15	1.6	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
16	0.7	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
19	2.8	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
21	6.1	Philips Elition	nc T1-w, T2-w Dix.
22	8.7	Philips Elition	nc T1-w, T2-w Dix.
28	0.5	Philips Achieva	nc T1-w, T2-w Dix.
29	6.2	Philips Ingenia	T1-w, nc T1-w, T2-w Dix.
30	12.2	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
32	0.7	Philips Achieva	T1-w, nc T1-w, T2-w Dix.
33	13.7	Philips Achieva	nc T1-w, T2-w Dix.
34	7.1	Siemens Avanto	nc T1-w, T2-w, STIR
35	2.7	Philips Ingenia	T1-w, nc T1-w, T2-w Dix.
36	2.8	Philips Achieva	nc T1-w, T2-w Dix.
37	1.6	Philips Achieva	nc T1-w, T2-w Dix.
38	60.8	Philips Achieva	nc T1-w, T2-w Dix.
39	1.1	Philips Achieva	T1-w, nc T1-w, T2-w Dix.

Table 3. Scan characteristics with image sequences per scanner vendor and model type for the follow-up scan. *T is the time interval between the two scans. **T1-w, nc T1-w, T2-w, T1-fs, STIR, and T2-w Dix. stand for T1-weighted contrast-enhanced, T1-weighted non-contrast-enhanced, T2-weighted, T1-weighted fat-saturated, short tau inversion recovery, and T2-weighted Dixon sequences.

Using a database like the herein presented, automatization in image analysis and processing could be further facilitated, which could directly influence the field of radiology and pave the way towards semi-automated and fully-automated algorithms for radiological diagnostics. As such, over the recent years, automated spine image analysis has seen a growing interest, particularly for the detection of vertebral fractures[29], assessment of spinal deformities[21], and computer-assisted surgical interventions[30]. As the lumbar spine is a particularly common site for various spinal disorders that can cause chronic low back pain (LBP), segmentation of anatomical structures along the lumbar spine is of considerable interest. While many factors can contribute to spinal disease and deformation, such as fractures, accidental injuries, osteoporosis, vertebral neoplasm or scoliosis[31], the primary cause of chronic LBP in most patients is believed to be related to spinal degeneration, particularly of lumbar intervertebral discs and endplates[32–37]. Spinal degeneration is a natural consequence of aging, but it can be accelerated by trauma, repetitive stress, systemic disease, and other factors[35,38–41]. Thus, investigating the relevance between chronic LBP pathology and lumbar vertebrae and intervertebral disc morphology and composition using specific segmentations could aid in developing appropriate methods to assist early diagnosis, provide better surgical planning, and facilitate individualized treatment strategies and patient phenotyping.

In summary, we offer a database of manually segmented lumbar vertebral bodies and intervertebral discs in MRI datasets collected from a variety of scanners and using different pulse sequence protocols. Besides the images that are part of standard scanning routines, such as non-contrast-enhanced and contrast-enhanced T1-weighted, T2-weighted, and short tau inversion recovery (STIR) images, we provide the access to images obtained with a DIXON turbo spin-echo (TSE) sequence. Thus, combined with other available imaging sequences, this data could help in the development of more efficient and robust methods of segmenting musculoskeletal structures, and it could facilitate quantitative MRI by achieving automated segmentation and analysis of datasets, which is useful for assessing vertebral bodies (e.g., fat fraction[42–46]), cartilage endplates (e.g., T2* [47,48]), and intervertebral discs (e.g., T1rho mapping[44,48,49]). The data including segmentations can be used as training and test datasets for (semi-)automated algorithms, which can potentially benefit from the heterogeneous character of this study's database, as it fosters the development of approaches that are more generalizable to data derived from multiple sites and scanners, which is known to be a major challenge for models that are typically trained on highly homogeneous data.

Methods

Patient cohort. This retrospective study was approved by the local Institutional Review Board (Ethikkommission der Technischen Universität München). The requirement for written informed study consent was waived due to the retrospective character of the specific analyses used for this publication. Yet, all patients provided standard informed consent for MRI scanning during clinical routine and agreed on the day of visit for MRI acquisition on an opt-in basis that their data may be used for scientific purposes. The patients were informed that their data may be anonymously publicly shared without any personal identifiable information, and the analyses for this study were performed on de-identified data.

