Reproducibility and Discriminative Power of MRI Liver Radiomics: Impact of Deep Learning-based Image Reconstruction

Louison Rey[a], Navid Rabbani[b,c], Benoit Chauveau[a], Constance Hordonneau[a], Bruno Pereira[d], Jean-Marc Garcier[a,c], Eleonore Pouget[a], Adrien Bartoli[b,c], Benoit Magnin[a,b,c]*

1 Radiology Department, Clermont-Ferrand University Hospital, Clermont-Ferrand, France

2 DIA2M, DRCI, Clermont-Ferrand University Hospital, Clermont-Ferrand, France

3 Institut Pascal, UMR 6602 CNRS, Université Clermont Auvergne, Clermont-Ferrand, France

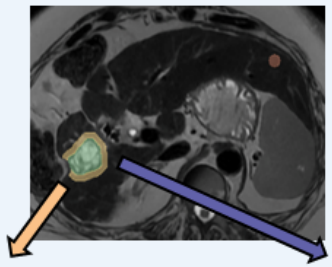
4 Biostatistics Department, DRCI, Clermont-Ferrand University Hospital, Clermont-Ferrand, France

* Corresponding author

Graphical Abstract

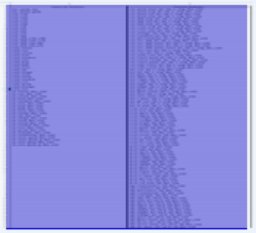
Reproducibility and Discriminative Power of MRI Liver Radiomics: Impact of Deep Learning-based Image Reconstruction

Segmentation and radiomics extraction



x 2

Non-DLIR radiomics

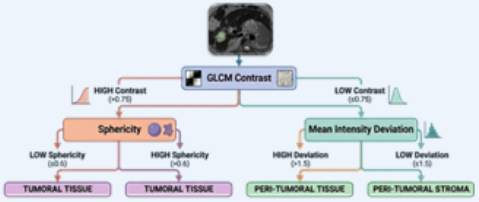


DLIR radiomics

Comparison of DLIR and non-DLIR radiomics

Reproducibility
(inter and intra observer)

Discriminative power
(Random Forest classifier)



No difference between DLIR and non-DLIR in radiomic feature **reproducibility** in liver lesion MRI

Radiomic **discriminative power** for tissue characterization is preserved following DLIR

CONCLUSION

DLIR does not alter radiomic feature reproducibility or discriminative power in liver MRI

Abstract

Objectives:

While Deep Learning-based Image Reconstruction (DLIR) is increasingly used to improve MRI quality, its influence on radiomic stability remains underexplored. This study aimed to assess the impact of DLIR on the reproducibility and discriminative power of radiomic features in liver MRI.

Material and Methods:

In this retrospective study, 95 patients with malignant liver lesions underwent 1.5T MRI (T2, T2FS, and Diffusion). Images were reconstructed using both traditional (non-DLIR) and DLIR (AIR Recon DL) methods. Three regions of interest (ROI) - tumor, peritumoral, and healthy liver - were segmented by two readers. Intra- and inter-observer reproducibility were evaluated using ICC. Discriminative power was assessed via Random Forest classifiers and Wilcoxon tests to distinguish tissue types.

Results:

Generalized Linear Mixed Models showed no significant difference in the number of reproducible features between DLIR and non-DLIR for intra-observer (Odds Ratio (OR): 0.860, 95%CI [0.653; 1.135]) or inter-observer (OR: 1.098, 95%CI [0.673; 1.791]) analyses. Inter-observer reproducibility remained low across all reconstructions, particularly for intensity features. Classifier performance for tissue differentiation remained comparable between reconstructions. A slight decrease (2–7%) in the percentage of discriminative features was observed with DLIR, suggesting a mild smoothing effect on subtle textures.

Conclusion:

DLIR does not significantly improve or degrade the reproducibility of radiomics in liver MRI. However, DLIR preserves the essential signatures required for tissue discrimination. These results support the

clinical integration of DLIR in radiomic workflows, confirming that it can be used without compromising the reliability of quantitative imaging biomarkers.

Keywords

Radiomics Reproducibility; Deep Learning Reconstruction; MRI; Liver Lesion

List of Abbreviations

- ARDL: AIR™ Recon Deep Learning
- ASIR: Adaptive Statistical Iterative Reconstruction
- TACE: Transcatheter arterial chemoembolization
- HCC: Hepatocellular carcinoma
- DL: Deep Learning
- DLIR: Deep Learning Image Reconstruction
- FS: Fat Saturation
- GLCM: Gray Level Co-occurrence Matrix
- GLDZM: Gray Level Distance Zone Matrix
- GLMM: Generalized Linear Mixed Model
- GLRLM: Gray Level Run Length Matrix
- GLSZM: Gray Level Size Zone Matrix
- IBSI: Image Biomarker Standardization Initiative
- ICC: Intraclass Correlation Coefficient

- MRI: Magnetic Resonance Imaging
- MIRP: Medical Image Radiomics Processor
- NGTDM: Neighborhood Gray Tone Difference Matrix
- NGLDM: Neighboring Gray Level Dependence Matrix
- OR: Odds Ratio
- ROI: Region of Interest
- T: Tesla
- CT: Computed Tomography
- PET: Positron Emission Tomography
- NET: Neuroendocrine Tumor

Introduction

Radiomics consists of extracting quantitative information, known as radiomic features, from medical images [1]. Used alone or combined with clinical, biological, or genetic data, these features can aid in diagnosis, evaluate therapeutic response, or predict prognosis [2]. Numerous studies have developed predictive models tailored to specific diseases using radiomic features derived from training cohorts [3–7].

However, despite a growing body of literature, radiomics still struggles to find widespread clinical application, mainly due to the challenge of reproducing results [8]. The reasons for this lack of reproducibility have been well documented [8,9]. Several methodological recommendations have emerged to address this limitation, notably the Image Biomarker Standardization Initiative (IBSI), which plays a central role [10,11].

Methodological studies have highlighted various factors that can affect the stability of radiomic features, particularly in CT imaging [12], PET imaging, and to a lesser extent, MRI. Indeed, it was shown that the extraction of radiomic features in MRI may be influenced by scanner-related factors, such as the manufacturer [13]. Moreover, acquisition parameters like spatial resolution or echo time also affect the extracted feature values [14,15]. Post-processing effects, such as smoothing or alteration of textures [16], and segmentation of regions of interest [17], further contribute to variability in radiomic features. These observations underscore the importance of standardizing every step in the radiomic workflow, from image acquisition to data analysis, to enhance the reliability and reproducibility of results.

DL-based image reconstruction methods are now available in both CT and MRI. These methods have demonstrated advantages over traditional reconstruction techniques, particularly in noise reduction [18] and diagnostic performance [19], and are now commonly used in clinical practice. The impact of DLIR on radiomic features has been studied in CT imaging. As expected, radiomic features obtained after DLIR differ from those from traditional iterative reconstructions [20]. The influence of DLIR on the stability of radiomic features has also been explored. A study using an anthropomorphic phantom showed improved reproducibility of radiomic features with DLIR [21]. Zhong et al. [20] partially addressed this issue in liver lesion CT scans, showing higher repeatability indices with DLIR compared to ASIR-V. In contrast, the impact of DLIR on radiomic features in MRI remains underexplored. A few studies [22–24] showed that radiomic features differ from those obtained with traditional non-DLIR, but the effect on reproducibility has never been studied in MRI.

We thus aimed to assess the impact of DLIR on the reproducibility of radiomic features, by comparing features from DLIR images with those from non-DLIR images. The primary objective of our study was to compare the intra- and inter-observer reproducibility of radiomic features according to reconstruction type on MRIs of malignant liver lesions. Secondary objectives were to compare the discriminative power of radiomic features between different tissue types (tumor, peritumoral area, and normal liver) across

reconstruction types, and to quantify the number of radiomic features differing between DLIR and non-DLIR.

Materials and Methods

Study Population

The study was approved by the local ethics committee (IRB00013412, “CHU de Clermont Ferrand IRB #1”, IRB number 2025-CF483) and complied with French data protection regulations.

We conducted a retrospective, single-center study at the Estaing University Hospital (CHU) in Clermont-Ferrand. All patients who underwent a liver MRI on the same scanner (Signa Artist; 1.5T, GE Healthcare™) between November 2022 and November 2023 were screened for inclusion.

Inclusion criteria were: patients who underwent a liver MRI in our institution on a Signa Artist scanner (1.5 T, GE Healthcare) including the three sequences of interest with both reconstructions, aged 18 years or older, and having at least one malignant hepatic lesion of primary or secondary origin (diagnosed histologically or radiologically) measuring more than 10 mm in its longest axis.

Exclusion criteria were a prior focal treatment on the lesion (surgery, thermal ablation, or radiation therapy), or the presence of artifacts on any of the sequences. Ultimately, 95 patients were included in our study (Figure 1). Population characteristics are described in Table 1.

A summary of the study design is shown in Figure 2.

Image Acquisition

The MRI protocol required three specific axial sequences: T2, T2 with fat saturation (T2FS), and diffusion-weighted imaging (Diffusion). Sequence parameters are provided in Table 2. Each sequence

was reconstructed both conventionally (non-DLIR) and using DLIR (AIR Recon DL, GE Healthcare™; vendor-specific algorithm implementation), yielding six image series per patient. Briefly, this algorithm includes a deep CNN trained on a database of paired low-noise and high-noise k-space data to suppress noise and truncation artifacts before image synthesis and takes raw k-space data as its input and generates high fidelity images as its output [25] and has no user-adjustable settings.

One lesion larger than 10 mm was randomly selected per patient. The image passing through the long axis of the lesion was extracted from each of the six series for analysis.

Segmentation

On each of the six series per patient, manual 2D segmentation was performed using 3D-Slicer software. Three different regions were segmented (Figure 3):

- Manual segmentation of the tumor lesion.
- Semi-automatic segmentation of the peritumoral region, defined as a 5 mm peripheral rim around the tumor ROI automatically generated with manual correction to remove non-parenchymal areas (e.g., vessels, extrahepatic structures).
- Manual segmentation of a circular ROI of healthy parenchyma (ROI), placed in the most anterior and left part of the liver on that slice, in normal appearing parenchyma, with a 10-mm diameter.

Segmentation was performed twice by Reader 1 (radiology resident) with a 3 month wash out period before the second segmentation and once by Reader 2 (senior radiologist with 13 years of abdominal imaging experience), resulting in three segmentations per sequence. Segmentation was performed with Slicer 5.6.1 (https://www.slicer.org/, [26]).

Radiomic Feature Extraction

Radiomic features were extracted using the MIRP (Medical Image Radiomics Processor) platform [27], which complies with the IBSI guidelines. The preprocessing settings included bias field inhomogeneity correction, intensity normalization by the standardization method, resampling to isotropic 1 mm^3 voxels, and discretization using 32 bins.

A total of 152 features were extracted: 57 intensity-based features (2 local intensity, 18 intensity-based statistical, 23 intensity histogram, 14 intensity-volume histogram) and 95 higher-order textural features (25 GLCM, 16 GLRLM, 16 GLSZM, 16 GLDZM, 5 NGTDM, 17 NGLDM) (see Supplementary material 1). Shape features were not included as they were deemed irrelevant to the study objective.

