Clinical Imaging 126 (2025) 110581

Contents lists available at ScienceDirect

Clinical Imaging

journal homepage: www.elsevier.com/locate/clinimag

Body Imaging

Optimizing contrast-enhanced abdominal MRI: A comparative study of deep learning and standard VIBE techniques

Alexander Herold [a,b], Nathaniel D. Mercaldo [a], Mark A. Anderson [a], Amirkasra Mojtahed [a], Aoife Kilcoyne [a], Wei-Ching Lo [c], Robert M. Sellers [c], Bryan Clifford [c], Marcel D. Nickel [d], Nabih Nakrour [a], Susie Y. Huang [a], Leo L. Tsai [a], Onofrio A. Catalano [a,1], Mukesh G. Harisinghani [a,*,1]

[a] Department of Radiology, Massachusetts General Hospital, Boston, MA, USA
[b] Department of Biomedical Imaging and Image-guided Therapy, Medical University of Vienna, Austria
[c] Siemens Medical Solutions USA, Boston, MA, USA
[d] Siemens Healthineers AG, Forchheim, Germany

ARTICLE INFO

Keywords:
Magnetic resonance imaging
Deep-learning image reconstruction
Liver
VIBE
3D Gradient Echo

ABSTRACT

Objective: To validate a deep learning (DL) reconstruction technique for faster post-contrast enhanced coronal Volume Interpolated Breath-hold Examination (VIBE) sequences and assess its image quality compared to conventionally acquired coronal VIBE sequences.

Methods: This prospective study included 151 patients undergoing clinically indicated upper abdominal MRI acquired on 3 T scanners. Two coronal T1 fat-suppressed VIBE sequences were acquired: a DL-reconstructed sequence ($VIBE_{DL}$) and a standard sequence ($VIBE_{SD}$). Three radiologists independently evaluated six image quality parameters: overall image quality, perceived signal-to-noise ratio, severity of artifacts, liver edge sharpness, liver vessel sharpness, and lesion conspicuity, using a 4-point Likert scale. Inter-reader agreement was assessed using Gwet's AC2. Ordinal mixed-effects regression models were used to compare $VIBE_{DL}$ and $VIBE_{SD}$.

Results: Acquisition times were 10.2 s for $VIBE_{DL}$ compared to 22.3 s for $VIBE_{SD}$. $VIBE_{DL}$ demonstrated superior overall image quality (OR 1.95, 95 % CI: 1.44–2.65, $p < 0.001$), reduced image noise (OR 3.02, 95 % CI: 2.26–4.05, $p < 0.001$), enhanced liver edge sharpness (OR 3.68, 95 % CI: 2.63–5.15, $p < 0.001$), improved liver vessel sharpness (OR 4.43, 95 % CI: 3.13–6.27, $p < 0.001$), and better lesion conspicuity (OR 9.03, 95 % CI: 6.34–12.85, $p < 0.001$) compared to $VIBE_{SD}$. However, $VIBE_{DL}$ showed increased severity of peripheral artifacts (OR 0.13, $p < 0.001$). $VIBE_{DL}$ detected 137/138 (99.3 %) focal liver lesions, while $VIBE_{SD}$ detected 131/138 (94.9 %). Inter-reader agreement ranged from good to very good for both sequences.

Conclusion: The DL-reconstructed VIBE sequence significantly outperformed the standard breath-hold VIBE in image quality and lesion detection, while reducing acquisition time. This technique shows promise for enhancing the diagnostic capabilities of contrast-enhanced abdominal MRI.

1. Introduction

Over the last decades, magnetic resonance imaging (MRI) has become the modality of choice for evaluating upper abdominal pathologies.[1] One of the cornerstones of multiphasic liver MRI protocols is the three-dimensional (3D) T1-weighted fat-suppressed gradient-recalled echo (GRE) sequence, particularly the volumetric interpolated breath-hold examination (VIBE), which provides high-resolution, contrast-enhanced T1-weighted images essential for detecting and characterizing focal liver lesions.[2] GRE sequences, with their shorter repetition times, allow for significant reductions in acquisition time compared to other sequences.[3] However, despite their widespread use in clinical practice, GRE sequences still typically require breath-holds of 15–20 s,[4] which can be challenging for many patients, particularly those with

* Corresponding author at: Department of Radiology, Division of Abdominal Imaging, Massachusetts General Hospital, Harvard Medical School, 55 Fruit St, White Building, Room 270, Boston, MA 02114, USA.
E-mail address: mharisinghani@mgh.harvard.edu (M.G. Harisinghani).
[1] These authors share last authorship.

https://doi.org/10.1016/j.clinimag.2025.110581
Received 9 December 2024; Received in revised form 28 July 2025; Accepted 4 August 2025
Available online 7 August 2025

compromised respiratory function or an inability to cooperate. This makes these sequences susceptible to motion artifacts, and image quality remains heavily dependent on patient cooperation during the examination.

Rapid imaging techniques that decrease breath-hold times or allow free-breathing acquisition have helped to overcome these challenges. A significant advancement came with the introduction of radial k-space sampling, dramatically improving the robustness against motion artifacts.[5] Parallel imaging methods such as Sensitivity Encoding (SENSE), Generalized Autocalibrating Partially Parallel Acquisitions (GRAPPA), and Controlled Aliasing in Parallel Imaging Results in Higher Acceleration (CAIPIRINHA) have also been adopted clinically, enabling significant reductions in scan duration.[6–8] The introduction of compressed sensing further advanced these efforts by leveraging image sparsity for even greater acceleration.[9]

The recent advent of deep learning (DL) reconstruction algorithms represents another paradigm shift in MRI acceleration strategies. DL reconstruction employs sophisticated neural network architectures to synthesize high-quality images from undersampled MRI data.[10]

Early applications of DL-based reconstructions in abdominal MRI, including 3D T1-weighted GRE sequences, have shown promise for improving image quality and reducing scan times.[11–16] However, validation of these DL reconstruction algorithms in clinical settings remains limited, and diagnostic performance and clinical applicability of DL-reconstructed VIBE sequences in prospective patient cohorts need to be further evaluated to determine their potential impact on routine abdominal imaging protocols.