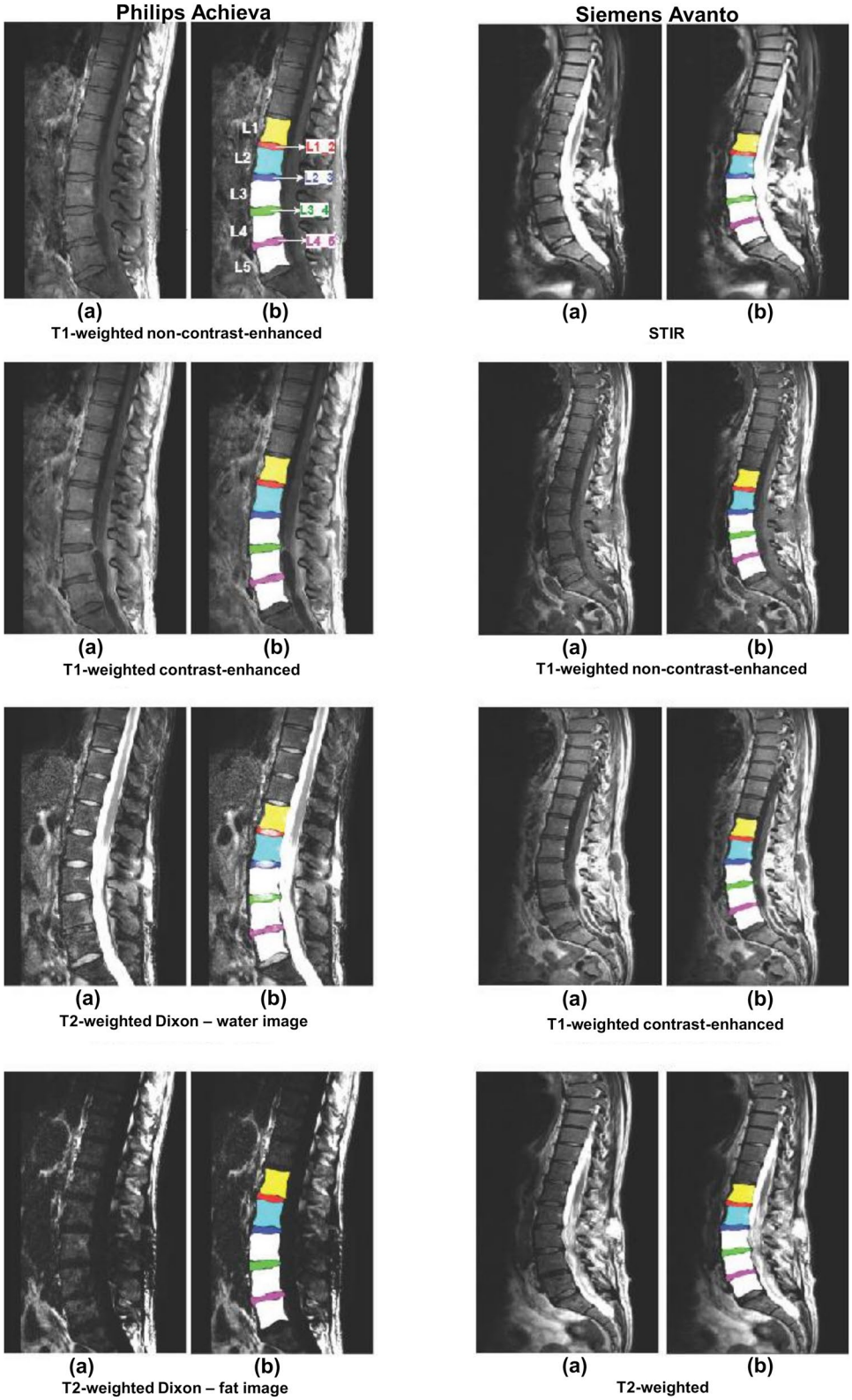

Fig. 1 Segmented lumbar vertebral bodies (L1 to L5) and intervertebral discs (L1_2, L2_3, L3_4, and L4_5) per sequence. One middle slice of each sequence is shown in (**a**) with corresponding segmentation masks in (**b**).

Thirty-four patients (mean age: 60.4 ± 15.2 years, age range: 30.0–88.1 years, 58.8% females) with the following medical indications for MRI acquisition during the clinical routine were included in this study: 1) LBP with or without radiculopathy due to suspected spinal degeneration (41.3% of patients), 2) postoperative or follow-up imaging after resection of a spinal tumor (23.5% of patients), 3) known malignancy with (suspected) spinal