Statistical Analysis

To evaluate the impact of DLIR on radiomic feature stability, the number of reproducible features was compared between DLIR and non-DLIR across three anatomical regions (Liver, Tumor, Peri-tumor) and three MRI sequences (Diffusion, T2, T2FS). Reproducibility analysis was performed using ICC (2,1). A feature was considered reproducible if the lower bound of the 95% CI was greater than 0.75.

Statistical significance was assessed using a Generalized Linear Mixed Model (GLMM) with a binomial logit link function. The stability of a feature (stable vs. unstable) was treated as the binary dependent variable. Reconstruction method (DLIR vs. non-DLIR) was included as a fixed effect, while anatomical region and sequence type were treated as random effects to account for the inherent clustering and potential variability across different imaging protocols and tissues. Odds Ratios (OR) and 95% Confidence Intervals (CI) were calculated to quantify the likelihood of achieving feature stability with DLIR compared to the non-DLIR.

Differences in feature values between DLIR and non-DLIR were also evaluated using the Wilcoxon signed-rank test with Bonferroni correction.

Tissue Type Discriminative Power

The tissue type discriminative power of radiomic features was assessed by two complementary approaches: statistical separability and predictive performance.

For statistical separability analysis, the number of significantly different features between the three tissue types for a given sequence and reconstruction was assessed using the Kruskal-Wallis test. Pairwise differences between regions were evaluated using the Wilcoxon signed-rank test with Bonferroni correction for multiple comparisons.

For predictive performance, Random Forest classifiers were employed to differentiate ROI from different tissue types using numerical radiomic features for each MRI sequence and reconstruction. For each combination, the dataset was repeatedly partitioned into training (70%) and testing (30%) subsets across 30 bootstrap iterations, ensuring that each patient was included in only one subset per iteration. Classifier performance was evaluated using accuracy scores, and statistical comparisons between reconstruction methods were conducted with paired t-tests.

All analyses were performed using R (v4.5.2, R Core Team, Vienna, Austria, [28]), with the lme4 package. A p-value < 0.05 was considered statistically significant.

Results

Intra-observer Reproducibility

The main results for intra-observer radiomic feature reproducibility are shown in Table 3 (detailed results in Supplementary material 2).

No reproducibility difference was found between the reconstructions with the GLMM: Odds Ratio (effect of DLIR vs non-DLIR) 0.860, 95%CI [0.653; 1.135].

Intra-observer analysis, regardless of reconstruction method, showed overall excellent reproducibility of radiomic features within tumor lesions, moderate reproducibility in the peritumoral region (varying by sequence), and almost no reproducibility in the healthy liver parenchyma. When comparing reconstructions, reproducibility in healthy liver was similar. In the peritumoral region, DLIR showed a moderate decrease in reproducibility across all three sequences, particularly in T2FS (67.76% of features reproducible vs. 74.34% with non-DLIR). This decline was mainly due to variation in higher-order features, especially GLSZM and GLDZM.

For healthy liver, reproducibility remained very low across all sequences and reconstructions, preventing reliable analysis in this region.

Inter-observer Reproducibility

The main results for inter-observer radiomic feature reproducibility are presented in Table 4 (detailed results in Supplementary material 3).

No reproducibility difference was found between the reconstructions with the GLMM: Odds Ratio (effect of DLIR vs non-DLIR) 1.098, 95%CI [0.673; 1.791].

Overall inter-observer reproducibility was low across all regions. In tumor lesions, only 6 to 8.5% of features were reproducible, with slightly higher rates for DLIR. Intensity features showed particularly poor reproducibility, even lower than textural features.

In both the peritumoral region and healthy liver, reproducibility was nearly zero, and DLIR did not improve inter-observer agreement.

Tissue Type Discriminative Power: Statistical Separability

Main results concerning percentage of feature differences between tissues are in Table 5, with detailed data in Supplementary material 4.

Evaluation of statistical separability between tissues revealed a high percentage of feature different between tissue types, especially involving tumor tissue. Across all features, three-tissue discrimination was high, ranging from 83.55% (T2FS DLIR) to 90.70% (Diffusion non-DLIR). Comparisons between peritumoral tissue and healthy liver also showed strong discrimination (around 80%). Tumor vs. healthy liver analysis also showed good results, ranging from 73.68% (T2 DLIR) to 80.92% (T2 non-DLIR).

However, tumor vs. peritumoral comparisons showed much lower discrimination rates: 40.13% (T2 DLIR) and as low as 15.79% (T2FS DLIR), highlighting the difficulty in differentiating these two tissues, especially in T2FS sequences.

When comparing reconstructions, a general trend was observed: DLIR showed a slight decrease in discriminative features compared to non-DLIR, particularly in tumor-involved comparisons. Most notable differences were in tumor vs. healthy liver analysis, especially in T2 and diffusion: 80.92% (T2 non-DLIR) vs. 73.68% (T2 DLIR), and 76.32% (Diffusion non-DLIR) vs. 69.74% (Diffusion DLIR). This 2–7% drop suggests that DLIR maintains overall discriminatory capability but may obscure subtle textural differences captured by conventional sequences.

Tissue Type Discriminative Power: Predictive Performance

Classifiers were trained to differentiate ROI from different tissue types using numerical radiomic features for each MRI sequence and reconstruction to differentiate ROI from different tissue types using numerical radiomic features for each MRI sequence and reconstruction, using a bootstrap cross validation. The performances of classifiers are shown in Table 6. There was no difference of

classification performance between DLIR and non-DLIR reconstructions for each sequence when considering all features. When considering independently intensity and texture features, only T2 non-DLIR had a better classification performance than T2 DLIR.

Differences Between Reconstructions

The main results showing the number of features significantly different between reconstructions are shown in Table 7, with detailed results in Supplementary material 5.

Analysis showed that reconstruction type significantly impacted certain radiomic feature categories. For all sequences and ROI location, the feature difference was more pronounced on higher-order features than on intensity features (43.16% vs 3.51% for T2 tumor ROIs). A graphical example of two affected features is presented in Figure 4. This difference, however, varied according to ROI location, the difference being major for peri tumoral ROIs, and less marked for normal liver ROIs. In diffusion sequences, differences in global features were minimal across all regions (tumor: 1.32%, peritumoral: 7.89%, healthy liver: 0.66%), showing minimal differences induced by the reconstruction.

Discussion

This study explored the impact of DLIR MRI on the stability of radiomic features of malignant liver lesions, compared to conventional reconstructions. To our knowledge, this is the first study conducted in liver MRI to integrate intra- and inter-observer reproducibility, discriminative power, and quantitative variations in radiomic features depending on reconstruction methods.

Our results show that there was no overall difference in the number of reproducible features between DLIR and non-DLIR for both inter- and intra-observer studies. This constitutes the main finding of our study, demonstrating that the integration of DL-based denoising does not introduce additional instability into the radiomic pipeline. It effectively shifts the focus away from reconstruction-related

variability, identifying other causes such as manual segmentation as the persistent and primary obstacle to radiomic reproducibility in MRI.

For inter observer study, intensity features tended to show a slight improvement under DLIR in T2FS. These results are consistent with previous CT studies showing that DLIR can improve the stability of certain radiomic features [21]. However, in the peritumoral region, we observed a moderate decrease in reproducibility across all sequences (T2, T2FS, Diffusion), mainly affecting higher-order features (GLSZM, GLDZM), suggesting that DLIR may alter certain fine textural patterns [22]. In healthy liver tissue, the low reproducibility observed under all conditions aligns with findings in the literature, which highlight the sensitivity of radiomic measurements to low contrast and artifacts in homogeneous healthy areas [29].

A crucial point of our study is the contrast between the relative contributions of reconstruction variability and segmentation variability to radiomic instability. The choice of reconstruction algorithm (DLIR vs. non-DLIR) had a negligible impact on overall feature reproducibility, as evidenced by the lack of statistical significance in our GLMM analysis. Conversely, segmentation variability emerged as the dominant source of error in the radiomic pipeline. This is demonstrated by the dramatic drop in feature stability when moving from intra-observer to inter-observer analysis: within tumor lesions, features exhibited high intra-observer reproducibility, yet only 6.0% to 8.5% of these features remained reproducible across different readers, regardless of whether DLIR was applied. This difference in variance underscores that human manual segmentation introduces a critical bottleneck that eclipses the subtle technical variations induced by advanced DL reconstruction methods. Consequently, DLIR does not exacerbate nor mitigate the segmentation-driven noise; it merely operates on a baseline that remains heavily compromised by inter-operator boundary definitions. In our results, intra-observer reproducibility of radiomic features was generally high within the tumor, regardless of reconstruction type. In contrast, inter-observer reproducibility was low across all regions of interest, regardless of reconstruction type. These results confirm the central role of segmentation as a major limiting factor in

radiomics, especially in MRI, where contrast boundaries are often less distinct than in other imaging modalities [30], even when segmentation is performed by experienced operators [31]. DLIR does not appear to offset this variability. These findings support the development of semi-automatic or automatic segmentation solutions [32] and the establishment of a rigorous methodology to limit inter-operator variability.

Discriminative power analysis was mainly evaluated by two complementary approaches: statistical separability and predictive performance. The predictive performances of the classifiers to distinguish tissue types were similar between reconstructions. The findings indicate that DLIR tends to maintain the integrity of the primary radiomic signatures. Specifically, the discriminatory information used for tissue characterization does not suffer from the oversmoothing or blurring effects often associated with advanced denoising methods. On the other hand, the statistical separability analysis revealed that a high percentage of radiomic features were significantly different between tissue types (tumor, peritumoral regions and healthy liver) for both reconstructions. However, a slight decrease in the percentage of discriminative features was observed after DLIR compared to non-DLIR, particularly in comparisons involving tumor vs. healthy liver. This 2 to 7% drop suggests that DLIR maintains overall discriminatory capability but may obscure subtle textural differences captured by conventional sequences. We hypothesize that this reduction may reflect a smoothing or intensity normalization effect that reduces the detection of fine tumor heterogeneities [33]. As expected, discrimination between tumor and peritumoral zones remained more challenging, with especially poor performance in T2FS DLIR. These findings suggest that DL reconstructions, while reducing noise, may smooth out subtle image variations, potentially blurring fine details. This effect is particularly noticeable in peritumoral regions, which may contain important diagnostic and prognostic information [34].