At our institution, a key sequence in the upper abdominal MRI protocol includes a delayed post-contrast enhanced, breath-hold, coronal VIBE sequence. As it is the last sequence of the protocol, it is acquired when patients may be fatigued and struggling to maintain the required 15–20 s breath-hold, potentially compromising image quality. This challenge underscores the need for techniques that can provide high-quality images with reduced breath-hold requirements.

We hypothesize that an accelerated VIBE with deep learning-based reconstruction will demonstrate superior image quality compared to the standard sequence while reducing acquisition and, therefore, breath-hold time with a particular focus on the liver. Additionally, we anticipate improved liver lesion conspicuity and detectability. Therefore, this study aimed to validate a DL reconstruction technique that utilizes an unrolled variational network and super-resolution algorithm for a post-contrast enhanced coronal VIBE sequence and assess its image quality compared to the conventional VIBE sequence.

2. Methods and materials

2.1. Study design

This prospective, single-center study was conducted in accordance with the Declaration of Helsinki and its amendments and received approval from the local institutional review board (IRB Nr.: 2018P000832). All participants provided informed consent to undergo additional sequences during their scheduled examination. Between May and August 2024, we consecutively enrolled 151 patients (≥18 years of age) referred for clinically indicated upper abdominal MRI. Siemens Healthineers collaborators provided research deep learning applications for the study. Exclusion criteria encompassed contraindications to MRI (such as claustrophobia or non-MRI-compatible implants), pregnancy, non-contrast examinations, use of hepatobiliary contrast media (as the DL-VIBE sequence was not yet implemented in our hepatobiliary protocol), and incomplete MRI datasets.

2.2. MRI protocol

All examinations were acquired on 3 T MRI scanners (MAGNETOM Vida, Siemens Healthineers, Forchheim, Germany) using phased array multi-channel body coils. Routine liver protocols included pre-contrast T1-weighted in- and opposed-phase imaging, T2-weighted imaging with and without fat suppression, diffusion-weighted imaging (*b* values: 100, 500, 1500 s/mm$^{2)}$ with corresponding apparent diffusion coefficient maps, and dynamic contrast-enhanced (DCE) images following weight-adjusted intravenous administration of gadoterate meglumine (Dotarem®, Guerbet, Princeton, NJ). Axial fat-suppressed T1-weighted DCE images were extracted from a golden-angle radial sparse parallel (GRASP) acquisition for time points at 25–30 s, 70 s and 180 s. Two breath-hold, coronal T1 fat-suppressed VIBE sequences after 300–600 s were acquired: a research sequence with DL-based reconstruction and eightfold CAIPIRINHA-acceleration (VIBE$_{DL}$) and a standard sequence with a fourfold acceleration (VIBE$_{SD}$), as detailed in Table 1.

2.3. Deep-learning based image reconstruction

The DL-based reconstruction involved two sequential steps (Fig. S1). The initial step utilized an architecture inspired by unrolled variational networks to reconstruct images at the initially acquired resolution.[10] The reconstruction algorithm uses undersampled k-space data and coil sensitivity maps as input. It then generates an image through six iterative steps that alternate between data consistency updates (ensuring the reconstructed image matches the acquired data) via parallel imaging reconstruction and image regularization using neural network evaluation. The network architecture was implemented using the PyTorch framework.[17] The network underwent supervised training using approximately 5000 training pairs derived from 500 fully sampled 3D datasets. These datasets were acquired from 1.5 T and 3 T MAGNETOM scanners (Siemens Healthineers, Forchheim, Germany) of the head, abdomen, and pelvis. The trained network was then integrated into the scanner reconstruction pipeline using ONNX Runtime.[18] This first reconstruction step, including necessary preprocessing, required approximately 15–20 s for the 3D T1-VIBE sequence, utilizing the scanner-integrated graphics processing units. The second step employed a DL-based super-resolution algorithm to further enhance image quality.[15,16,19] This algorithm performs an initial upsampling by a factor of 2 in all spatial dimensions. Additionally, it enforces data consistency on the acquired frequencies and is optimized for partial Fourier acquisition in both phase encode directions chosen as 6/8 (Figs. 1–4).

This reconstruction approach leverages the key strength of unrolled

Table 1
Technical parameters of T1-weighted Volume Interpolated Breath-hold Examination (VIBE) Sequences. DL: deep learning, SD: standard, VIBE sequences. ms: milliseconds, Hz: Hertz, mm: millimeters, s: seconds. CAIPIRINHA: Controlled Aliasing in Parallel Imaging Results in Higher Acceleration, PE: Phase Encoding direction, 3D: Three-dimensional.

Parameter	T1 VIBE$_{DL}$	T1 VIBE$_{SD}$
Orientation	Coronal	Coronal
Acquisition type	Breath-hold	Breath-hold
Repetition time (ms)	3.5	3.9
Echo time (ms)	1.3	1.7
Flip angle (degree)	9	9
Pixel bandwidth (Hz/pixel)	256	370
Field of view (mm^2)	325 × 400	350 × 400
Matrix	288	384
Slice thickness (mm)	1.5	3
Number of slices	180	82
Phase Resolution	0.75	0.7
Slice Resolution	0.5	0.5
Phase Partial Fourier	6/8	7/8
Slice Partial Fourier	6/8	6/8
Acquired voxel size (mm^3)	1.9 × 1.4 × 3	1.5 × 1 × 6
Reconstructed voxel size (mm^3)	0.7 × 0.7 × 1.5	1 × 1 × 3
Echo train length	1	1
Acceleration mode	DL-enhanced CAIPIRINHA	CAIPIRINHA
Acceleration PE x 3D	4 × 2	2 × 2
Acquisition time (s)	10.2	22.3