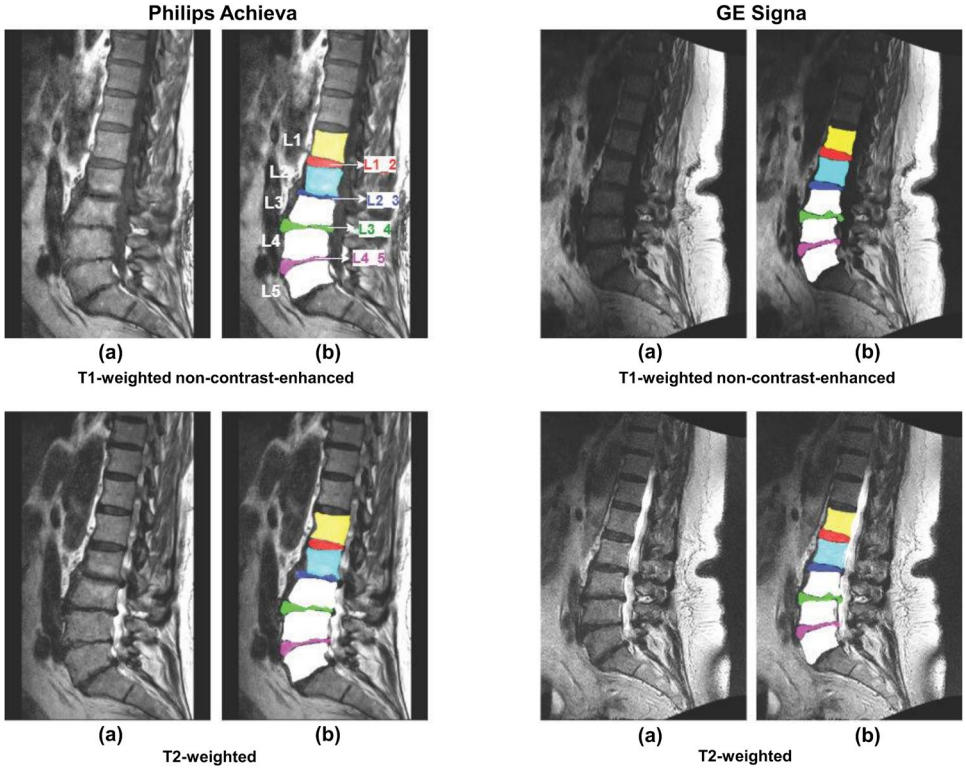

Fig. 2 Segmented lumbar vertebral bodies (L1 to L5) and intervertebral discs (L1_2, L2_3, L3_4, and L4_5) per sequence acquired from two different scanner vendors. Significant qualitative differences arise due to scanner and protocol variation. One middle slice of each sequence is shown in (**a**) with corresponding segmentation masks in (**b**).

metastases (17.6% of patients), 4) spondylodiscitis or other spinal inflammatory/infectious diseases (14.7% of patients), and 5) spinal fracture (2.9% of patients). Patient characteristics (sex and age) and medical indications are listed in Table 1.

The database contains MRI datasets collected from these 34 patients, whereby 21 patients were rescanned at a second time point using a different MRI scanner model and/or sequence protocol (mean interval between first and second MRI scan: 8.3 ± 12.6 months, interval range: 0.5–60.8 months). From the remaining 13 patients only MRI datasets of one imaging time point were used. MRI scans acquired between August 2014 and October 2019 were considered in this study, with the data being acquired either on institutional-intern scanners or on scanners of other institutions, thus being available after image transfer due to clinical requests. A detailed list of pulse sequences and MRI vendors per patient, together with the time intervals between scans, are shown in Tables 2 and 3.

Magnetic resonance imaging. Various scanner models from different vendors were utilized for collecting a number of images, including non-contrast-enhanced T1-weighted, contrast-enhanced T1-weighted, T2-weighted, STIR, T2-weighted DIXON, and T1-weighted fat-saturated (fs) images. 70.9% of included datasets were derived from Philips scanners (Achieva, Ingenia, and Elition), 27.3% were derived from Siemens scanners (Avanto, Verio, Espree, Symphony, Amira, Aera, and Magnetom). One dataset (1.8%) was taken from a GE scanner (Signa).

The field of view (FOV) covered at least the lumbar spine. Only scans that were performed in supine position were included. Each patient had at least one MRI scan performed on a Philips system with the following reference sequences:

- Sagittal non-contrast-enhanced T1-weighted sequence: repetition time (TR)/echo time (TE) = 600/8 ms, FOV = 180 × 275 × 49 mm, acquisition voxel size = 0.80 × 1.00 × 3.00 mm^3, acquisition duration = 3 min 3 s.
- Sagittal T2-weighted DIXON TSE sequence: TR/TE = 2,500/100 ms, FOV = 180 × 275 × 49 mm, acquisition voxel size = 0.70 × 0.98 × 3.00 mm^3, acquisition duration = 3 min 25 s.

Image segmentation. Non-contrast-enhanced T1-weighted images were used for obtaining the segmentations of lumbar vertebral bodies (L1 to L5) and intervertebral discs (L1/2 to L4/5) for each patient and each respective MRI dataset. Segmentations were performed manually and without (semi-)automatic software support. Specifically, non-contrast-enhanced T1-weighted images were successively opened in MITK, an open-access image viewer software package that allows a simultaneous visualization of the images in all three image planes ([http://mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK)]; German Cancer Research Center,

