

# How Deep Learning Reconstruction Impacts Radiomics Features in Abdominal MRI

Jingyu Zhong, M.D.<sup>1,2</sup>; Yang Song, Ph.D.<sup>3</sup>; Minda Lu<sup>4</sup>; Dominik Nickel, Ph.D.<sup>5</sup>

<sup>1</sup> Department of Imaging, Tongren Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

<sup>2</sup> Shanghai Key Laboratory of Flexible Medical Robotics, Tongren Hospital, Institute of Medical Robotics, Shanghai Jiao Tong University, Shanghai, China

<sup>3</sup> MR Research Collaboration Team, Siemens Healthineers, Shanghai, China

<sup>4</sup> MR Application, Siemens Healthineers, Shanghai, China

<sup>5</sup> Research & Clinical Translation, Magnetic Resonance, Siemens Healthineers, Erlangen, Germany

## Why did we conduct the research?

Magnetic resonance imaging (MRI), especially T2-weighted imaging with fat suppression (T2 FS) or without (T2WI), plays a pivotal role in abdominal imaging [1]. However, conventional MRI protocols have the inherent limitation of prolonged acquisition time, which may lead to motion artifacts [2, 3]. One solution is the half-Fourier acquisition single-shot turbo spin echo (HASTE) technique, which shortens the acquisition time. However, this comes at the cost of reduced signal-to-noise ratio, reduced image contrast, and reduced sharpness [4, 5]. To address this drawback, HASTE with deep learning reconstruction (DLR)<sup>1</sup> was introduced to accelerate the acquisition and shorten echo trains for improved T2 weighting [6, 7].

Radiomics, which extracts quantitative features from medical images for clinical decision-making, has gained in popularity [8, 9]. Yet it faces challenges caused by its sensitivity to acquisition parameters, reconstruction algorithms, and post-processing techniques [10–12]. DLR enables acceptable image quality and faster acquisitions, and it is therefore expected to be widely adopted in the future. However, the influence of DLR and the combined effects of DLR and accelerated acquisition on radiomics features have not yet been investigated. It is necessary to evaluate in advance the potential changes that the new technique could cause to radiomics features.