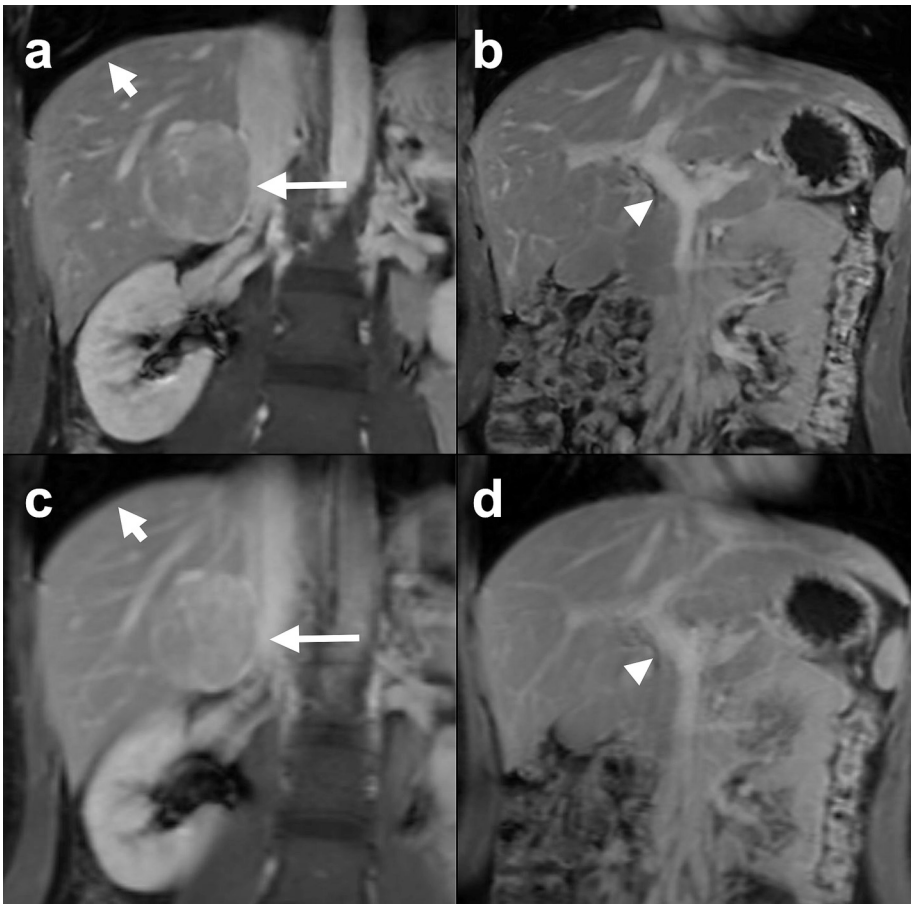

Fig. 1. Clinical Example 1. 38 y.o. Female presenting for further characterization of liver mass. Coronal post-contrast $VIBE_{DL}$ images (a-b) demonstrate better lesion conspicuity of a well circumscribed hypervascular lesion (long arrows), liver edge sharpness (short arrow), and vessel sharpness (arrowheads) compared to $VIBE_{SD}$ (c-d).

variational networks: their ability to directly incorporate MRI physics of the data acquisition scheme into the neural network architecture. By unrolling iterative optimization algorithms into neural network blocks, this method can effectively reconstruct images from undersampled k-space data while maintaining fidelity to the acquired data and leveraging learned image characteristics to compensate for undersampling artifacts, making it particularly well-suited for accelerated VIBE sequences.

2.4. Image analysis

Three board-certified abdominal radiologists (**, **, and **, with 6, 8, and 5 years of experience, respectively) performed independent visual evaluations of coronal $VIBE_{DL}$ and $VIBE_{SD}$ images. The assessment was conducted using a standardized four-point Likert scale for the following criteria, with particular emphasis on liver-specific parameters: (a) overall image quality, (b) perceived contrast-to-noise ratio, (c) severity of imaging artifacts, (d) liver edge sharpness, (e) liver vessel sharpness, and (f) conspicuity of up to two largest focal lesions, when present. A detailed description of the Likert scale criteria is provided in Supplemental Table S1. In cases where focal lesions were identified, lesion size and location of the two largest lesions were documented. The reference standard for focal liver lesion assessment was established using the complete MRI protocol, excluding the two coronal VIBE sequences being evaluated. For each patient, the radiologists performed a side-by-side comparison of $VIBE_{DL}$ and $VIBE_{SD}$ sequences. All patient identifiers and sequence labels were removed from the images prior to analysis. Readers were blinded to the acquisition technique and to each other's assessments. The order of image presentation was randomized on a per-patient basis to mitigate potential order effects. The visual analysis was conducted on a high-resolution, DICOM-calibrated workstation equipped with dedicated medical imaging software (Visage Client Version 7.1.18, Visage Imaging, Inc., San Diego, CA).