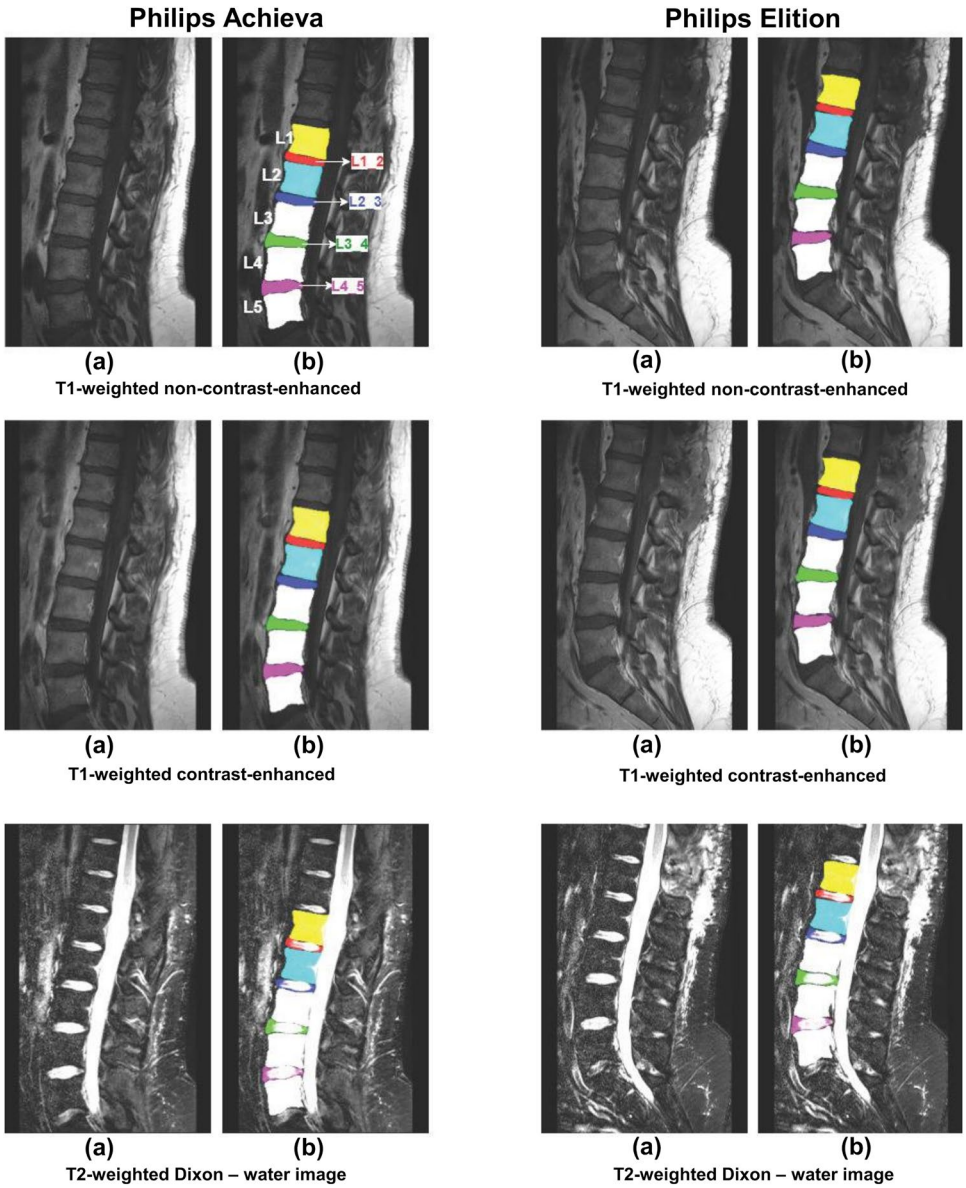

Fig. 3 Segmented lumbar vertebral bodies (L1 to L5) and intervertebral discs (L1_2, L2_3, L3_4, and L4_5) per sequence acquired from two different scanner models belonging to the same vendor. One middle slice of each sequence is shown in (**a**) with corresponding segmentation masks in (**b**).

Division of Medical and Biological Informatics, Medical Imaging Interaction Toolkit, Heidelberg, Germany). The manual segmentations were done in the sagittal plane for each vertebral body and intervertebral disc, respectively. Segmentation was performed by a medical doctor. Regions of interest (ROIs) were carefully drawn at the boundaries of the vertebrae and intervertebral discs, respectively. The other image planes were used to check for accidentally included structures not belonging to the vertebral bodies or intervertebral discs. In these cases, ROIs were corrected accordingly. For lumbar vertebral bodies, each vertebrae was separately enclosed in sagittal slices, avoiding any inclusion of paraspinal tissue or fluid. Posterior elements were not considered, thus restricting the segmentations to the vertebral corpora only. Analogously, lumbar intervertebral discs were also separately segmented in sagittal slices, considering the entire disc of the whole circumference. All segmentations were supervised by a radiologist with eleven years of experience. Segmentation time per one T1-weighted image series amounted to 60–90 min.

The obtained segmentation labels, generated in non-contrast-enhanced T1-weighted images of all patients of one or two MRI sessions, were then overlaid over the rest of the available sequences acquired with different MRI scanners and pulse protocols. Due to the differences in imaging and scanning parameters, the original images were registered to the non-contrast-enhanced T1-weighted images using the 3D linear registration tool in ITK-SNAP (https://www.itksnap.org), ensuring that labels were accurately overlaid. We resampled each scan in the space of the respective non-contrast-enhanced T1-weighted image using linear interpolation with an identity transform[50,51], given that we have the case of two or more scans of the same patient with different