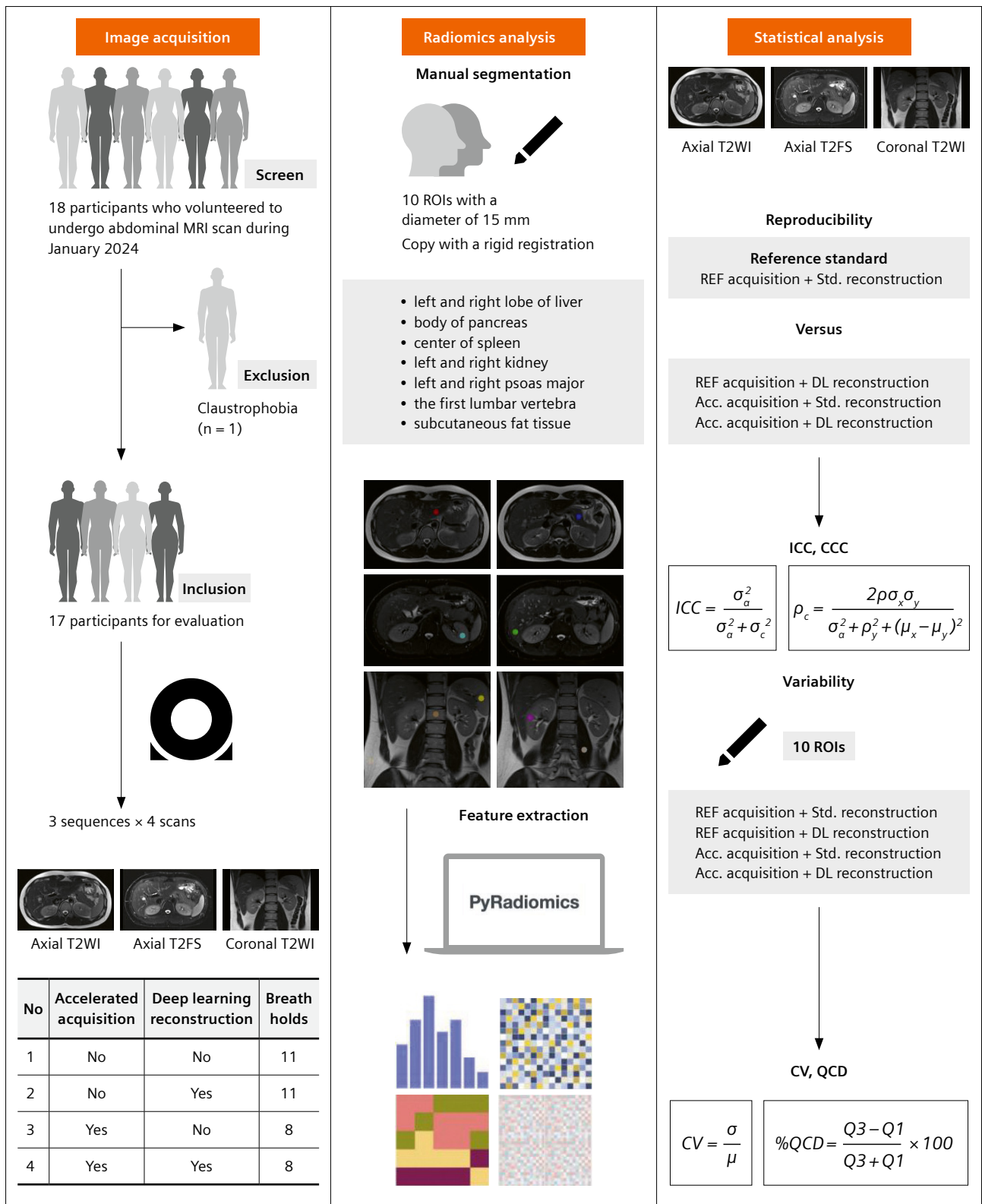
Our study provides a basis for future research into establishing radiomics models using images from an accelerated HASTE sequence with DLR. We investigated the impact of DLR and accelerated acquisitions on the robustness of radiomics features in abdominal scans acquired using a HASTE sequence.

## How did we conduct the research?

We performed abdominal MRI scans on volunteers with no history of abdominal surgery or malignancies (Fig. 1). We used a 3T MRI system (MAGNETOM Vida, Siemens Healthineers, Erlangen, Germany) with an 18-channel body matrix coil and the integrated spine array. We ran the research HASTE sequence<sup>1</sup> three times with different orientations and contrasts: axial T2W, axial T2 FS, and coronal T2W. Each protocol was acquired and reconstructed four times: (1) clinical reference acquisition with standard reconstruction (REF + STD), (2) clinical reference acquisition with deep learning reconstruction (REF + DLR), (3) accelerated acquisition with standard reconstruction (ACC + STD), and (4) accelerated acquisition with deep learning reconstruction (ACC + DLR) (Table 1). The clinical reference acquisition with standard reconstruction (REF+STD) served as the reference standard for the evaluation. The impact of accelerated acquisition and DLR were then assessed by quantifying how these two factors changed radiomics features.

To extract the radiomics features, the regions of interest (ROIs) were drawn for 10 anatomical sites with rigid registrations. The ROIs, which had a fixed diameter of 15 pixels (approximately 15 mm), were placed on 10 anatomical sites: the left and right lobes of the liver, the body of the pancreas, the center of the spleen, the left and right kidneys, the left and right psoas major muscles, the first lumbar vertebra, and the subcutaneous fat tissue. The ROIs were selected to cover as much as possible of the parenchyma of the organ or tissue, avoiding contact with the vessels, ducts, and lesions. Ninety-three radiomics features, including histogram-based and texture features from the original image, were extracted using PyRadiomics after z-score normalization. All these features were extracted in line with the Image Biomarker Standardization Initiative (IBSI).