2.5. Statistical analysis

Descriptive summaries were generated for the entire cohort, including the median and interquartile range for continuous variables and the frequency (proportion) for categorical/ordinal variables. Contingency tables were generated for all pairs of ratings and are provided in Supplemental Table S2. Interrater agreement was assessed using Gwet's agreement coefficient version 2 (AC2) using ordinal weights [1 = x, 4 = y].[20] Benchmarking using Altman's rules was applied to characterize the agreement (0.8–1.0 = Very Good, 0.6–0.8 = Good, 0.4–0.6 = Moderate, 0.2–0.4 = Fair, −1.0-0.2 = Poor).[21] Benchmark labels were assigned using the lower 95 % confidence interval estimate of the agreement coefficient. Ordinal cumulative logit mixed-effects regression models were constructed to quantify the association between rating and reconstruction method (DL vs Standard). Rater and lesion number (conspicuity only) were also included as fixed effects, while patient identifiers were accounted for using separate random effects (intercept-only). Each model was summarized using the exponentiated coefficient (odds ratio, OR) associated with the reconstruction method, its 95 % confidence interval (CI), and *p*-value. These estimates may be interpreted as the fold-change in the odds of having a higher rating when comparing $VIBE_{DL}$ to $VIBE_{SD}$ (reference). Rater-specific analyses were

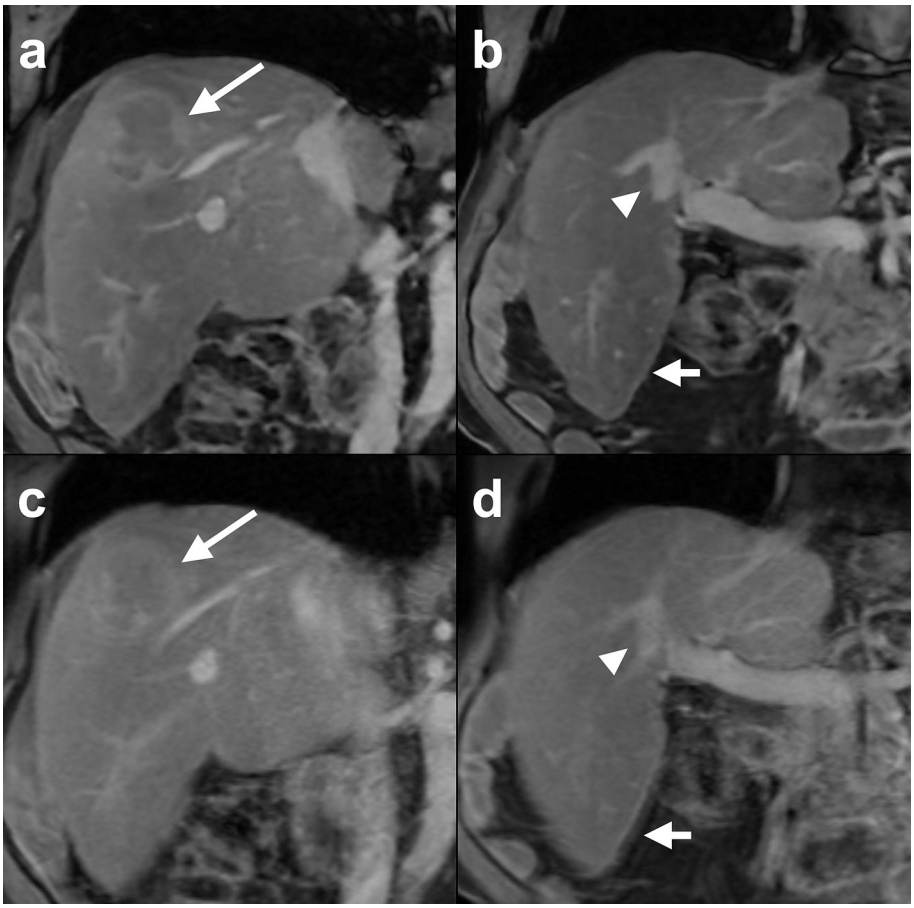

Fig. 2. Clinical Example 2. 68 y.o. Male with known chronic liver disease (CLD) and hepatocellular carcinoma (HCC). Coronal post-contrast $VIBE_{DL}$ images (a-b) demonstrate superior lesion conspicuity of a well-circumscribed lesion with washout and pseudo capsule (long arrows), superior liver edge sharpness (short arrow) and vessel sharpness (arrowheads) compared to $VIBE_{SD}$ (c-d).

also performed, and the modeling approach was akin to that described above except that rater was omitted from the model formulation. In addition to ORs, the proportion of times the $VIBE_{DL}$ ranking was greater than, equal to, and less than the corresponding $VIBE_{SD}$ measurement was reported. All analyses were performed using R version 4.4.1.

3. Results

A total of 151 patients were included in the study (57.1 ± 17.6 years, 54.3 % female). A total of 11 studies were excluded from the analysis: 5 due to the use of hepatobiliary contrast media and 6 due to incomplete MRI datasets. Acquisition times were 10.2 s for $VIBE_{DL}$ compared to 22.3 s for $VIBE_{SD}$. The most common reasons for examination were abnormal pancreas (26.5 %), chronic liver disease surveillance (13.3 %), and characterization of liver lesions (15.2 %). Focal liver lesions were identified in 60.3 % of patients, with 28.5 % having one lesion and 31.8 % having at least two lesions. The median lesion size was 15 mm (IQR:19). Detailed patient demographics are presented in Table 2.

3.1. Qualitative assessment

Comparison of $VIBE_{DL}$ and $VIBE_{SD}$ revealed significant differences across several image quality parameters (Table 3). Unadjusted regression summaries by raters can be found in Supplemental Table S3. $VIBE_{DL}$ demonstrated superior overall image quality, with 55 % of cases assigned an equal image quality score and 33.8 % of cases receiving a higher image quality score (OR 1.95, 95 % CI: 1.44–2.65, $p < 0.001$). Furthermore, $VIBE_{DL}$ showed higher odds of reduced image noise (OR 3.02, 95 % CI: 2.26–4.05, $p < 0.001$), enhanced liver edge sharpness (OR 3.68, 95 % CI: 2.63–5.15, $p < 0.001$), and improved liver vessel sharpness (OR 4.43, 95 % CI: 3.13–6.27, $p < 0.001$) compared to $VIBE_{SD}$. Notably, lesion conspicuity for focal lesions was substantially improved with $VIBE_{DL}$ (OR 9.03, 95 % CI: 6.34–12.85, $p < 0.001$). Interestingly, $VIBE_{DL}$ showed significantly lower odds of artifact reduction compared to $VIBE_{SD}$ (OR 0.13, 95 % CI: 0.09–0.18, $p < 0.001$). In terms of lesion detection, $VIBE_{DL}$ identified 137/138 (99.3 %) focal liver lesions, while $VIBE_{SD}$ detected 131/138 (94.9 %). The single lesion undetectable on $VIBE_{DL}$ measured 15 mm, whereas the median size of missed lesions on $VIBE_{SD}$ was 4 mm (IQR: 1) (Figs. 1–4). Due to the low number of non-detectable lesions, formal inferential comparisons were not conducted.