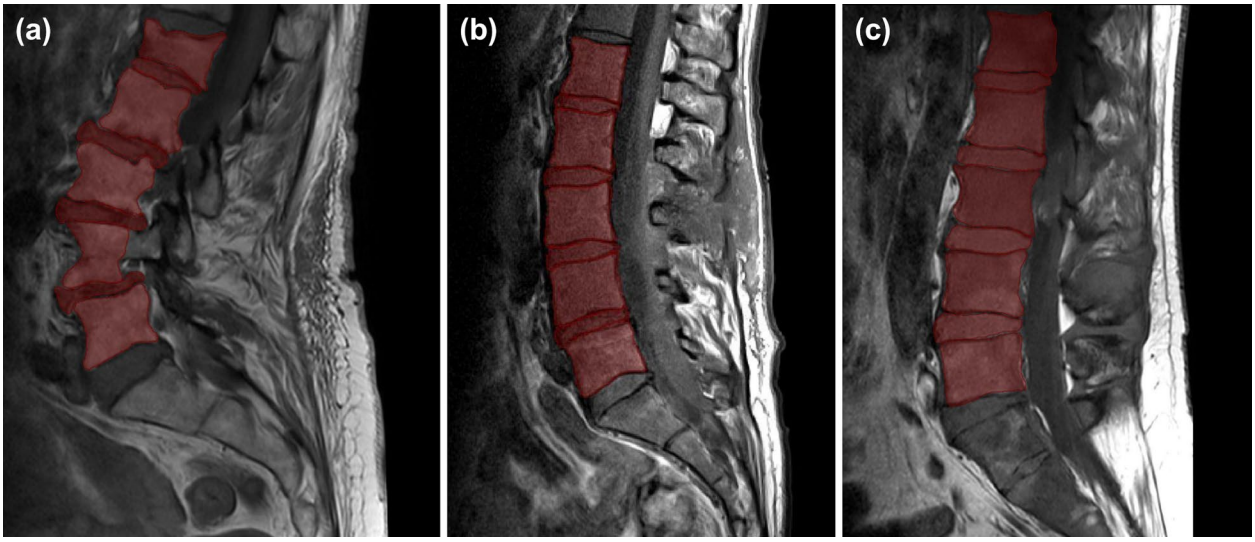

Fig. 4 Segmented lumbar vertebral bodies (L1 to L5) and intervertebral discs (L1_2, L2_3, L3_4, and L4_5) using sagittal T1-weighted sequences in a patient with a fractured vertebral body L1 (**a**), a patient with spondylodiscitis of the segment L4/L5 (**b**), and a patient with diffuse metastatic lesions in vertebral bodies (**c**). Individual segmentation masks for single vertebral bodies and intervertebral discs are outlined in red.

coordinate transformations to patient space. Once resampled, additional verification was performed to ensure all segmentation masks are accurately overlaid over all available sequences per patient. The verified data were subsequently uploaded to the database.

Examples of segmented lumbar vertebral bodies and intervertebral discs are shown in Figs. 1–4. In detail, Fig. 1 shows two scans acquired on scanners from differing vendors (Philips and Siemens) per each available sequence. Similar is shown in Fig. 2 for another vendor (Philips and GE). Furthermore, Fig. 3 shows the differences between scans of the same patient acquired from the same vendor's scanners (Philips), which differ in models. Figure 4 provides three exemplary patient cases with segmentations shown on T1-weighted imaging in presence of pathological findings that may render manual segmentations particularly demanding (e.g., vertebral fracture, spondylodiscitis, and metastatic bone lesions).

The manual segmentations of each vertebral body and intervertebral disc are available as separate binary masks, where pixels with an intensity value of 1 correspond to the tissue of interest, while pixels of an intensity value of 0 belong to the background. Each mask of each image volume was stored as a separate *.nii file. In total, each available volume is accompanied with the corresponding segmentations of lumbar vertebral bodies and intervertebral discs.

Data Records

The database is available online within the OSF Repository, including the different imaging datasets, segmentation files, as well as metadata comprising patient characteristics and details on the available pulse sequences and used MRI systems per patient (Identifier: https://doi.org/10.17605/OSF.IO/QX5RT)[52]. Furthermore, Python scripts to read and visualize the data by utilizing open-source image analysis libraries – SimpleITK (https://simpleitk.org/) and NiBabel (https://nipy.org/nibabel/) – are provided.

Non-contrast-enhanced and contrast-enhanced T1-weighted, T2-weighted, STIR, and T1-weighted fs images are stored as separate datasets for each patient per scanner model. In the case of the T2-weighted DIXON sequences, sagittal water, fat, and in-phase images are deposited as separate datasets for each patient, if available. The segmentation maps of each vertebra (L1 to L5) and intervertebral discs (L1/2 to L4/5) are stored as *.nii files. All imaging data, as well as the segmentation maps, are saved using the Neuroimaging Informatics Technology Initiative (NIfTI) format (https://nifti.nimh.nih.gov/). NIfTI files have several features, such as raw data saved in 3D, containing two affine coordinates to relate voxel to spatial index, as well as additional data such as key acquisition parameters, encoding directions, and grid spacing, saved as a part of the header. NIfTI files can be directly read in a number of programming environments, such as Python, Matlab, and R, as well as directly visualized using tools such as ImageJ (https://imagej.nih.gov/ij/), ITK-SNAP (http://www.itksnap.org/), FSL (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki), AFNI (https://afni.nimh.nih.gov/), and FreeSurfer (https://surfer.nmr.mgh.harvard.edu/).