<sup>1</sup> Work in progress. The application is currently under development and is not for sale in the U.S. and in other countries. Its future availability cannot be ensured.

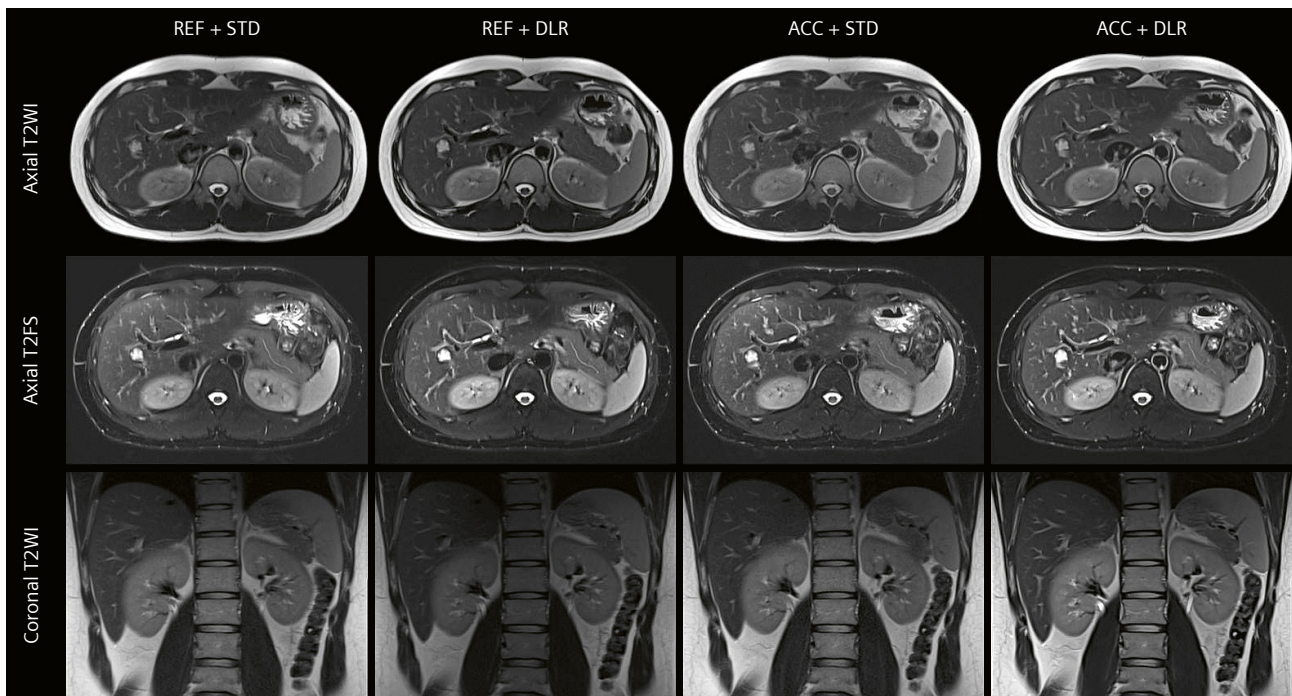


**1 Study workflow**

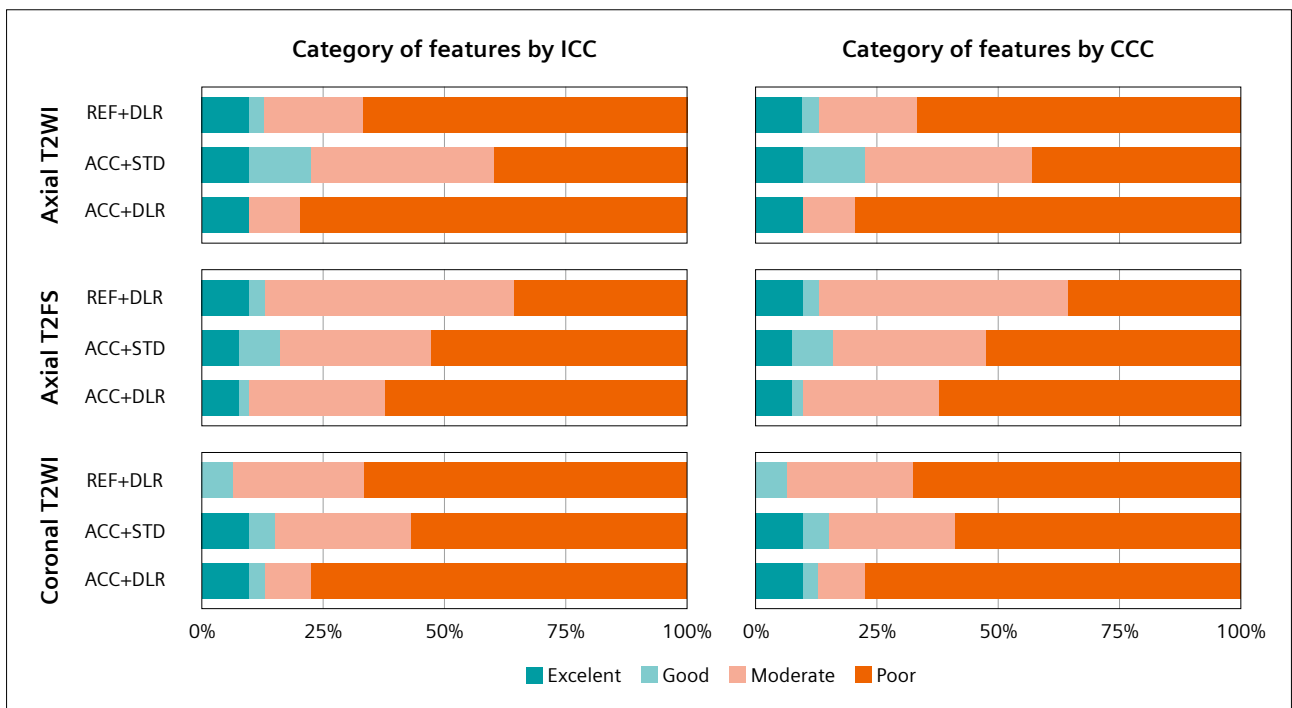
The study had three steps: image acquisition, radiomics analysis, and statistical analysis.

REF = original; Std. = standard; Acc. = accelerated; DL = deep learning; ICC = intraclass correlation coefficient;

CCC = concordance correlation coefficient; CV = coefficient of variation; QCD = quartile coefficient of dispersion



**2** The representative case was four scans of three sequences from a 29-year-old male participant with a height of 1.86 m, a weight of 80 kg, and a body mass index of 23.1 kg/m<sup>2</sup>. He had no history of abdominal surgery or cancer, and no current acute abdominal injury or disease. The conventional abdominal MRI scan identified a hepatic cystic lesion, but it was situated outside the region of interest.



**3** Reproducibility of radiomics features according to intraclass correlation coefficient (ICC) and concordance correlation coefficient (CCC) values.