The analysis of differences between reconstruction methods identified the most DLIR-sensitive families of features. Higher-order textural features were the most affected, especially outside the tumor, with up to 55% of textural features in the peritumoral region significantly different between both

reconstructions. To anchor the mathematical variations induced by the difference of reconstruction in a radiological context, specific higher-order features disrupted by DLIR in the peritumoral region warrant interpretation. For example, GLSZM - Small Zone Emphasis, which quantifies the prevalence of small, fine-grained zones of identical intensity, significantly decreased after DLIR (Figure 4). Factually, this indicates that the DL-based denoising process removes micro-textures and fine pixel-to-pixel variations, reducing perceived tissue micro-heterogeneity. Similarly, NGTDM - Busyness, which measures rapid spatial changes in intensity characteristic of highly textured areas or irregular borders, was significantly lower on DLIR images (Figure 4). This decrease reflects an intensity-smoothing effect that softens transition zones between tissue interfaces. While this smoothing reduces high-frequency technical noise and improves visual image quality for clinical reading, it might limit the capacity of radiomic models to detect subtle, high-frequency textural patterns critical for characterizing peritumoral infiltration. In contrast, intensity features were relatively unchanged. The diffusion sequence showed remarkable stability, with limited variation induced by DLIR. Such variability indicates that DLIR performance is inherently linked to the specific imaging sequence (and possibly anatomical site), likely reflecting the biases within the underlying training datasets. Consequently, these sequence-dependent variations in reconstruction can differentially affect the robustness and reproducibility of extracted radiomic features.

The generalizability of our findings to other commercial DLIR algorithms warrants careful consideration. Our quantitative results are specific to GE Healthcare's AIR Recon DL (1.5T), which operates as a deep CNN applied directly to complex k-space data, while other vendors employ distinct neural network architectures, training datasets, and target output optimization strategies [35,36]. For instance, Siemens Healthineers (Deep Resolve) combines k-space denoising with image-space super-resolution neural network [37], Philips Healthcare (SmartSpeed) uses a CNN designed to replace the iterative optimization process, and the traditional wavelet transform in compress sensing in a an undersampled k-space [38], and Canon Medical Systems (AiCE) utilizes DL noise-matching networks operating predominantly in the image domain [39]. Because k-space filtering directly alters high-frequency spatial

spectra prior to image generation, it selectively modifies fine textural relationships differently than post-reconstruction image-domain filters. Despite these proprietary differences, clinical DLIR engines share the fundamental objective of noise reduction and signal-to-noise ratio enhancement [40–42]. Consequently, the qualitative findings of our study, specifically that DLIR preserves baseline radiomic reproducibility while inducing a mild smoothing of micro-scale textural patterns, are expected to generalize as inherent properties of DL denoising. However, the exact quantitative feature shifts, such as specific numerical drops in GLSZM and NGTDM indices, do not directly translate to other algorithms. These quantitative variations remain dependent on proprietary network designs, training datasets, and user-selectable denoising strength levels (e.g., low, medium, or high). Therefore, while the broad methodological principles apply across platforms, feature-level numerical thresholds must be considered algorithm- and setting-specific.

Furthermore, these findings have significant implications for multicenter radiomics harmonization studies. As DLIR algorithms are increasingly adopted across different institutions and vendors, they introduce a new layer of technical heterogeneity across institutions. Because different vendors utilize distinct neural network architectures and noise-reduction thresholds, combining conventional and vendor-specific DL-reconstructed series acquired with varying parameters across multiple sites risks introducing systematic batch effects. In a multicenter context, these reconstruction-induced variations, particularly in higher-order textural features, can obscure subtle biological signals or lead to false-positive associations. To ensure the portability of radiomic signatures across centers, multicenter protocols should, as already mentioned, first establish standardized segmentation workflows to minimize observer-dependent variance. Building on a consistent segmentation baseline, studies must systematically implement feature-level post-processing harmonization methods, such as ComBat [43] or other feature normalization approaches [44,45], to adjust for reconstruction type alongside scanner model, acquisition parameters, and field strength [11].

These sequence- and region-dependent variations have implications for AI models trained on mixed reconstruction images. In real-world clinical repositories, datasets are frequently heterogeneous, combining historical conventional reconstructions with modern DLIR images. Our results reveal that up to 55% of higher-order textural features in the peritumoral zone significantly differ between DLIR and non-DLIR series. Consequently, if a radiomic AI model is trained on such a mixed dataset without specific constraints, the underlying machine learning algorithm may inadvertently leverage these reconstruction-driven technical variations as hand-crafted features. This creates a high risk of shortcut learning where the model classifies tissues based on the reconstruction type (e.g., detecting the smoothness of DLIR) rather than the true physiological or pathological heterogeneity of the tumor [46]. To mitigate this risk and ensure the generalizability of AI models, future workflows utilizing mixed pipelines must implement strict data-augmentation strategies (such as training models with explicit domain-adversarial regularization) or utilize style-transfer techniques (e.g., CycleGANs) to harmonize image textures prior to feature extraction [47].

Based on the significant variations we observed in higher-order textural features between reconstruction methods, we advocate for the implementation of standardized reconstruction reporting in all future radiomics literature. Currently, many radiomic studies meticulously report acquisition parameters (such as repetition time, echo time, and spatial resolution) but tend to omit the specifics of the reconstruction pipeline. Our data show that the introduction of DLIR introduces a distinct technical signature. Therefore, authors should systematically report: the type of reconstruction engine used (conventional or DL-based); the specific vendor and software version; and the exact reconstruction strength or denoising level applied (e.g., low, medium, high), as these settings directly dictate the degree of textural smoothing.

Based on those findings, we sum up four practical recommendations for future radiomics workflows incorporating DL reconstructions to ensure study reproducibility and comply with the evolving IBSI and EuSoMII guidelines: 1) Standardized reporting of reconstruction parameters: researchers should

systematically report the specific reconstruction algorithm, vendor, software version, and applied DL strength level (such as low, medium, or high) [48]. 2) Prioritizing automated segmentation: because inter-observer segmentation variability remains the dominant source of feature instability, future studies should implement validated automated or semi-automated segmentation pipelines to minimize reader-dependent noise before evaluating technical image parameters. 3) Harmonization: when datasets combine conventional and DL-reconstructed series, or images from multiple vendors, feature-level harmonization methods or generative style-transfer models should be applied prior to classifier training to prevent shortcut learning. 4) Feature selection and stability filtering: Analysts should exercise caution when utilizing higher-order textural features (specifically GLSZM and NGTDM indices) derived from DLIR images, as these metrics are sensitive to denoising-induced smoothing. Conducting explicit feature-stability filtering across reconstruction types is recommended prior to constructing predictive radiomic signatures.

Our study has several limitations. First, segmentations were performed in 2D on a single slice per sequence. Although 2D analysis reduces computational burden and manual segmentation time, 3D volumetric analysis would capture whole-tumor spatial heterogeneity more completely and avoid slice-selection bias. Second, our study is constrained by a single-field-strength (1.5T) and single-vendor design. While restricting our protocol to a single platform (Signa Artist, GE Healthcare™) was methodologically intentional to eliminate hardware-induced confounding factors, it restricts the immediate generalizability of our quantitative results. Imaging at 1.5T features a lower baseline signal-to-noise ratio (SNR) and different susceptibility artifact profiles compared to 3T systems. Because DLIR algorithms are intrinsically designed to suppress noise and optimize image sharpness, their impact on higher-order textural features might differ on 3T systems, where the baseline noise is lower and spatial resolution is often inherently higher. Furthermore, because different manufacturers implement distinct CNN architectures and proprietary training datasets, intensity normalization, and smoothing effects observed here are representative of GE's AIR Recon DL but may not directly translate to alternative commercial algorithms. Consequently, our specific percentage thresholds regarding feature

modifications should be interpreted as vendor- and field-strength-dependent, highlighting the need for future multi-vendor, multi-field-strength trials to validate these interactions across the broader clinical landscape. Third, the patient cohort was histologically heterogeneous, including both primary liver cancers (such as hepatocellular carcinoma) and secondary metastatic lesions, situated within varying underlying liver condition (cirrhotic versus non-cirrhotic liver). Tissue-specific baseline texture might influence the magnitude of DLIR smoothing differently across pathologies. Lastly, while we evaluated statistical separability and tissue classification, we did not evaluate downstream clinical prediction tasks, such as predicting microvascular invasion, treatment response, or overall survival. Further studies incorporating clinical endpoints are necessary to determine whether reconstruction-induced feature shifts alter actual predictive model performance. Nevertheless, this methodological study appears essential for improving our understanding of the factors influencing radiomics and for guiding the development of robust models.

In conclusion, our study demonstrates that Deep Learning-based Image Reconstruction does not fundamentally alter the overall reproducibility of radiomic features compared to conventional methods. Both intra- and inter-observer variability remain primarily driven by segmentation challenges rather than reconstruction algorithms. Notably, DLIR maintains the discriminative power necessary to distinguish between malignant lesions and healthy tissue, despite a slight attenuation of subtle textural features in peritumoral regions. These findings suggest that DLIR can be integrated into radiomic workflows without compromising diagnostic integrity, though researchers should remain cautious of its potential smoothing effects on high-order textural signatures. Ultimately, establishing standard operating procedures that include mandatory, detailed reporting of reconstruction algorithms and their specific strength settings will be an essential step toward achieving reproducible radiomics in clinical practice.

During the preparation of this work the authors used Gemini (Google) for language editing and text refinement to improve the clarity and flow of the manuscript. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

References

[1] P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G.P.M. van Stiphout, P. Granton, C.M.L. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J.W.L. Aerts, Radiomics: extracting more information from medical images using advanced feature analysis, Eur J Cancer 48 (2012) 441–446. https://doi.org/10.1016/j.ejca.2011.11.036.

[2] H.J.W.L. Aerts, E.R. Velazquez, R.T.H. Leijenaar, C. Parmar, P. Grossmann, S. Carvalho, J. Bussink, R. Monshouwer, B. Haibe-Kains, D. Rietveld, F. Hoebers, M.M. Rietbergen, C.R. Leemans, A. Dekker, J. Quackenbush, R.J. Gillies, P. Lambin, Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat Commun 5 (2014) 4006. https://doi.org/10.1038/ncomms5006.

[3] K. Deng, T. Chen, Z. Leng, F. Yang, T. Lu, J. Cao, W. Pan, Y. Zheng, Radiomics as a tool for prognostic prediction in transarterial chemoembolization for hepatocellular carcinoma: a systematic review and meta-analysis, Radiol Med 129 (2024) 1099–1117. https://doi.org/10.1007/s11547-024-01840-9.

[4] J. Choe, S.M. Lee, K.-H. Do, J.B. Lee, S.M. Lee, J.-G. Lee, J.B. Seo, Prognostic value of radiomic analysis of iodine overlay maps from dual-energy computed tomography in patients with resectable lung cancer, European Radiology (2018). https://doi.org/10.1007/s00330-018-5639-0.

[5] T. Chen, Z. Ning, L. Xu, X. Feng, S. Han, H.R. Roth, W. Xiong, X. Zhao, Y. Hu, H. Liu, J. Yu, Y. Zhang, Y. Li, Y. Xu, K. Mori, G. Li, Radiomics nomogram for predicting the malignant potential of gastrointestinal stromal tumours preoperatively, European Radiology (2018). https://doi.org/10.1007/s00330-018-5629-2.

[6] L. Dercle, B. Zhao, M. Gönen, C.S. Moskowitz, A. Firas, V. Beylergil, D.E. Connors, H. Yang, L. Lu, T. Fojo, R. Carvajal, S. Karovic, M.L. Maitland, G.V. Goldmacher, G.R. Oxnard, M.A. Postow, L.H. Schwartz, Early Readout on Overall Survival of Patients With Melanoma Treated With Immunotherapy Using a Novel Imaging Analysis, JAMA Oncol 8 (2022) 385. https://doi.org/10.1001/jamaoncol.2021.6818.