Datasets of patients and corresponding segmentation masks are labeled with the same patient ID. Masks of each vertebra are labeled as L1 to L5, while the masks for each intervertebral disc are labeled as L1_2, L2_3, L3_4, and L4_5.

Technical Validation

Acquisition of MRI was medically indicated and image quality control was performed during and/or immediately after completion of the exam by the technologist and radiologist in charge. Furthermore, inclusion of image data in the current study was performed by a board-certified radiologist with eleven years of experience based on a sufficient image quality.

Code availability

The database is available online within the OSF Repository, including the different imaging datasets, segmentation files, as well as metadata comprising patient characteristics and details on the available pulse sequences and used MRI systems per patient (Identifier: https://doi.org/10.17605/OSF.IO/QX5RT)[52]. We additionally provide Python scripts to read and visualize the data by utilizing open-source image analysis libraries – SimpleITK (https://simpleitk.org/) and NiBabel (https://nipy.org/nibabel/), available within the repository wiki page.

Received: 11 September 2021; Accepted: 3 March 2022;
Published online: 23 March 2022

References

1. Sheehan, N. J. Magnetic resonance imaging for low back pain: indications and limitations. *Postgrad Med J* **86**, 374–378, https://doi.org/10.1136/ard.2009.110973 (2010).
2. Chou, D. *et al.* Degenerative magnetic resonance imaging changes in patients with chronic low back pain: a systematic review. *Spine* **36**, S43–53, https://doi.org/10.1097/BRS.0b013e31822ef700 (2011).
3. Wassenaar, M. *et al.* Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic systematic review. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* **21**, 220–227, https://doi.org/10.1007/s00586-011-2019-8 (2012).
4. Roudsari, B. & Jarvik, J. G. Lumbar spine MRI for low back pain: indications and yield. *AJR. American journal of roentgenology* **195**, 550–559, https://doi.org/10.2214/AJR.10.4367 (2010).
5. Gawel, D., Glowka, P., Kotwicki, T. & Nowak, M. Automatic Spine Tissue Segmentation from MRI Data Based on Cascade of Boosted Classifiers and Active Appearance Model. *Biomed Res Int* **2018**, 7952946, https://doi.org/10.1155/2018/7952946 (2018).
6. Hille, G., Saalfeld, S., Serowy, S. & Tonnies, K. Vertebral body segmentation in wide range clinical routine spine MRI data. *Comput Methods Programs Biomed* **155**, 93–99, https://doi.org/10.1016/j.cmpb.2017.12.013 (2018).
7. Liu, D., Zucherman, M. & Tulloss, W. B. Jr. Six characteristics of effective structured reporting and the inevitable integration with speech recognition. *J Digit Imaging* **19**, 98–104, https://doi.org/10.1007/s10278-005-8734-0 (2006).
8. Kahn, C. E. Jr. *et al.* Toward best practices in radiology reporting. *Radiology* **252**, 852–856, https://doi.org/10.1148/radiol.2523081992 (2009).
9. European Society of Radiology. ESR paper on structured reporting in radiology. *Insights Imaging* **9**, 1–7, https://doi.org/10.1007/s13244-017-0588-8 (2018).
10. Pfirrmann, C. W., Metzdorf, A., Zanetti, M., Hodler, J. & Boos, N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine* **26**, 1873–1878 (2001).
11. Hwang, D. *et al.* Quantitative magnetic resonance imaging of the lumbar intervertebral discs. *Quant Imaging Med Surg* **6**, 744–755, https://doi.org/10.21037/qims.2016.12.09 (2016).
12. Mwale, F., Iatridis, J. C. & Antoniou, J. Quantitative MRI as a diagnostic tool of intervertebral disc matrix composition and integrity. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* **17**(Suppl 4), 432–440, https://doi.org/10.1007/s00586-008-0744-4 (2008).
13. Lotz, J. C. *et al.* New treatments and imaging strategies in degenerative disease of the intervertebral disks. *Radiology* **264**, 6–19, https://doi.org/10.1148/radiol.12110339 (2012).