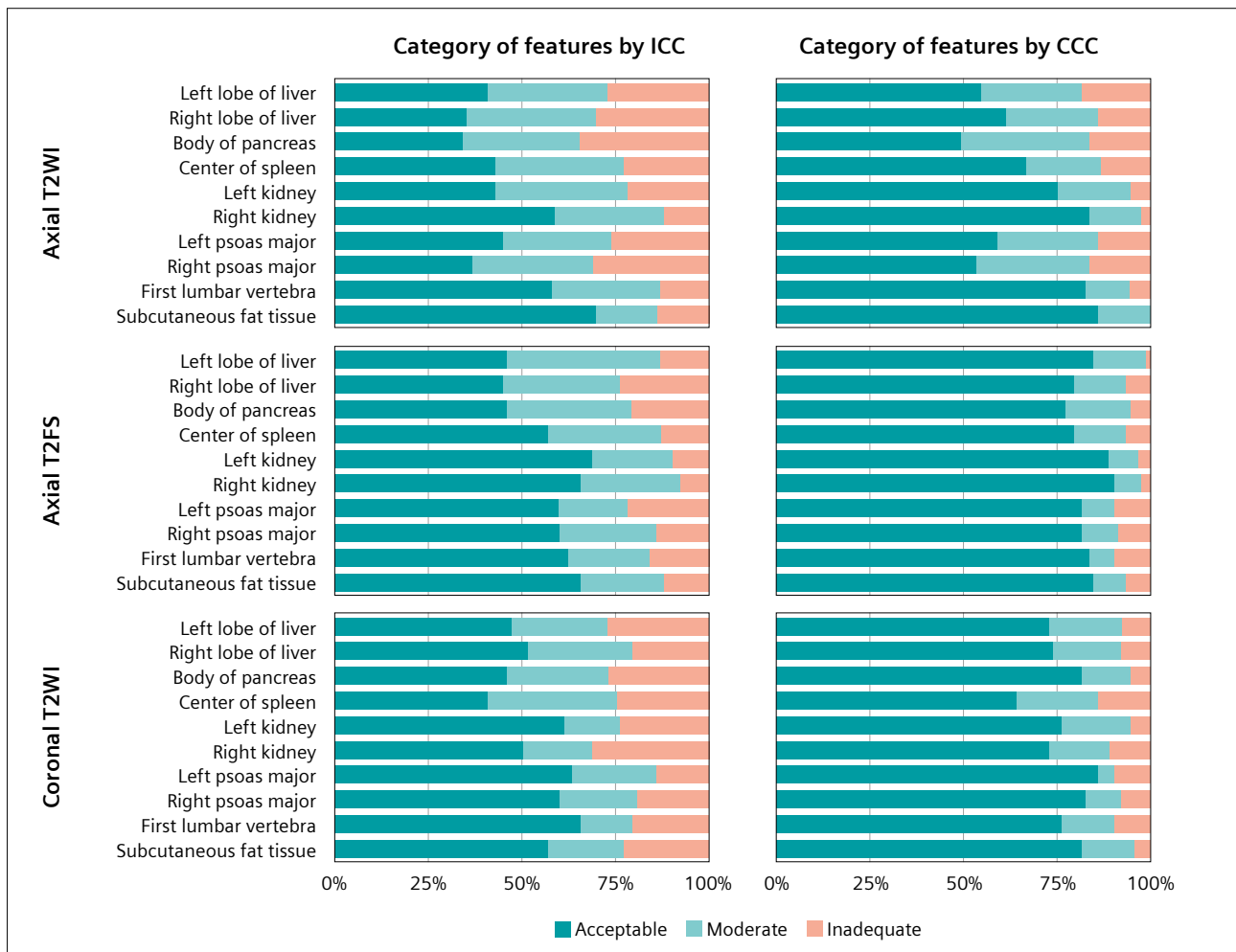
The reproducibility was evaluated with the intraclass correlation coefficient (ICC) and the concordance correlation coefficient (CCC), using REF+STD as the reference. The variability between four scans was assessed by coefficient of variation (CV) and quartile coefficient of dispersion (QCD).

### How might deep learning reconstruction impact radiomics features in abdominal MRI?

One author of this study (JZ) also participated in the study and served as a representative case. The REF+STD is used as the reference standard for the image quality and showed good signal-to-noise ratio, reduced image contrast, and reduced sharpness for assessment. The hepatic cystic lesion was clearly visible on the REF+STD images. The REF+DLR provided even better contrast and sharpness for the lesion characterization. However, the ACC+STD did not show satisfactory image quality due to the increased image noise, blurry edges of the lesion, and suboptimal

contrast. The image quality of ACC+DLR was greatly improved thanks to DLR, even for the accelerated acquisition protocol. Therefore, we suspected that both the accelerated acquisition protocol and DLR could change the radiomics features. However, for both, the changed radiomics features might be recovered.