[7] R. Ortiz-Ramón, A. Larroza, S. Ruiz-España, E. Arana, D. Moratal, Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study, Eur Radiol 28 (2018) 4514–4523. https://doi.org/10.1007/s00330-018-5463-6.

[8] D. Pinto Dos Santos, M. Dietzel, B. Baessler, A decade of radiomics research: are images really data or just patterns in the noise?, Eur Radiol 31 (2021) 1–4. https://doi.org/10.1007/s00330-020-07108-w.

[9] A. Zwanenburg, Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis, Eur J Nucl Med Mol Imaging 46 (2019) 2638–2655. https://doi.org/10.1007/s00259-019-04391-8.

[10] M. Hatt, M. Vallieres, D. Visvikis, A. Zwanenburg, IBSI: an international community radiomics standardization initiative, Journal of Nuclear Medicine 59 (2018) 287–287.

[11] J. Santinha, D. Pinto Dos Santos, F. Laqua, J.J. Visser, K.B.W. Groot Lipman, M. Dietzel, M.E. Klontzas, R. Cuocolo, S. Gitto, T. Akinci D'Antonoli, ESR Essentials: radiomics—practice

recommendations by the European Society of Medical Imaging Informatics, Eur Radiol (2024). https://doi.org/10.1007/s00330-024-11093-9.
[12] A. Zwanenburg, M. Vallières, M.A. Abdalah, H.J.W.L. Aerts, V. Andrearczyk, A. Apte, S. Ashrafinia, S. Bakas, R.J. Beukinga, R. Boellaard, M. Bogowicz, L. Boldrini, I. Buvat, G.J.R. Cook, C. Davatzikos, A. Depeursinge, M.-C. Desseroit, N. Dinapoli, C.V. Dinh, S. Echegaray, I. El Naqa, A.Y. Fedorov, R. Gatta, R.J. Gillies, V. Goh, M. Götz, M. Guckenberger, S.M. Ha, M. Hatt, F. Isensee, P. Lambin, S. Leger, R.T.H. Leijenaar, J. Lenkowicz, F. Lippert, A. Losnegård, K.H. Maier-Hein, O. Morin, H. Müller, S. Napel, C. Nioche, F. Orlhac, S. Pati, E.A.G. Pfaehler, A. Rahmim, A.U.K. Rao, J. Scherer, M.M. Siddique, N.M. Sijtsema, J. Socarras Fernandez, E. Spezi, R.J.H.M. Steenbakkers, S. Tanadini-Lang, D. Thorwarth, E.G.C. Troost, T. Upadhaya, V. Valentini, L.V. van Dijk, J. van Griethuysen, F.H.P. van Velden, P. Whybra, C. Richter, S. Löck, The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping, Radiology 295 (2020) 328–338. https://doi.org/10.1148/radiol.2020191145.
[13] J. Peerlings, H.C. Woodruff, J.M. Winfield, A. Ibrahim, B.E. Van Beers, A. Heerschap, A. Jackson, J.E. Wildberger, F.M. Mottaghy, N.M. DeSouza, P. Lambin, Stability of radiomics features in apparent diffusion coefficient maps from a multi-centre test-retest trial, Sci Rep 9 (2019) 4800. https://doi.org/10.1038/s41598-019-41344-5.
[14] L. Bianchini, J. Santinha, N. Loução, M. Figueiredo, F. Botta, D. Origgi, M. Cremonesi, E. Cassano, N. Papanikolaou, A. Lascialfari, A multicenter study on radiomic features from T_2 -weighted images of a customized MR pelvic phantom setting the basis for robust radiomic models in clinics, Magnetic Resonance in Med 85 (2021) 1713–1726. https://doi.org/10.1002/mrm.28521.
[15] B. Baeßler, K. Weiss, D. Pinto dos Santos, Robustness and Reproducibility of Radiomics in Magnetic Resonance Imaging: A Phantom Study, Investigative Radiology 54 (2019) 221–228. https://doi.org/10.1097/RLI.0000000000000530.
[16] B.D. Wichtmann, F.N. Harder, K. Weiss, S.O. Schönberg, U.I. Attenberger, H. Alkadhi, Influence of Image Processing on Radiomic Features From Magnetic Resonance Imaging, Investigative Radiology 58 (2023).
[17] A. Traverso, M. Kazmierski, Z. Shi, P. Kalendralis, M. Welch, H.D. Nissen, D. Jaffray, A. Dekker, L. Wee, Stability of radiomic features of apparent diffusion coefficient (ADC) maps for locally advanced rectal cancer in response to image pre-processing, Physica Medica 61 (2019) 44–51. https://doi.org/10.1016/j.ejmp.2019.04.009.
[18] T. Tsuboyama, H. Onishi, A. Nakamoto, K. Ogawa, Y. Koyama, H. Tarewaki, N. Tomiyama, Impact of Deep Learning Reconstruction Combined With a Sharpening Filter on Single-Shot Fast Spin-Echo T2-Weighted Magnetic Resonance Imaging of the Uterus, Invest Radiol 57 (2022) 379–386. https://doi.org/10.1097/RLI.0000000000000847.
[19] A. Kanan, B. Pereira, C. Hordonneau, L. Cassagnes, E. Pouget, L.A. Tianhoun, B. Chauveau, B. Magnin, Deep learning CT reconstruction improves liver metastases detection, Insights Imaging 15 (2024) 167. https://doi.org/10.1186/s13244-024-01753-1.
[20] J. Zhong, Y. Xia, Y. Chen, J. Li, W. Lu, X. Shi, J. Feng, F. Yan, W. Yao, H. Zhang, Deep learning image reconstruction algorithm reduces image noise while alters radiomics features in dual-energy CT in comparison with conventional iterative reconstruction algorithms: a phantom study, Eur Radiol (2022). https://doi.org/10.1007/s00330-022-09119-1.
[21] F. Michallek, U. Genske, S.M. Niehues, B. Hamm, P. Jahnke, Deep learning reconstruction improves radiomics feature stability and discriminative power in abdominal CT imaging: a phantom study, Eur Radiol 32 (2022) 4587–4595. https://doi.org/10.1007/s00330-022-08592-y.
[22] H. Li, V.V. Alves, A. Pednekar, M.K. Manhard, J. Greer, A.T. Trout, L. He, J.R. Dillman, Impact of Emerging Deep Learning–Based MR Image Reconstruction Algorithms on Abdominal MRI Radiomic Features, J Comput Assist Tomogr (2024). https://doi.org/10.1097/RCT.0000000000001648.
[23] J. Zhong, Y. Xing, Y. Hu, X. Liu, S. Dai, D. Ding, J. Lu, J. Yang, Y. Song, M. Lu, D. Nickel, W. Lu, H. Zhang, W. Yao, Assessment of Robustness of MRI Radiomic Features in the Abdomen: Impact of

Deep Learning Reconstruction and Accelerated Acquisition, J Digit Imaging. Inform. Med. (2025). https://doi.org/10.1007/s10278-025-01503-9.
[24] J. Dana, E. McNabb, J. Castro, I. Al-Qanoobi, Y. Omiya, K. Ah-Lan, V. Fortier, G. Artho, C. Reinhold, S. Gauvin, Deep learning reconstruction in biparametric prostate MRI: Impact on qualitative and radiomics analyses, Res Diagn Interv Imaging 14 (2025) 100059. https://doi.org/10.1016/j.redii.2025.100059.
[25] R.M. Lebel, Performance characterization of a novel deep learning-based MR image reconstruction pipeline, (2020). http://arxiv.org/abs/2008.06559 (accessed October 28, 2024).
[26] A. Fedorov, R. Beichel, J. Kalpathy-Cramer, J. Finet, J.-C. Fillion-Robin, S. Pujol, C. Bauer, D. Jennings, F. Fennessy, M. Sonka, J. Buatti, S. Aylward, J.V. Miller, S. Pieper, R. Kikinis, 3D Slicer as an image computing platform for the Quantitative Imaging Network, Magn Reson Imaging 30 (2012) 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001.
[27] A. Zwanenburg, S. Löck, MIRP: A Python package for standardisedradiomics, JOSS 9 (2024) 6413. https://doi.org/10.21105/joss.06413.
[28] R Core Team, R: A Language and Environment for Statistical Computing, (n.d.). https://www.R-project.org/.
[29] T. Perrin, A. Midya, R. Yamashita, J. Chakraborty, T. Saidon, W.R. Jarnagin, M. Gonen, A.L. Simpson, R.K.G. Do, Short-term reproducibility of radiomic features in liver parenchyma and liver malignancies on contrast-enhanced CT imaging, Abdom Radiol 43 (2018) 3271–3278. https://doi.org/10.1007/s00261-018-1600-6.
[30] M.G. Poirot, M.W.A. Caan, H.G. Ruhe, A. Bjørnerud, I. Groote, L. Reneman, H.A. Marquering, Robustness of radiomics to variations in segmentation methods in multimodal brain MRI, Sci Rep 12 (2022) 16712. https://doi.org/10.1038/s41598-022-20703-9.
[31] E.C. Covert, K. Fitzpatrick, J. Mikell, R.K. Kaza, J.D. Millet, D. Barkmeier, J. Gemmete, J. Christensen, M.J. Schipper, Y.K. Dewaraja, Intra- and inter-operator variability in MRI-based manual segmentation of HCC lesions and its impact on dosimetry, EJNMMI Phys 9 (2022) 90. https://doi.org/10.1186/s40658-022-00515-6.
[32] Ortiz-Ramon R, Larroza A, Arana E, Moratal D. A radiomics evaluation of 2D and 3D MRI texture features to classify brain metastases from lung cancer and melanoma. Annu Int Conf IEEE Eng Med Biol Soc. 2017 Jul;2017:493-496. doi: 10.1109/EMBC.2017.8036869. PMID: 29059917., (n.d.).
[33] L. Duron, D. Balvay, S. Vande Perre, A. Bouchouicha, J. Savatovsky, J.-C. Sadik, I. Thomassin-Naggara, L. Fournier, A. Lecler, Gray-level discretization impacts reproducible MRI radiomics texture features, PLoS ONE 14 (2019) e0213459. https://doi.org/10.1371/journal.pone.0213459.
[34] G. Azemi, A. Di Ieva, Radiomic Fingerprinting of the Peritumoral Edema in Brain Tumors, Cancers 17 (2025) 478. https://doi.org/10.3390/cancers17030478.
[35] P. Rai, I.T. Mark, N. Soni, F. Diehn, S.A. Messina, J.C. Benson, A. Madhavan, A. Agarwal, G. Bathla, Deep Learning–Based Acceleration in MRI: Current Landscape and Clinical Applications in Neuroradiology, AJNR Am J Neuroradiol 47 (2026) 9–16. https://doi.org/10.3174/ajnr.A8943.
[36] S. Kiryu, H. Akai, K. Yasaka, T. Tajima, A. Kunimatsu, N. Yoshioka, M. Akahane, O. Abe, K. Ohtomo, Clinical Impact of Deep Learning Reconstruction in MRI, RadioGraphics 43 (2023) e220133. https://doi.org/10.1148/rg.220133.
[37] N. Behl, Deep Resolve – Mobilizing the Power of Networks, (n.d.).
[38] M. Wang, Y. Ma, L. Li, X. Pan, Y. Wen, Y. Qiu, D. Guo, Y. Zhu, J. Lian, D. Tong, Compressed Sensitivity Encoding Artificial Intelligence Accelerates Brain Metastasis Imaging by Optimizing Image Quality and Reducing Scan Time, AJNR Am J Neuroradiol 45 (2024) 444–452. https://doi.org/10.3174/ajnr.A8161.
[39] M. Kidoh, K. Shinoda, M. Kitajima, K. Isogawa, M. Nambu, H. Uetani, K. Morita, T. Nakaura, M. Tateishi, Y. Yamashita, Y. Yamashita, Deep Learning Based Noise Reduction for Brain MR Imaging: Tests on Phantoms and Healthy Volunteers, Magn Reson Med Sci 19 (2019) 195–206. https://doi.org/10.2463/mrms.mp.2019-0018.