14. Wehrli, F. W., Song, H. K., Saha, P. K. & Wright, A. C. Quantitative MRI for the assessment of bone structure and function. *NMR in biomedicine* **19**, 731–764, https://doi.org/10.1002/nbm.1066 (2006).
15. Sollmann, N. *et al.* MRI-Based Quantitative Osteoporosis Imaging at the Spine and Femur. *Journal of magnetic resonance imaging: JMRI* https://doi.org/10.1002/jmri.27260 (2020).
16. Stern, D., Likar, B., Pernus, F. & Vrtovec, T. Parametric modelling and segmentation of vertebral bodies in 3D CT and MR spine images. *Phys Med Biol* **56**, 7505–7522, https://doi.org/10.1088/0031-9155/56/23/011 (2011).
17. Litjens, G. *et al.* A survey on deep learning in medical image analysis. *Med Image Anal* **42**, 60–88, https://doi.org/10.1016/j.media.2017.07.005 (2017).
18. Hosny, A., Parmar, C., Quackenbush, J., Schwartz, L. H. & Aerts, H. Artificial intelligence in radiology. *Nat Rev Cancer* **18**, 500–510, https://doi.org/10.1038/s41568-018-0016-5 (2018).
19. Lundervold, A. S. & Lundervold, A. An overview of deep learning in medical imaging focusing on MRI. *Zeitschrift fur medizinische Physik* **29**, 102–127, https://doi.org/10.1016/j.zemedi.2018.11.002 (2019).
20. Ching, T. *et al.* Opportunities and obstacles for deep learning in biology and medicine. *J R Soc Interface* **15**, https://doi.org/10.1098/rsif.2017.0387 (2018).
21. Forsberg, D., Sjoblom, E. & Sunshine, J. L. Detection and Labeling of Vertebrae in MR Images Using Deep Learning with Clinical Annotations as Training Data. *J Digit Imaging* **30**, 406–412, https://doi.org/10.1007/s10278-017-9945-x (2017).
22. Gaonkar, B. *et al.* Multi-Parameter Ensemble Learning for Automated Vertebral Body Segmentation in Heterogeneously Acquired Clinical MR Images. *IEEE J Transl Eng Health Med* **5**, 1800412, https://doi.org/10.1109/JTEHM.2017.2717982 (2017).
23. Korez, R., Likar, B., Pernuš, F. & Vrtovec, T. Model-Based Segmentation of Vertebral Bodies from MR Images with 3D CNNs. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 433–441, https://doi.org/10.1007/978-3-319-46723-8_50 (2016).
24. Chu, C. *et al.* Fully Automatic Localization and Segmentation of 3D Vertebral Bodies from CT/MR Images via a Learning-Based Method. *PloS one* **10**, e0143327, https://doi.org/10.1371/journal.pone.0143327 (2015).
25. Al-Kafri, A. S. *et al.* Boundary Delineation of MRI Images for Lumbar Spinal Stenosis Detection Through Semantic Segmentation Using Deep Neural Networks. *IEEE Access* **7**, 43487–43501, https://doi.org/10.1109/ACCESS.2019.2908002 (2019).
26. Huang, J. *et al.* Spine Explorer: a deep learning based fully automated program for efficient and reliable quantifications of the vertebrae and discs on sagittal lumbar spine MR images. *Spine J* **20**, 590–599, https://doi.org/10.1016/j.spinee.2019.11.010 (2020).
27. Zhou, J. *et al.* Automatic Vertebral Body Segmentation Based on Deep Learning of Dixon Images for Bone Marrow Fat Fraction Quantification. *Front Endocrinol (Lausanne)* **11**, 612, https://doi.org/10.3389/fendo.2020.00612 (2020).
28. Burns, J. E., Yao, J. & Summers, R. M. Artificial Intelligence in Musculoskeletal Imaging: A Paradigm Shift. *J Bone Miner Res* **35**, 28–35, https://doi.org/10.1002/jbmr.3849 (2020).
29. Yao, J., Burns, J. E., Munoz, H. & Summers, R. M. Detection of vertebral body fractures based on cortical shell unwrapping. *Med Image Comput Comput Assist Interv* **15**, 509–516, https://doi.org/10.1007/978-3-642-33454-2_63 (2012).
30. Knez, D., Likar, B., Pernus, F. & Vrtovec, T. Computer-Assisted Screw Size and Insertion Trajectory Planning for Pedicle Screw Placement Surgery. *IEEE Trans Med Imaging* **35**, 1420–1430, https://doi.org/10.1109/TMI.2016.2514530 (2016).
31. Hoy, D., Brooks, P., Blyth, F. & Buchbinder, R. The Epidemiology of low back pain. *Best Pract Res Clin Rheumatol* **24**, 769–781, https://doi.org/10.1016/j.berh.2010.10.002 (2010).