Inter-rater agreement (Table 4) for $VIBE_{SD}$ and $VIBE_{DL}$ ranged from good to very good across all evaluated parameters. For $VIBE_{SD}$, Gwet's AC2 ranged from 0.78 (95 % CI: 0.76–0.80) for artifacts to 0.87 (95 % CI: 0.85–0.89) for overall image quality. For $VIBE_{DL}$, Gwet's AC2 ranged from 0.82 (95 % CI: 0.78–0.86) for noise to 0.95 (95 % CI: 0.92–0.98) for lesion conspicuity.

4. Discussion

This prospective study demonstrates the efficacy of an accelerated T1 VIBE sequence with deep learning-based reconstruction for post-contrast enhanced coronal imaging of the upper abdomen. Qualitative evaluations, with a focus on liver-specific parameters, revealed that $VIBE_{DL}$ significantly outperformed $VIBE_{SD}$ in overall image quality and diagnostic performance, perceived SNR, liver edge sharpness, liver vessel delineation, lesion detection, and lesion conspicuity. Importantly,

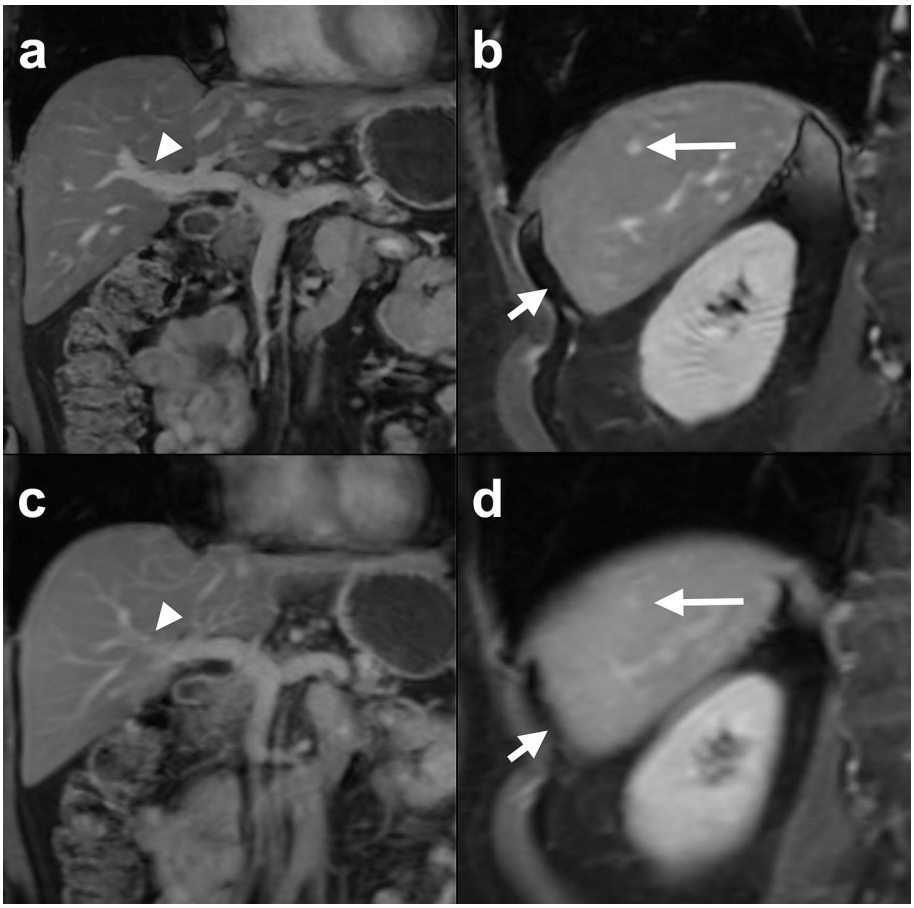

Fig. 3. Clinical Example 3. 55 y.o. Female with known hemangioma. Coronal post-contrast $VIBE_{DL}$ images (a-b) demonstrates superior conspicuity of a 10 mm hypervascular lesion (long arrows), liver edge sharpness (short arrow) and vessel sharpness (arrowheads) compared to $VIBE_{SD}$ (c-d).

these enhancements were achieved with a significantly reduced acquisition time.

Our findings align with and extend previous research on DL-reconstruction algorithms for 3D GRE sequences, further validating the efficacy of DL super-resolution reconstruction in abdominal MRI. Several studies have demonstrated significant enhancements in image quality, sharpness, lesion conspicuity, and diagnostic confidence across various imaging phases.[11–16,22,23] Gassenmaier et al.[11] and Almansour et al.[12] reported improved image quality and diagnostic confidence using iterative denoising and image enhancement techniques. Building on this, Afat et al.[15] and Wessling et al.[16] investigated DL-based super-resolution algorithms tailored to partial Fourier acquisitions, demonstrating not only improvements in image quality but also the potential for scan time reduction. In the context of liver imaging, Wei et al.[13] and Takayama et al.[22] both demonstrated superior image quality and enhanced lesion detection using DL reconstruction techniques in gadoxetic acid-enhanced MRI on hepatobiliary phase. Yun et al.[23] extended these findings to multi-arterial phase imaging, reporting improvements in both qualitative and quantitative parameters for small hypervascular hepatic masses. Brendel et al.[14] showed that DL reconstruction could facilitate a 50 % reduction in breath-hold duration while maintaining or improving image quality in upper abdominal MRI.