[40] M. Tanabe, M. Higashi, T. Yonezawa, T. Yamaguchi, E. Iida, M. Furukawa, M. Okada, K. Shinoda, K. Ito, Feasibility of high-resolution magnetic resonance imaging of the liver using deep learning reconstruction based on the deep learning denoising technique, Magnetic Resonance Imaging 80 (2021) 121–126. https://doi.org/10.1016/j.mri.2021.05.001.
[41] K. Ozaki, H. Hasegawa, S. Ishida, S. Haseyama, Y. Tanahashi, S. Goshima, Deep learning reconstruction for liver DWI: impact on image quality and ADC quantification, European Journal of Radiology 202 (2026) 112925. https://doi.org/10.1016/j.ejrad.2026.112925.
[42] N. Nishioka, N. Fujima, S. Tsuneta, D. Kato, T. Kamiishi, M. Yoshikawa, R. Kimura, K. Sakamoto, R. Matsumoto, T. Abe, J. Kwon, M. Yoneyama, K. Kudo, Comparative evaluation of four reconstruction techniques for prostate T2-weighted MRI: Sensitivity encoding, compressed sensing, deep learning, and super-resolution, Eur J Radiol Open 15 (2025) 100671. https://doi.org/10.1016/j.ejro.2025.100671.
[43] F. Orlhac, How can we combat multicenter variability in MR radiomics? Validation of a correction procedure, Eur Radiol (2021) 9.
[44] M. Foltyn-Dumitru, M. Schell, A. Rastogi, F. Sahm, T. Kessler, W. Wick, M. Bendszus, G. Brugnara, P. Vollmuth, Impact of signal intensity normalization of MRI on the generalizability of radiomic-based prediction of molecular glioma subtypes, Eur Radiol 34 (2024) 2782–2790. https://doi.org/10.1007/s00330-023-10034-2.
[45] E. Stamoulou, C. Spanakis, G.C. Manikis, G. Karanasiou, G. Grigoriadis, T. Foukakis, M. Tsiknakis, D.I. Fotiadis, K. Marias, Harmonization Strategies in Multicenter MRI-Based Radiomics, J. Imaging 8 (2022) 303. https://doi.org/10.3390/jimaging8110303.
[46] R. Geirhos, J.-H. Jacobsen, C. Michaelis, R. Zemel, W. Brendel, M. Bethge, F.A. Wichmann, Shortcut learning in deep neural networks, Nat Mach Intell 2 (2020) 665–673. https://doi.org/10.1038/s42256-020-00257-z.
[47] A. Chartsias, T. Joyce, G. Papanastasiou, S. Semple, M. Williams, D.E. Newby, R. Dharmakumar, S.A. Tsaftaris, Disentangled representation learning in cardiac image analysis, Med Image Anal 58 (2019) 101535. https://doi.org/10.1016/j.media.2019.101535.
[48] B. Kocak, B. Baessler, S. Bakas, R. Cuocolo, A. Fedorov, L. Maier-Hein, N. Mercaldo, H. Müller, F. Orlhac, D. Pinto dos Santos, A. Stanzione, L. Ugga, A. Zwanenburg, CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII, Insights Imaging 14 (2023) 75. https://doi.org/10.1186/s13244-023-01415-8.

Tables

Table 1: Population characteristics.

	Total (n=95)
Age, in years mean (min-max)	66.96 (35-87)
Gender (%, M/F)	76.84 / 23.16
Histology type, n (%)	
Primary lesions	**38 (40.00%)**
HCC	35 (36.84 %)
Cholangiocarcinoma	3 (3.16 %)
Secondary lesion	**57 (60.00%)**
Colorectal adenocarcinoma	27 (28.42 %)
Neuroendocrine tumor	24 (25.26 %)
Pancreatic adenocarcinoma	3 (3.16 %)
Melanoma	3 (3.16%)
Lesion size in mm mean (min-max)	39 (11-200)
History of systemic treatment, n (%)	48 (50.53%)
History of TACE, n (%)	16 (16.84 %)

Table 2: MRI acquisition parameters.

Reference acquisition	In plane resolution (mm x mm)	Matrix	Slice thickness (mm)	TR (ms)	TE (ms)	NEX	b
axial T2	1.3 x 1.9	320 x 224	4.5	Minimum	108	1	/
axial T2FS	1.6 x 2.1	256 x 200	4	17645	Min Full	1	/
axial diffusion	3.8 x 2.5	112 x 150	4.5	17650	Min	12	800

Table 3: Analysis of intra-observer reproducibility

		All features		Intensity features		Texture features	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Tumor	T2	93.42 (142)	93.42 (142)	92.98 (53)	91.23 (52)	93.68 (89)	94.74 (90)
	T2FS	91.45 (139)	90.79 (138)	87.72 (50)	89.47 (51)	93.68 (89)	91.58 (87)
	Diffusion	90.79 (138)	90.79 (138)	91.23 (52)	89.47 (51)	90.53 (86)	91.58 (87)
Peri tumor	T2	88.16 (134)	91.45 (139)	91.23 (52)	87.72 (50)	86.32 (82)	93.68 (89)
	T2FS	67.76 (103)	74.34 (113)	75.44 (43)	80.70 (46)	63.16 (60)	70.53(67)
	Diffusion	84.87 (129)	86.18 (131)	85.96 (49)	82.46 (47)	84.21 (80)	88.42 (84)
Liver	T2	0	0	0	0	0	0
	T2FS	0	0	0	0	0	0
	Diffusion	0.66 (1)	0	1.75 (1)	0	0	0

Data are presented as number of features with ICC > 0.75 (%, (n)).

Table 4: Analysis of inter-observer reproducibility

		All features		Intensity features		Texture features	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Tumor	T2	6.58 (10)	5.92 (9)	3.51 (2)	1.75 (1)	8.42 (8)	8.42 (8)
	T2FS	8.55 (13)	7.24 (11)	3.51 (2)	3.51 (2)	11.58 (11)	9.47 (9)
	Diffusion	7.24 (11)	7.24 (11)	0	0	11.58 (11)	11.58 (11)
Peri tumor	T2	0.66 (1)	0.66 (1)	0	0	1.05 (1)	1.05 (1)
	T2FS	0.66 (1)	0.66 (1)	0	0	1.05 (1)	1.05 (1)
	Diffusion	0	0	0	0	0	0
Liver	T2	0	0	0	0	0	0
	T2FS	0	0	0	0	0	0
	Diffusion	0	0	0	0	0	0

Data are presented as number of features with ICC > 0.75 (%, (n)).

		All features		Intensity features		Texture features	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Difference between 3 tissus	**T2**	88.16 (134)	89.47 (136)	85.96 (49)	85.96 (49)	89.47 (85)	91.58 (87)
	T2FS	83.55 (127)	86.84 (132)	78.95 (45)	82.46 (47)	86.32 (82)	89.47 (85)
	Diffusion	89.47 (136)	90.79 (138)	85.96 (49)	87.72 (50)	91.58 (87)	92.63 (88)
Difference between tumor and peri tumor	**T2**	40.13 (61)	33.55 (51)	42.11 (24)	36.84 (21)	38.95 (37)	31.58 (30)
	T2FS	15.79 (24)	17.11 (26)	17.54 (10)	14.04 (8)	14.74 (14)	18.95 (18)
	Diffusion	34.21 (52)	37.50 (57)	36.84 (21)	36.84 (21)	32.63 (31)	37.89 (36)
Difference between tumor and liver	**T2**	73.68 (112)	80.92 (123)	61.40 (35)	66.67 (38)	81.05 (77)	89.47 (85)
	T2FS	78.29 (119)	79.61 (121)	70.18 (40)	73.68 (42)	83.16 (79)	83.16 (79)
	Diffusion	69.74 (106)	76.32 (116)	59.65 (34)	63.16 (36)	75.79 (72)	84.21 (80)
Difference between peri tumor and liver	**T2**	80.26 (122)	76.32 (116)	80.70 (46)	71.93 (41)	80.00 (76)	78.95 (75)
	T2FS	80.26 (122)	80.92 (123)	75.44 (43)	75.44 (43)	83.16 (79)	84.21 (80)
	Diffusion	80.92 (123)	80.92 (123)	71.93 (41)	71.93 (41)	86.32 (82)	86.32 (82)

Table 5: Tissue type discriminative power - statistical separability: percentage of feature differences between tissues

Data are presented as % (n).

Table 6: Tissue type discriminative power: predictive performance: performance of classifiers in differentiating tissue types across reconstruction methods.

		All features				Intensity features				
Sequence	**Reconstruction**	**Mean Accuracy**	**T-test p-value**	**Mean Sensitivity**	**Mean Specificity**	**Mean Accuracy**	**T-test p-value**	**Mean Sensitivity**	**Mean Specificity**	**A**
Diffusion	**DLIR**	0.865		0.865	0.932	0.770		0.771	0.885	
	non-DLIR	0.863	0.687	0.863	0.931	0.774	0.509	0.774	0.887	
T2	**DLIR**	0.812		0.812	0.906	0.704		0.704	0.852	
	non-DLIR	0.803	0.100	0.803	0.901	0.731	**<0.001**	0.731	0.865	
T2FS	**DLIR**	0.796		0.796	0.898	0.733		0.733	0.866	
	non-DLIR	0.804	0.117	0.804	0.902	0.733	0.954	0.733	0.867	

Bold characters indicate significative difference (after Bonferroni correction for multiple comparisons)

Table 7: Relative frequency of radiomic parameters affected by the reconstruction technique

		All features	Intensity features	Higher order features
Tumor	**T2**	28.29 (43)	3.51 (2)	43.16 (41)
	T2FS	28.95 (44)	7.02 (4)	42.11 (40)
	Diffusion	1.32 (2)	0	2.11 (2)
Peri tumor	**T2**	38.16 (58)	8.77 (5)	55.79 (53)
	T2FS	37.5 (57)	8.77 (5)	54.74 (52)
	Diffusion	7.89 (12)	0	12.63 (12)
liver	**T2**	34.21 (52)	17.54 (10)	44.21 (42)
	T2FS	36.84 (56)	17.54 (10)	48.42 (46)
	Diffusion	0.66 (1)	0	1.05 (1)

Data are presented as % (n).