32. Ellingson, A. M., Shaw, M. N., Giambini, H. & An, K. N. Comparative role of disc degeneration and ligament failure on functional mechanics of the lumbar spine. *Comput Methods Biomech Biomed Engin* **19**, 1009–1018, https://doi.org/10.1080/10255842.2015.1088524 (2016).
33. Hartvigsen, J. *et al*. What low back pain is and why we need to pay attention. *Lancet* **391**, 2356–2367, https://doi.org/10.1016/S0140-6736(18)30480-X (2018).
34. Brinjikji, W. *et al*. MRI Findings of Disc Degeneration are More Prevalent in Adults with Low Back Pain than in Asymptomatic Controls: A Systematic Review and Meta-Analysis. *AJNR. American journal of neuroradiology* **36**, 2394–2399, https://doi.org/10.3174/ajnr.A4498 (2015).
35. An, H. S. *et al*. Introduction: disc degeneration: summary. *Spine* **29**, 2677–2678, https://doi.org/10.1097/01.brs.0000147573.88916.c6 (2004).
36. Dudli, S., Fields, A. J., Samartzis, D., Karppinen, J. & Lotz, J. C. Pathobiology of Modic changes. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* **25**, 3723–3734, https://doi.org/10.1007/s00586-016-4459-7 (2016).
37. Fields, A. J., Ballatori, A., Liebenberg, E. C. & Lotz, J. C. Contribution of the endplates to disc degeneration. *Curr Mol Biol Rep* **4**, 151–160, https://doi.org/10.1007/s40610-018-0105-y (2018).
38. Panagopoulos, J. *et al*. Prospective Comparison of Changes in Lumbar Spine MRI Findings over Time between Individuals with Acute Low Back Pain and Controls: An Exploratory Study. *AJNR. American journal of neuroradiology* **38**, 1826–1832, https://doi.org/10.3174/ajnr.A5357 (2017).
39. Clark, S. & Horton, R. Low back pain: a major global challenge. *Lancet* **391**, 2302, https://doi.org/10.1016/S0140-6736(18)30725-6 (2018).
40. Buchbinder, R. *et al*. Low back pain: a call for action. *Lancet* **391**, 2384–2388, https://doi.org/10.1016/S0140-6736(18)30488-4 (2018).
41. Andersson, G. B. Epidemiological features of chronic low-back pain. *Lancet* **354**, 581–585, https://doi.org/10.1016/S0140-6736(99)01312-4 (1999).
42. Fields, A. J. *et al*. Measurement of vertebral endplate bone marrow lesion (Modic change) composition with water-fat MRI and relationship to patient-reported outcome measures. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* https://doi.org/10.1007/s00586-021-06738-y (2021).
43. Fields, A. J. *et al*. Measuring and reporting of vertebral endplate bone marrow lesions as seen on MRI (Modic changes): recommendations from the ISSLS Degenerative Spinal Phenotypes Group. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* **28**, 2266–2274, https://doi.org/10.1007/s00586-019-06119-6 (2019).
44. Krug, R. *et al*. Associations between vertebral body fat fraction and intervertebral disc biochemical composition as assessed by quantitative MRI. *Journal of magnetic resonance imaging: JMRI* **50**, 1219–1226, https://doi.org/10.1002/jmri.26675 (2019).
45. Sollmann, N. *et al*. Associations Between Lumbar Vertebral Bone Marrow and Paraspinal Muscle Fat Compositions-An Investigation by Chemical Shift Encoding-Based Water-Fat MRI. *Front Endocrinol (Lausanne)* **9**, 563, https://doi.org/10.3389/fendo.2018.00563 (2018).
46. Dieckmeyer, M. *et al*. Vertebral Bone Marrow Heterogeneity Using Texture Analysis of Chemical Shift Encoding-Based MRI: Variations in Age, Sex, and Anatomical Location. *Front Endocrinol (Lausanne)* **11**, 555931, https://doi.org/10.3389/fendo.2020.555931 (2020).
47. Fields, A. J., Han, M., Krug, R. & Lotz, J. C. Cartilaginous end plates: Quantitative MR imaging with very short echo times-orientation dependence and correlation with biochemical composition. *Radiology* **274**, 482–489, https://doi.org/10.1148/radiol.14141082 (2015).
48. Wang, L. *et al*. Evaluation of human cartilage endplate composition using MRI: Spatial variation, association with adjacent disc degeneration, and *in vivo* repeatability. *J Orthop Res* **39**, 1470–1478, https://doi.org/10.1002/jor.24787 (2021).
49. Auerbach, J. D. *et al*. *In vivo* quantification of human lumbar disc degeneration using T(1rho)-weighted magnetic resonance imaging. *European spine journal: official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society* **15**(Suppl 3), S338–344, https://doi.org/10.1007/s00586-006-0083-2 (2006).
50. Yushkevich, P. A. *et al*. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage* **31**, 1116–1128, https://doi.org/10.1016/j.neuroimage.2006.01.015 (2006).
51. Yushkevich, P. A. *et al*. User-Guided Segmentation of Multi-modality Medical Imaging Datasets with ITK-SNAP. *Neuroinformatics* **17**, 83–102, https://doi.org/10.1007/s12021-018-9385-x (2019).
52. Al Khalil, Y. *et al*. Lumbar vertebral body and intervertebral disc segmentation in multi-scanner and multi-modal MRI – a ground truth database. *OSF Repository* https://osf.io/qx5rt/?view_only=d6dc0ac92f0e4ecb88d6de0cdfea6c47, https://doi.org/10.17605/OSF.IO/QX5RT (2021).

Acknowledgements

The present work was supported by the European Research Council (grant agreement No 677661 – ProFatMRI and grant agreement No 637164 – iBack), the German Society of Musculoskeletal Radiology (Deutsche Gesellschaft für Muskuloskelettale Radiologie, DGMSR), the German Research Foundation (Deutsche Forschungsgemeinschaft, project 432290010), and the OpenGTN project, part of the Marie Curie Innovative Training Network (ITN) fellowship program (project 236 764465).

Author contributions

Y.K.: data post-processing, drafting of manuscript. E.B.: data post-processing, critical revision of manuscript. J.S.K.: data acquisition, critical revision of manuscript. D.C.K.: data acquisition, critical revision of manuscript. M.B.: concept of study design, critical revision of manuscript. T.B.: concept of study design, data acquisition, data post-processing, critical revision of manuscript. N.S.: concept of study design, data acquisition, data post-processing, critical revision of manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to N.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.