Unfortunately, the radiomics features were not recovered. Compared to the REF+STD images, the REF+DLR, ACC+STD, and ACC+DLR images showed substantial change in the radiomics features. Our study found that the median (first and third quartile) of overall ICC and CCC values were 0.451 (0.305, 0.583) and 0.450 (0.304, 0.582). The overall percentage of radiomics features with ICC > 0.90 and CCC > 0.90 was 8.1% and 8.1%, which was considered acceptable (Fig. 3). The median (first and third quartile) of overall CV and QCD values was 9.4% (4.9%, 17.2%) and 4.9% (2.5%, 9.7%). The overall percentage of radiomics features with CV < 10% and QCD < 10% was 51.9% and 75.0%, which was considered acceptable (Fig. 4).



4 Variability of radiomics features according to coefficient of variation (CV) and quartile coefficient of dispersion (QCD) values.

Our study showed that the radiomics features were significantly influenced by the accelerated acquisition and DLR. The reproducibility of radiomics features was low between the standard and accelerated acquisitions, and between the standard and DLR acquisitions. Although the REF+STD and ACC+DLR images remained unchanged to the naked eye, the in-depth radiomics features did change. The brittle changes might be too small to be detected by the naked eye.

### **What should we do before introducing deep learning reconstruction into radiomics analysis?**

We expected that DLR would allow faster acquisition with higher image quality and comparable diagnostic performance. The DLR algorithm used in our study reconstructed images using a variational network to generate images with higher signal-to-noise ratio and thereby allow faster acquisition with higher parallel imaging acceleration. The algorithm enhanced the image contrast and reduced the image noise for diagnostic purposes. However, it was not designed to keep the values of the radiomics features stable. We hypothesized that the increased image noise caused by accelerated acquisition and the reduced image noise from DLR might offset each other and make the radiomics features more reproducible. However, the combination of these two factors did not result in stable values of the radiomics features. Although the subjective image-quality and diagnostic assessments favored the accelerated acquisition with DLR, the radiomics analysis for different acquisition protocols and reconstruction algorithms was not interchangeable.

To reduce the impact on the generalizability of radiomics analysis, we need to try something different to harmonize DLR and non-DLR images. First, standardize protocols: Radiomics analysis relies on consistent image acquisition to ensure feature reproducibility. Our study revealed that protocol variations, even minor differences in repetition time, echo time, and parallel imaging factors, can cause severe variations in texture features like the gray-level co-occurrence matrix (GLCM) contrast. To mitigate this, we suggest a standardization framework in at least a multi-scanner or multi-center study. This approach

can harmonize data at the very beginning without concerns about changing the in-depth information due to the later pre-processing steps. However, it cannot be used here, because the DLR itself is a game-changer in MRI scanning.

Second, choose appropriate image pre-processing before performing radiomics analysis: It is hard to use a standardized protocol across multiple centers. Even with standardized protocols, image pre-processing is critical to address residual variability. Our workflow only used z-score normalization, but other workflows may integrate bias field correction, anisotropic diffusion filtering, and so on. These steps reduce intensity inhomogeneity and noise, and they may change the information in the image. It is necessary to assess their influence as well. By treating pre-processing as an integral step rather than an afterthought, radiomics models gain resilience to acquisition-specific noise patterns, ensuring features reflect true biological heterogeneity rather than technical variability.

Third, use only radiomics features with higher reproducibility and less variability: Feature selection bridges raw data and actionable biomarkers. Ideally, we would only use features that have acceptable reproducibility and variability to establish the radiomics models. We did find that some features are robust against many factors. Further, if these features are biologically relevant to lesion properties, corresponding radiomics models will achieve greater generalizability across diverse clinical settings. However, the question is whether these features are informative enough to establish a radiomics model. We need to balance the robustness of features and the diagnostic performance of models. Prioritizing stability and biological relevance will make radiomics models more generalizable across diverse clinical settings.