Figures

Figure 1: Patient selection flowchart

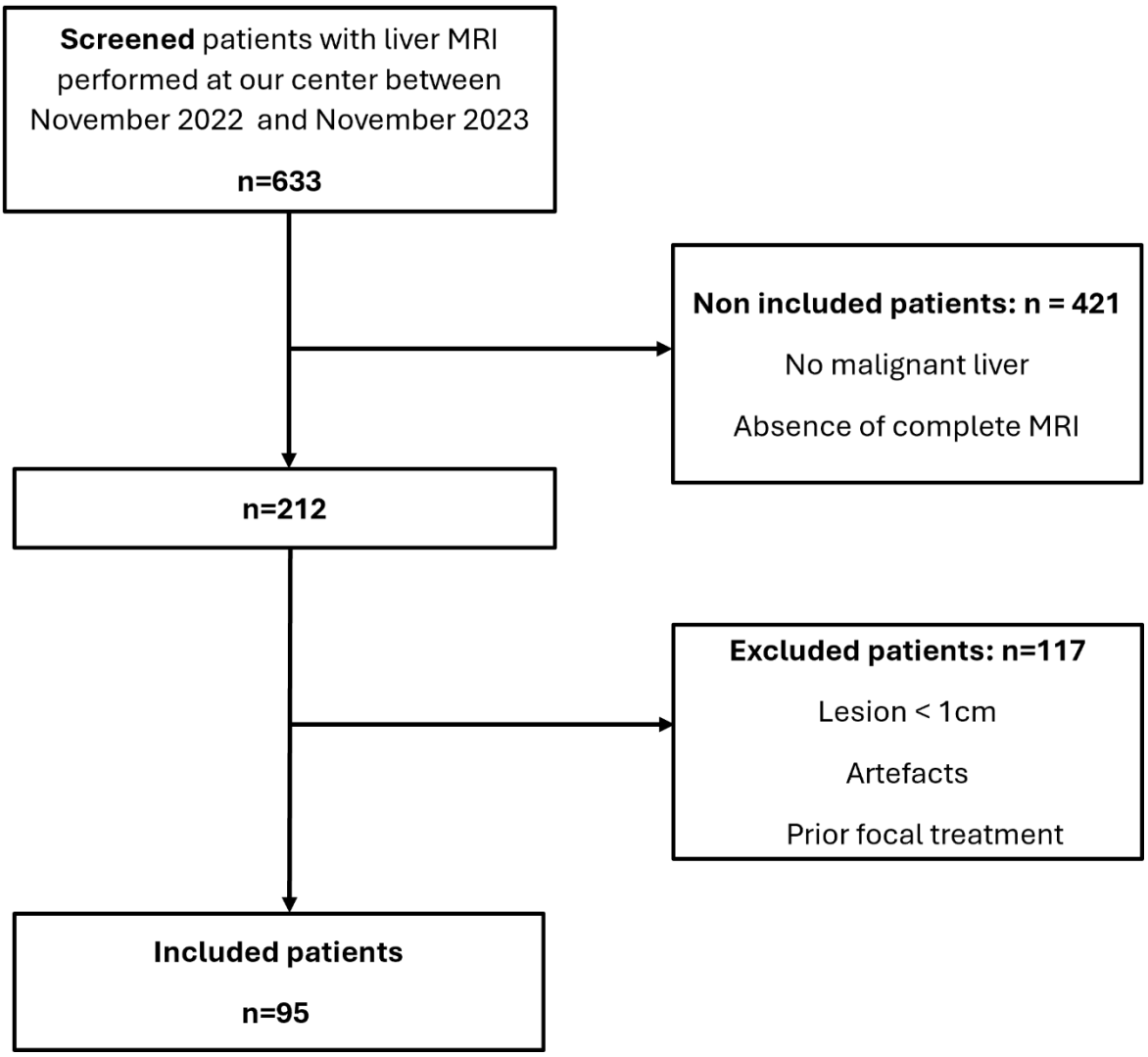


Figure 2: Study design

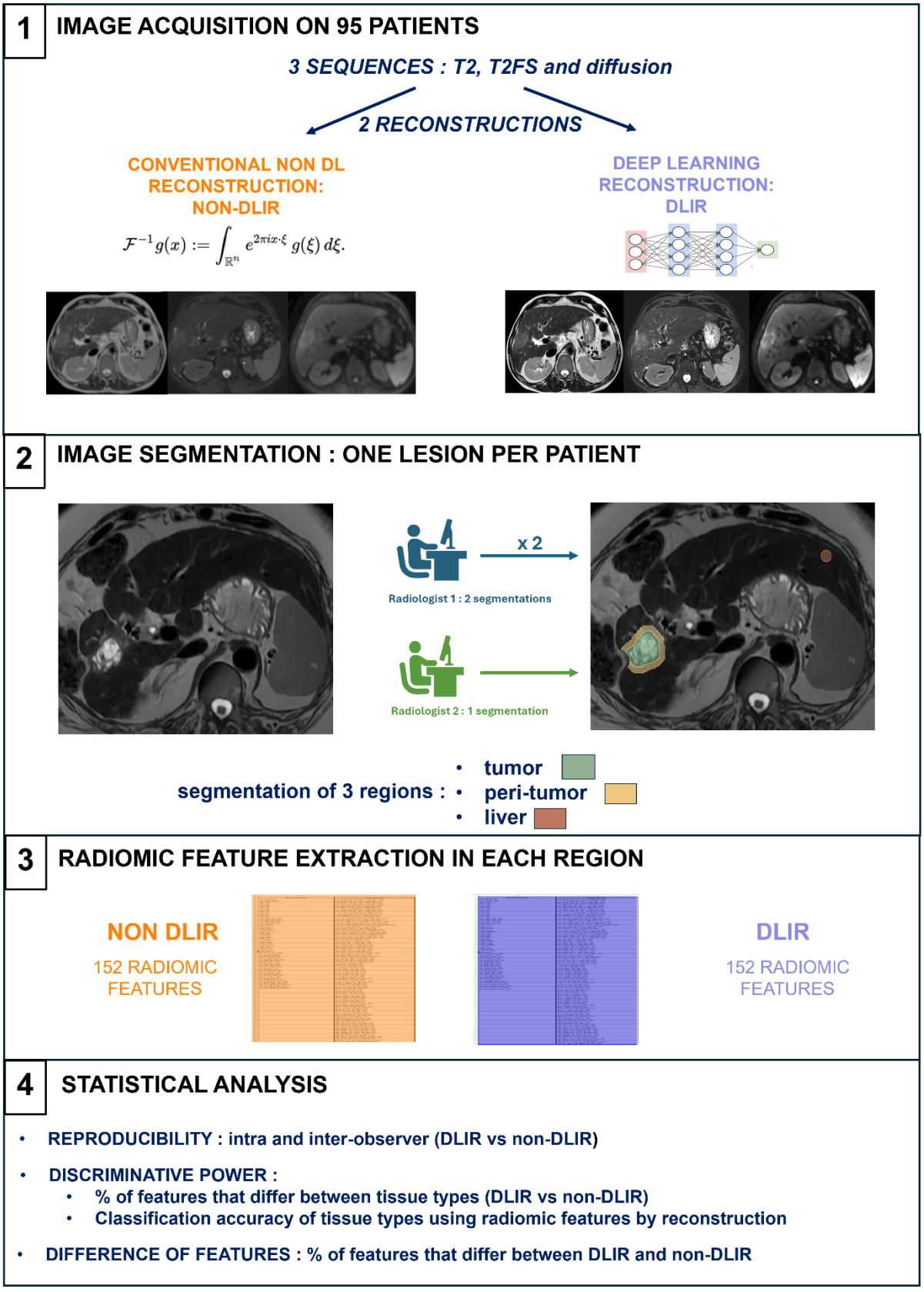


Figure 3: Graphical example of the segmentation of the three ROIs on an axial T2 image