Our study distinguishes itself from prior research in several key aspects. Many previous investigations were limited by smaller sample sizes (32–55 patients),[11,12,14–16,22,23] retrospective simulation of scan time reductions,[11,12,15,16] focus on 1.5 Tesla scanners[14–16] and/or use of hepatobiliary contrast media.[13,22,23] In contrast, our study leverages a substantially larger patient cohort and employs a prospective acquisition with extracellular contrast media-enhanced $VIBE_{DL}$ on a 3 T scanner. The findings demonstrate both reproducibility and clinical applicability of DL reconstruction techniques, while capturing the full spectrum of clinical scenarios encountered in daily practice.

Our study demonstrated that the $VIBE_{DL}$ sequence significantly outperformed the standard $VIBE_{SD}$ in terms of image quality, while substantially reducing acquisition time from 22.3 to 10.2 s. This improvement stems from the sophisticated DL reconstruction algorithm, which employs an unrolled variational network architecture for initial k-space to image conversion, followed by a super-resolution step. The $VIBE_{DL}$ images exhibited markedly enhanced liver edge definition and hepatic vessel delineation. These improvements can be attributed to the DL algorithm's ability to mitigate noise and the super-resolution technique enabling a reduction in slice thickness from 3 mm to 1.5 mm. Consequently, the perceived signal-to-noise ratio was also significantly higher in $VIBE_{DL}$, contributing to the overall superior image quality.

Lesion conspicuity and detection rates also demonstrated substantial improvement with $VIBE_{DL}$. The combination of enhanced image sharpness and reduced partial volume averaging due to thinner slices resulted in better visualization of focal liver lesions. $VIBE_{DL}$ detected 137 out of 138 lesions (99.3 %), compared to 131 lesions (94.9 %) with $VIBE_{SD}$, with a median size of 4 mm for lesions missed on $VIBE_{SD}$ but detected on $VIBE_{DL}$, suggesting superior detectability of $VIBE_{DL}$ for subcentimeter lesions. Such enhanced lesion detection capability has significant clinical implications, potentially enabling earlier diagnosis, more accurate staging, and improved treatment planning.

Interestingly, in our study $VIBE_{DL}$ images showed a higher prevalence of peripheral Moiré fringe artifacts, commonly associated with GRE sequences. This finding contrasts with several previous studies reporting reduced artifacts with DL reconstruction, although most

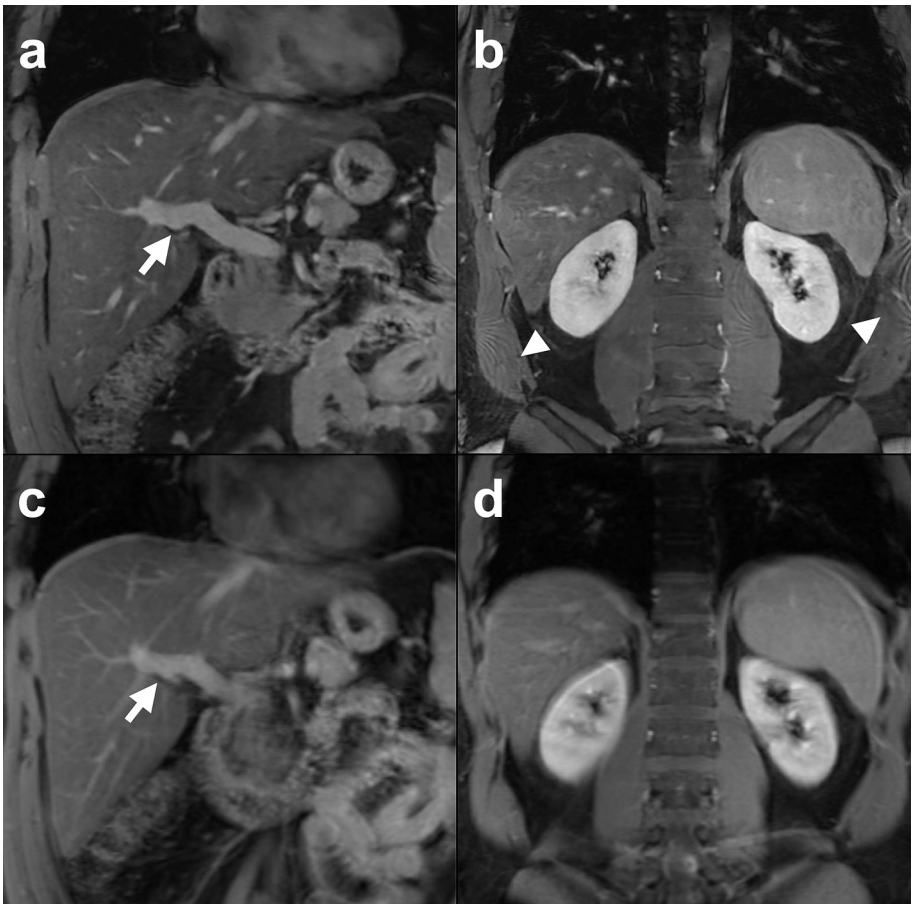

Fig. 4. Moiré Fringe Artifacts in VIBE-DL Reconstruction. Example of common peripheral Moiré fringe artifacts (arrowheads), more pronounced on the coronal post-contrast $VIBE_{DL}$ images (a-b) than on $VIBE_{SD}$ (c-d). However, the overall image quality and diagnostic confidence were not compromised in this case, with sharper vessel delineation (short arrows) on $VIBE_{DL}$ versus $VIBE_{SD}$.

Table 2
Patient demographics. Data is presented as median (interquartile range) for continuous variables or n (%) for categorical variables. BMI: Body Mass Index, CLD: Chronic Liver Disease, HCC: Hepatocellular Carcinoma, s/p OLT: status post Orthotopic Liver Transplantation, MR: Magnetic Resonance, MRCP: Magnetic Resonance Cholangiopancreatography, mm: millimeters.