### **Conclusion**

Deep learning reconstruction is a double-edged sword for radiomics. While it addresses critical bottlenecks in MRI workflows, its feature perturbations demand rigorous validation. By adhering to standardized protocols and investing in robustness testing, the radiology community can harness the full potential of DLR without compromising the diagnostic promise of radiomics.

## References

- 1 Klessen C, Asbach P, Kroencke TJ, Fischer T, Warmuth C, Stemmer A, et al. Magnetic resonance imaging of the upper abdomen using a free-breathing T2-weighted turbo spin echo sequence with navigator triggered prospective acquisition correction. *J Magn Reson Imaging*. 2005;21(5):576–82.
- 2 Nanko S, Oshima H, Watanabe T, Sasaki S, Hara M, Shibamoto Y. Usefulness of the application of the BLADE technique to reduce motion artifacts on navigation-triggered prospective acquisition correction (PACE) T2-weighted MRI (T2WI) of the liver. *J Magn Reson Imaging*. 2009;30(2):321–6.
- 3 Nakayama Y, Yamashita Y, Matsuno Y, Tang Y, Namimoto T, Kadota M, et al. Fast breath-hold T2-weighted MRI of the kidney by means of half-Fourier single-shot turbo spin echo: comparison with high resolution turbo spin echo sequence. *J Comput Assist Tomogr*. 2001;25(1):55–60.
- 4 Lee MG, Jeong YK, Kim JC, Kang EM, Kim PN, Auh YH, et al. Fast T2-weighted liver MR imaging: comparison among breath-hold turbo-spin-echo, HASTE, and inversion recovery (IR) HASTE sequences. *Abdom Imaging*. 2000;25(1):93–9.
- 5 Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans Med Imaging*. 2018;37(2):491–503.
- 6 Hammernik K, Klatzer T, Kobler E, Recht MP, Sodickson DK, Pock T, et al. Learning a variational network for reconstruction of accelerated MRI data. *Magn Reson Med*. 2018;79(6):3055–3071.
- 7 Herrmann J, Gassenmaier S, Nickel D, Arberet S, Afat S, Lingg A, et al. Diagnostic confidence and feasibility of a deep learning accelerated HASTE sequence of the abdomen in a single breath-hold. *Invest Radiol*. 2021;56(5):313–319.
- 8 Volpe S, Mastroleo F, Krengli M, Jereczek-Fossa BA. Quo vadis radiomics? Bibliometric analysis of 10-year Radiomics journey. *Eur Radiol*. 2023;33(10):6736–6745.
- 9 Kocak B, Baessler B, Cuocolo R, Mercaldo N, Pinto Dos Santos D. Trends and statistics of artificial intelligence and radiomics research in radiology, nuclear medicine, and medical imaging: bibliometric analysis. *Eur Radiol*. 2023;33(11):7542–7555.
- 10 Kocak B, Baessler B, Bakas S, Cuocolo R, Fedorov A, Maier-Hein L, et al. CheckList for Evaluation of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMI. *Insights Imaging*. 2023;14(1):75.
- 11 Kocak B, Akinci D'Antonoli T, Mercaldo N, Alberich-Bayarri A, Baessler B, Ambrosini I, et al. METHodological Radiomics Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMI. *Insights Imaging*. 2024;15(1):8.
- 12 Akinci D'Antonoli T, Cuocolo R, Baessler B, Pinto Dos Santos D. Towards reproducible radiomics research: introduction of a database for radiomics studies. *Eur Radiol*. 2024;34(1):436–443.




---

**Contact**

Jingyu Zhong, M.D., Ph.D.  
 Department of Imaging  
 Tongren Hospital  
 Shanghai Jiao Tong University  
 School of Medicine  
 No. 1111 Xianxia Road  
 Shanghai 200336  
 China  
 ZJY4623@shtrhospital.com

---