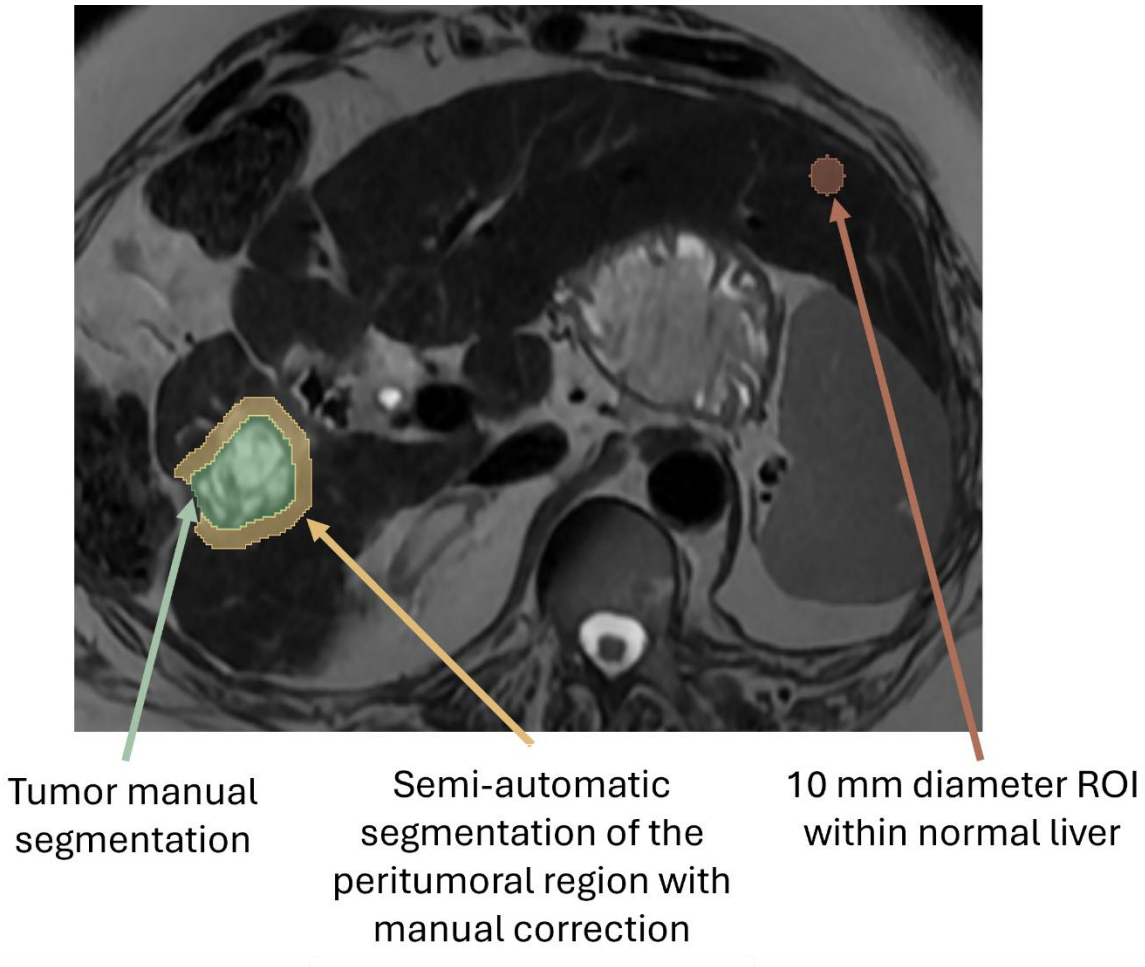
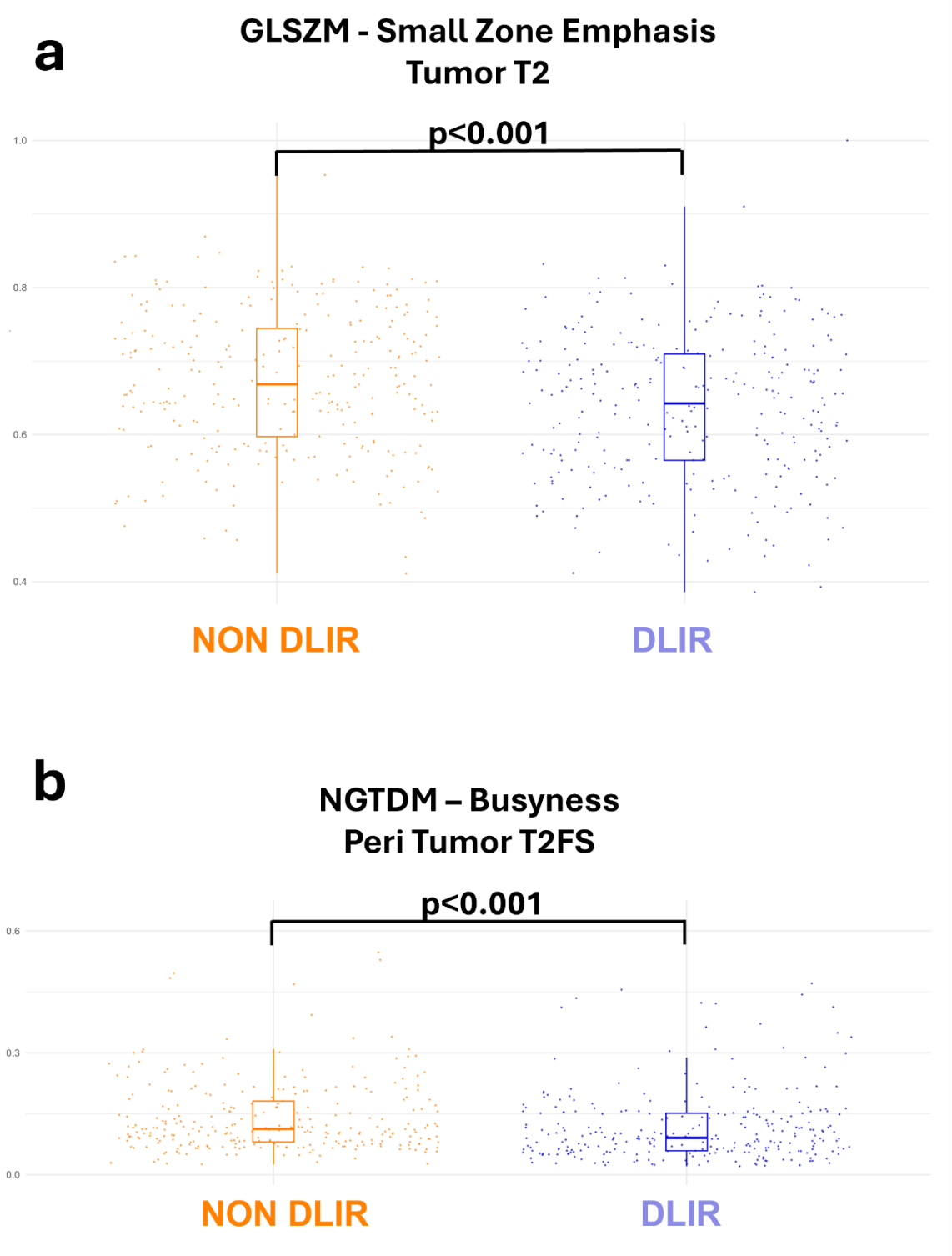


Figure 4: Graphical example of the differences of feature distribution depending on reconstruction. a: Boxplot of tumor T2 GLSZM - Small Zone Emphasis values with DLIR and non-DLIR, DLIR values being significantly lower (p<0.001); b: Boxplot of peri tumor T2FS NGTDM - Busyness values with DLIR and non-DLIR, DLIR values being significantly lower (p<0.001)



Supplementary Material

Intensity features	Texture features	
loc_peak_loc	cm_joint_max_d1_3d_v_mrg_fbn_n32	dzm_sde_3d_fbn_n32
loc_peak_glob	cm_joint_avg_d1_3d_v_mrg_fbn_n32	dzm_lde_3d_fbn_n32
ivh_v10	cm_joint_var_d1_3d_v_mrg_fbn_n32	dzm_lgze_3d_fbn_n32
ivh_v25	cm_joint_entr_d1_3d_v_mrg_fbn_n32	dzm_hgze_3d_fbn_n32
ivh_v50	cm_diff_avg_d1_3d_v_mrg_fbn_n32	dzm_sdlge_3d_fbn_n32
ivh_v75	cm_diff_var_d1_3d_v_mrg_fbn_n32	dzm_sdhge_3d_fbn_n32
ivh_v90	cm_diff_entr_d1_3d_v_mrg_fbn_n32	dzm_ldlge_3d_fbn_n32
ivh_i10	cm_sum_avg_d1_3d_v_mrg_fbn_n32	dzm_ldhge_3d_fbn_n32
ivh_i25	cm_sum_var_d1_3d_v_mrg_fbn_n32	dzm_glnu_3d_fbn_n32
ivh_i50	cm_sum_entr_d1_3d_v_mrg_fbn_n32	dzm_glnu_norm_3d_fbn_n32
ivh_i75	cm_energy_d1_3d_v_mrg_fbn_n32	dzm_zdnu_3d_fbn_n32
ivh_i90	cm_contrast_d1_3d_v_mrg_fbn_n32	dzm_zdnu_norm_3d_fbn_n32
ivh_diff_v10_v90	cm_dissimilarity_d1_3d_v_mrg_fbn_n32	dzm_z_perc_3d_fbn_n32
ivh_diff_v25_v75	cm_inv_diff_d1_3d_v_mrg_fbn_n32	dzm_gl_var_3d_fbn_n32
ivh_diff_i10_i90	cm_inv_diff_norm_d1_3d_v_mrg_fbn_n32	dzm_zd_var_3d_fbn_n32
ivh_diff_i25_i75	cm_inv_diff_mom_d1_3d_v_mrg_fbn_n32	dzm_zd_entr_3d_fbn_n32
stat_mean	cm_inv_diff_mom_norm_d1_3d_v_mrg_fbn_n32	ngt_coarseness_3d_fbn_n32
stat_var	cm_inv_var_d1_3d_v_mrg_fbn_n32	ngt_contrast_3d_fbn_n32
stat_skew	cm_corr_d1_3d_v_mrg_fbn_n32	ngt_busyness_3d_fbn_n32
stat_kurt	cm_auto_corr_d1_3d_v_mrg_fbn_n32	ngt_complexity_3d_fbn_n32
stat_median	cm_clust_tend_d1_3d_v_mrg_fbn_n32	ngt_strength_3d_fbn_n32
stat_min	cm_clust_shade_d1_3d_v_mrg_fbn_n32	ngl_lde_d1_a0.0_3d_fbn_n32
stat_p10	cm_clust_prom_d1_3d_v_mrg_fbn_n32	ngl_hde_d1_a0.0_3d_fbn_n32
stat_p90	cm_info_corr1_d1_3d_v_mrg_fbn_n32	ngl_lgce_d1_a0.0_3d_fbn_n32
stat_max	cm_info_corr2_d1_3d_v_mrg_fbn_n32	ngl_hgce_d1_a0.0_3d_fbn_n32
stat_iqr	rlm_sre_3d_v_mrg_fbn_n32	ngl_ldlge_d1_a0.0_3d_fbn_n32
stat_range	rlm_lre_3d_v_mrg_fbn_n32	ngl_ldhge_d1_a0.0_3d_fbn_n32
stat_mad	rlm_lgre_3d_v_mrg_fbn_n32	ngl_hdlge_d1_a0.0_3d_fbn_n32
stat_rmad	rlm_hgre_3d_v_mrg_fbn_n32	ngl_hdhge_d1_a0.0_3d_fbn_n32
stat_medad	rlm_srlge_3d_v_mrg_fbn_n32	ngl_glnu_d1_a0.0_3d_fbn_n32
stat_cov	rlm_srhge_3d_v_mrg_fbn_n32	ngl_glnu_norm_d1_a0.0_3d_fbn_n32
stat_qcod	rlm_lrlge_3d_v_mrg_fbn_n32	ngl_dcnu_d1_a0.0_3d_fbn_n32
stat_energy	rlm_lrhge_3d_v_mrg_fbn_n32	ngl_dcnu_norm_d1_a0.0_3d_fbn_n32
stat_rms	rlm_glnu_3d_v_mrg_fbn_n32	ngl_dc_perc_d1_a0.0_3d_fbn_n32

ih_mean_fbn_n32	rlm_glnu_norm_3d_v_mrg_fbn_n32	ngl_gl_var_d1_a0.0_3d_fbn_n32
ih_var_fbn_n32	rlm_rlnu_3d_v_mrg_fbn_n32	ngl_dc_var_d1_a0.0_3d_fbn_n32
ih_skew_fbn_n32	rlm_rlnu_norm_3d_v_mrg_fbn_n32	ngl_dc_entr_d1_a0.0_3d_fbn_n32
ih_kurt_fbn_n32	rlm_r_perc_3d_v_mrg_fbn_n32	ngl_dc_energy_d1_a0.0_3d_fbn_n32
ih_median_fbn_n32	rlm_gl_var_3d_v_mrg_fbn_n32	
ih_min_fbn_n32	rlm_rl_var_3d_v_mrg_fbn_n32	
ih_p10_fbn_n32	rlm_rl_entr_3d_v_mrg_fbn_n32	
ih_p90_fbn_n32	szm_sze_3d_fbn_n32	
ih_max_fbn_n32	szm_lze_3d_fbn_n32	
ih_mode_fbn_n32	szm_lgze_3d_fbn_n32	
ih_iqr_fbn_n32	szm_hgze_3d_fbn_n32	
ih_range_fbn_n32	szm_szlge_3d_fbn_n32	
ih_mad_fbn_n32	szm_szhge_3d_fbn_n32	
ih_rmad_fbn_n32	szm_lzlge_3d_fbn_n32	
ih_medad_fbn_n32	szm_lzhge_3d_fbn_n32	
ih_cov_fbn_n32	szm_glnu_3d_fbn_n32	
ih_qcod_fbn_n32	szm_glnu_norm_3d_fbn_n32	
ih_entropy_fbn_n32	szm_zsnu_3d_fbn_n32	
ih_uniformity_fbn_n32	szm_zsnu_norm_3d_fbn_n32	
ih_max_grad_fbn_n32	szm_z_perc_3d_fbn_n32	
ih_max_grad_g_fbn_n32	szm_gl_var_3d_fbn_n32	
ih_min_grad_fbn_n32	szm_zs_var_3d_fbn_n32	
ih_min_grad_g_fbn_n32	szm_zs_entr_3d_fbn_n32	

Supplementary material 1: List of radiomic features extracted by MIRP.