Patient characteristics	($N = 151$)
Age, years	62 (28)
Female	82 (54.3 %)
BMI	27.6 (6.9)
Reason for exam, n (%)	
Abnormal pancreas	40 (26.5 %)
CLD, HCC screening	20 (13.2 %)
Assessment for metastases	19 (12.6 %)
Evaluation of liver lesion	23 (15.2 %)
Cholestatic liver disease	10 (6.6 %)
s/p OLT	4 (2.6 %)
Intrahepatic cholangiocarcinoma	2 (1.3 %)
Other	33 (21.9 %)
Exam type, n (%)	
MR liver	93 (62.0 %)
MR liver including MRCP	57 (38.0 %)
Number of assessed lesions, n (%)	
0	60 (39.7 %)
1	44 (29.1 %)
2	47 (31.1 %)
Lesion size, mm	15 (19)

Table 3
Unadjusted regression summary comparing deep-learning (DL) based and standard reconstruction methods. Estimates and 95 % confidence intervals (CI) from ordinal mixed effects models, with DL as the predictor and Non-DL as the reference. Models account for the rater as a fixed effect (and lesion for conspicuity), with the patient as a random effect. Estimates >1 indicate higher scores for DL compared to standard reconstruction. Est: Estimate, CI: Confidence Interval.

Variable	Est (95 % CI)	p-value
Overall image quality	1.95 (1.44, 2.65)	<0.001
Perceived signal-to-noise ratio	3.02 (2.26, 4.05)	<0.001
Severity of artifacts	0.13 (0.09, 0.18)	<0.001
Liver edge sharpness	3.68 (2.63, 5.15)	<0.001
Liver vessel sharpness	4.43 (3.13, 6.27)	<0.001
Lesion conspicuity	9.03 (6.34, 12.85)	<0.001

studies report motion-related artifacts.[11,13,14] It is worth noting that our study utilized an unmodified version of the DL model without specific artifact reduction algorithms. However, these artifacts did not significantly impact the overall image quality or diagnostic performance of $VIBE_{DL}$, as evidenced by the consistently superior ratings for overall image quality and diagnostic performance. The presence of these artifacts highlights the need for continued refinement of DL reconstruction techniques to address specific artifact types while maintaining substantial improvements in image quality and acquisition time.

Table 4
Inter-reader agreement for $VIBE_{SD}$ and $VIBE_{DL}$ sequences. Values represent Gwet's AC2 agreement coefficients with 95 % confidence intervals. Ordinal weights (1–4) were applied. Agreement benchmarks: Very Good (0.8–1.0), Good (0.6–0.8), Moderate (0.4–0.6), Fair (0.2–0.4), Poor (−1.0–0.2). Benchmark assignment based on lower bound of 95 % confidence interval. VIBE: Volume Interpolated Breath-hold Examination, DL: Deep Learning, SD: Standard.

Variable	$VIBE_{SD}$	$VIBE_{DL}$
Overall image quality	0.87 (0.85, 0.89)	0.89 (0.87, 0.91)
Perceived signal-to-noise ratio	0.80 (0.76, 0.83)	0.82 (0.78, 0.86)
Severity of artifacts	0.78 (0.76, 0.80)	0.92 (0.90, 0.94)
Liver edge sharpness	0.84 (0.81, 0.86)	0.91 (0.88, 0.94)
Liver vessel sharpness	0.84 (0.81, 0.87)	0.92 (0.89, 0.94)
Lesion 1 conspicuity	0.82 (0.78, 0.86)	0.94 (0.92, 0.97)
Lesion 2 conspicuity	0.85 (0.80, 0.88)	0.95 (0.92, 0.98)

4.1. Limitations

This study has several limitations. First, although our DL reconstruction network was trained on a diverse external dataset, the single-center design may limit generalizability to other institutions or MRI systems. Different scanners, field strengths, or institutional protocols could potentially yield varying results, underscoring the need for multi-center validation studies. Second, while we employed standardized evaluation criteria, the subjective nature of qualitative image assessment introduces potential variability between readers. This is reflected in the slightly varying inter-rater agreement observed for some parameters. The variability could be attributed to differences in reader experience, personal preferences in image interpretation, or the inherent complexity of evaluating certain image features. Third, we did not differentiate between various types of artifacts in our analysis, which could provide more nuanced insights into the specific strengths and weaknesses of the DL reconstruction approach. A more detailed categorization could help identify areas for focused improvement in the DL algorithm and provide a more comprehensive understanding of its performance across different imaging challenges.

Fourth, there was a difference in slice thickness between $VIBE_{DL}$ (1.5 mm) and $VIBE_{SD}$ (3 mm) sequences. While this difference represents a real advantage of the DL reconstruction technique in clinical application, it does introduce a variable that independently affects image quality, particularly regarding partial volume effects and lesion conspicuity.

Lastly, our study included patients with a wide range of clinical indications, which may have introduced heterogeneity in the results. While this diverse cohort reflects real-world clinical practice, it limited our ability to perform meaningful subgroup analyses based on specific pathologies or lesion types. Furthermore, while we demonstrated improved lesion detection with $VIBE_{DL}$, we only evaluated the two largest lesions per patient, which introduces a bias towards more conspicuous lesions. Additionally, we did not conduct a comprehensive analysis of lesion characterization or compare our findings to a gold standard such as histopathology. Such an analysis could provide further insights into the clinical impact of the improved image quality and lesion conspicuity offered by the DL reconstruction technique. Future prospective studies should focus on investigating whether improvements in image quality and lesion detection with DL reconstruction vary across different parenchymal conditions and lesion characteristics. Moreover, studies focusing on portal vein or inferior vena cava thrombosis would provide important insights into how DL reconstruction performs in complex hepatobiliary imaging, especially when evaluating patients with decompensated liver disease whose limited breath-holding ability poses significant diagnostic challenges.