		Local intensity		Intensity based statistical		Intensity histogram		Intensity volume histogram		GLCM		GLRLM		GLSZM		GLDZM		NGTDM		NGLDM	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Tumor	**T2**	100 (2)	100 (2)	94,44 (17)	88,89 (16)	86,96 (20)	86,96 (20)	100 (24)	100 (24)	100 (25)	100 (25)	100 (16)	100 (16)	93,75 (15)	100 (16)	75 (12)	75 (12)	100 (5)	100 (5)	94,12 (16)	94,12 (16)
	T2FS	100 (2)	100 (2)	83,33 (15)	88,89 (16)	82,61 (19)	82,61 (19)	100 (24)	100 (24)	100 (25)	100 (25)	100 (16)	100 (16)	93,75 (15)	81,25 (13)	75 (12)	75 (12)	100 (5)	100 (5)	94,12 (16)	94,12 (16)
	Diff	100 (2)	100 (2)	94,44 (17)	94,44 (17)	82,61 (19)	78,26 (18)	100 (24)	100 (24)	96 (24)	100 (25)	100 (16)	100 (16)	87,5 (14)	93,75 (15)	68,75 (11)	62,5 (10)	100 (5)	100 (5)	94,12 (16)	94,12 (16)
Peri tumor	**T2**	100 (2)	100 (2)	88,89 (16)	83,33 (15)	86,96 (20)	82,61 (19)	100 (24)	100 (24)	100 (25)	100 (25)	100 (16)	100	75 (12)	93,75 (15)	56,25 (9)	75 (12)	100 (5)	100 (5)	88,23 (15)	94,12 (16)
	T2FS	100 (2)	100 (2)	83,33 (15)	83,33 (15)	60,87 (14)	65,22 (15)	85,71 (12)	100 (24)	84 (21)	88 (22)	50 (8)	50 (8)	62,5 (10)	75 (12)	43,75 (7)	68,75 (11)	100 (5)	100 (5)	52,94 (9)	52,94 (9)
	Diff	100 (2)	100 (2)	88,89 (16)	88,89 (16)	78,26 (18)	78,26 (18)	92,86 (13)	78,57 (11)	88 (22)	100 (25)	93,75 (15)	93,75 (15)	87,5 (14)	93,75 (15)	75 (12)	68,75 (11)	80 (4)	80 (4)	76,47 (13)	82,35 (14)
Liver	**T2**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	T2FS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Diff	0	0	5,56 (1)	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Supplementary material 2: Detailed analysis of intra-observer reproducibility, number of features with ICC > 0.7 (%, (n)).

		Local intensity		Intensity based statistical		Intensity histogram		Intensity volume histogram		GLCM		GLRLM		GLSZM		GLDZM		NGTDM		NGLDM	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Tumor	T2	0	0	0	0	8,70 (2)	4,35(1)	0	0	0	0	12,5 (2)	12,5 (2)	6,25 (1)	6,25 (1)	12,5 (2)	12,5 (2)	20 (1)	20 (1)	11,76 (2)	11,76 (2)
	T2FS	0	0	0	0	8,70(2)	8,70 (2)	0	0	0	0	12,5 (2)	12,5 (2)	25 (4)	12,5 (2)	12,5 (2)	12,5 (2)	0	20 (1)	17,65 (3)	11,76 (2)
	Diff	0	0	0	0	0	0	0	0	0	0	12,5 (2)	12,5 (2)	18,75 (3)	18,75 (3)	12,5 (2)	12,5 (2)	0	0	23,53(4)	23,53 (4)
Peri tumor	T2	0	0	0	0	0	0	0	0	0	0	6,25 (1)	6,25 (1)	0	0	0	0	0	0	0	0
	T2FS	0	0	0	0	0	0	0	0	0	0	6,25 (1)	6,25 (1)	0	0	0	0	0	0	0	0
	Diff	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Liver	T2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	T2FS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	Diff	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Supplementary material 3: Detailed analysis of inter-observer reproducibility, number of features with ICC > 0.75 (%, (n)).

		Local intensity		Intensity based statistical		Intensity histogram		Intensity volume histogram		GLCM		GLRLM		GLSZM		GLDZM		NGTDM		NGLDM	
		DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR	DLIR	Non-DLIR
Difference between 3 tissus	**T2**	100 (2)	100 (2)	83,33 (15)	83,33 (15)	86,96 (20)	82,61 (19)	85,71 (12)	92,86 (13)	100 (25)	92 (23)	87,5 (14)	93,75 (15)	93,75 (15)	100 (16)	62,5 (10)	75 (12)	100 (5)	100 (5)	94,12 (16)	94,12 (16)
	T2FS	100 (2)	100 (2)	88,89 (16)	100,00 (18)	69,57 (16)	69,57 (16)	78,57 (11)	78,57 (11)	92 (23)	96 (24)	75 (12)	81,25 (13)	100 (16)	100 (16)	68,75 (11)	75 (12)	100 (5)	100 (5)	88,24 (15)	88,24 (15)
	Diff	100 (2)	100 (2)	94,44 (17)	94,44 (17)	82,61 (19)	82,61 (19)	78,57 (11)	85,71 (12)	88 (22)	92 (23)	100 (16)	100 (16)	100 (16)	100 (16)	75 (12)	75 (12)	100 (5)	100 (5)	94,12 (16)	94,12 (16)
Difference between tumor and peri tumor	**T2**	0	0	33,33 (6)	33,33 (6)	47,83 (11)	43,48 (10)	50 (7)	35,71 (5)	20 (5)	20 (5)	50 (8)	31,25 (5)	50 (8)	50 (8)	50 (8)	31,25 (5)	20 (1)	40 (2)	41,18 (7)	29,41 (5)
	T2FS	0	0	33,33 (6)	33,33 (6)	17,39 (4)	8,7 (2)	0	0	4 (1)	0	12,5 (2)	25 (4)	31,25 (5)	43,75 (7)	12,5 (2)	18,75 (3)	40 (2)	40 (2)	11,76 (2)	11,76 (2)
	Diff	0	0	44,44 (8)	44,44 (8)	39,13 (9)	34,78 (9)	28,57 (4)	35,71 (5)	40 (10)	36 (9)	31,25 (5)	50 (8)	37,5 (6)	43,75 (7)	12,5 (2)	18,75 (3)	40 (2)	40 (2)	35,29 (6)	41,18 (7)

Difference between tumor and liver	**T2**	100 (2)	100 (2)	55,56 (10)	61,11 (11)	65,22 (15)	69,57 (16)	57,14 (8)	64,29 (9)	84 (21)	92 (23)	68,75 (11)	87,5 (14)	93,75 (15)	100 (16)	62,5 (10)	75 (12)	100 (5)	100 (5)	88,24 (15)	88,24 (15)
	T2FS	100 (2)	100 (2)	83,33 (15)	83,33 (15)	65,22 (15)	69,57 (16)	57,14 (8)	64,29 (9)	92 (23)	92 (23)	68,75 (11)	75 (12)	93,75 (15)	93,75 (15)	62,5 (10)	56,25 (9)	100 (5)	100 (5)	88,24 (15)	88,24 (15)
	Diff	100 (2)	100 (2)	72,22 (13)	72,22 (13)	56,52 (13)	56,52 (13)	42,86 (6)	57,14 (8)	64 (16)	72 (18)	75 (12)	87,5 (14)	93,75 (15)	100 (16)	56,25 (9)	75 (12)	100 (5)	100 (5)	88,24 (15)	88,24 (15)
Difference between peri tumor and liver	**T2**	100 (2)	100 (2)	77,78 (14)	72,22 (13)	82,61 (19)	69,57 (16)	78,57 (11)	71,43 (10)	88 (22)	80 (20)	75 (12)	75 (12)	93,75 (15)	93,75 (15)	50 (8)	56,25 (9)	100 (5)	100 (5)	82,35 (14)	82,35 (14)
	T2FS	100 (2)	100 (2)	77,78 (14)	83,33 (15)	73,91 (17)	69,57 (16)	71,43 (10)	71,43 (10)	92 (23)	96 (24)	81,25 (13)	81,25 (13)	93,75 (15)	93,75 (15)	50 (8)	56,25 (9)	100 (5)	100 (5)	88,24 (15)	82,35 (14)
	Diff	100 (2)	100 (2)	77,78 (14)	77,78 (14)	69,57 (16)	69,57 (16)	64,29 (9)	64,29 (9)	80 (20)	72 (18)	93,75 (15)	93,75 (15)	100 (16)	100 (16)	75 (12)	75 (12)	80 (4)	100 (5)	88,24 (15)	94,12 (16)

Supplementary material 4: Detailed analysis of tissue type discriminative power - statistical separability: percentage of feature differences between tissues (% (n)).

Supplementary material 5: Detailed analysis of relative frequency of radiomic parameters affected by the reconstruction technique (% (n)).

		Local intensity	Intensity based statistical	Intensity histogram	Intensity volume histogram	GLCM	GLRLM	GLSZM	GLDZM	NGTDM	NGLDM
Tumor	**T2**	0	5,56 (1)	4,35 (1)	0	56 (14)	31,25 (5)	56,25 (9)	18,75 (3)	40 (2)	47,06 (8)
	T2FS	0	16,67 (3)	4,35 (1)	0	56 (14)	31,25 (5)	50 (8)	18,75 (3)	40 (2)	47,06 (8)
	Diff	0	0	0	0	8 (2)	0	0	0	0	0
Peri tumor	**T2**	0	0,00	17,39 (4)	7,14 (1)	64 (16)	56,25 (9)	62,5 (10)	18,75 (3)	80 (4)	64,71 (11)
	T2FS	0	5,56 (1)	17,39 (4)	0	64 (16)	37,5 (6)	62,5 (10)	18,75 (3)	100 (5)	70,59 (12)
	Diff	0	0,00	0	0	8 (2)	18,75 (3)	18,75 (3)	6,25 (1)	0	17,65 (3)
Liver	**T2**	0	33,33 (6)	13,04 (3)	7,14 (1)	64 (16)	31,25 (5)	50 (8)	18,75 (3)	80 (4)	35,29 (6)
	T2FS	0	38,89 (7)	8,70 (2)	7,14 (1)	64 (16)	43,75 (7)	56,25 (9)	18,75 (3)	80 (4)	41,18 (7)
	Diff	0	0	0	0	4 (1)	0	0	0	0	0