5. Conclusion

In conclusion, this study provides compelling evidence for the clinical utility of deep learning-based reconstruction in abdominal MRI. The $VIBE_{DL}$ sequence offers a synergistic combination of improved image quality, enhanced lesion detection, and modestly reduced acquisition time, potentially enhancing the diagnostic capabilities of contrast-enhanced abdominal MRI. As DL algorithms continue to evolve, their integration into clinical workflows promises to further refine and augment the diagnostic performance of MRI in abdominal imaging.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.clinimag.2025.110581.

CRediT authorship contribution statement

Alexander Herold: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Nathaniel D. Mercaldo:** Writing – review & editing, Formal analysis. **Mark A. Anderson:** Writing – review & editing, Data curation. **Amirkasra Mojtahed:** Writing – review & editing, Data curation. **Aoife Kilcoyne:** Writing – review & editing, Data curation. **Wei-Ching Lo:** Writing – review & editing, Software, Methodology. **Robert M. Sellers:** Writing – review & editing, Software. **Bryan Clifford:** Writing – review & editing, Software. **Marcel D. Nickel:** Writing – review & editing, Software, Resources, Data curation. **Nabih Nakrour:** Writing – review & editing, Methodology. **Susie Y. Huang:** Writing – review & editing, Supervision, Project administration. **Leo L. Tsai:** Writing – review & editing, Supervision, Methodology, Investigation. **Onofrio A. Catalano:** Writing – review & editing, Supervision, Project administration, Methodology. **Mukesh G. Harisinghani:** Writing – review & editing, Supervision, Project administration, Methodology.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used "Claude 3.5 Sonnet AI Assistant. Anthropic, PBC" in order to improve the readability, language and quality of the writing. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

Wei-Ching Lo, Robert M. Sellers, Bryan Clifford are employees at Siemens Medical Solutions. Marcel D. Nickel is employee at Siemens Healthineers. Susie Y. Huang is a consultant for Siemens Healthineers. The remaining authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence.

References

1 Caraiani C, Yi D, Petresc B, Dietrich C. Indications for abdominal imaging: when and what to choose? J Ultrason 2020;20(80):e43–54. https://doi.org/10.15557/JoU.2020.0008.
2 Rofsky NM, Lee VS, Laub G, et al. Abdominal MR imaging with a volumetric interpolated breath-hold examination. Radiology 1999;212(3):876–84. https://doi.org/10.1148/radiology.212.3.r99se34876.
3 Markl M, Leupold J. Gradient echo imaging. J Magn Reson Imaging 2012;35(6):1274–89. https://doi.org/10.1002/jmri.23638.
4 Zaitsev M, Maclaren J, Herbst M. Motion artifacts in MRI: a complex problem with many partial solutions. J Magn Reson Imaging 2015;42(4):887–901. https://doi.org/10.1002/jmri.24850.
5 Block KTCH, Milla S, Bruno M, et al. Towards routine clinical use of radial stack-of-stars 3D gradient-echo sequences for reducing motion sensitivity. J Korean Soc Magn Reson Med 2014;18(2):87–106. https://doi.org/10.13104/jksmrm.2014.18.2.87.
6 Lee Y, Yoon S, Paek M, Han D, Choi MH, Park SH. Advanced MRI techniques in abdominal imaging. Abdom Radiol (NY) 2024;49(10):3615–36. https://doi.org/10.1007/s00261-024-04369-7.
7 Yoon JH, Nickel MD, Peeters JM, Lee JM. Rapid imaging: recent advances in abdominal MRI for reducing acquisition time and its clinical applications. Korean J Radiol 2019;20(12):1597–615. https://doi.org/10.3348/kjr.2018.0931.
8 Deshmane A, Gulani V, Griswold MA, Seiberlich N. Parallel MR imaging. J Magn Reson Imaging 2012;36(1):55–72. https://doi.org/10.1002/jmri.23639.

9 Feng L, Benkert T, Block KT, Sodickson DK, Otazo R, Chandarana H. Compressed sensing for body MRI. J Magn Reson Imaging 2017;45(4):966–87. https://doi.org/10.1002/jmri.25547.

10 Hammernik K, Klatzer T, Kobler E, et al. Learning a variational network for reconstruction of accelerated MRI data. Magn Reson Med 2018;79(6):3055–71. https://doi.org/10.1002/mrm.26977.

11 Gassenmaier S, Afat S, Nickel D, et al. Application of a novel iterative denoising and image enhancement technique in T1-weighted precontrast and postcontrast gradient echo imaging of the abdomen: improvement of image quality and diagnostic confidence. Investig Radiol 2021;56(5):328–34. https://doi.org/10.1097/RLI.0000000000000746.

12 Almansour H, Herrmann J, Gassenmaier S, et al. Combined deep learning-based super-resolution and partial Fourier reconstruction for gradient echo sequences in abdominal MRI at 3 tesla: shortening breath-hold time and improving image sharpness and lesion conspicuity. Acad Radiol 2023;30(5):863–72. https://doi.org/10.1016/j.acra.2022.06.003.

13 Wei H, Yoon JH, Jeon SK, et al. Enhancing gadoxetic acid-enhanced liver MRI: a synergistic approach with deep learning CAIPIRINHA-VIBE and optimized fat suppression techniques. Eur Radiol 2024;34(10):6712–25. https://doi.org/10.1007/s00330-024-10693-9.

14 Brendel JM, Jacoby J, Dehdab R, et al. Prospective deployment of deep learning reconstruction facilitates highly accelerated upper abdominal MRI. Acad Radiol 2024;(12). https://doi.org/10.1016/j.acra.2024.05.044.

15 Afat S, Wessling D, Afat C, et al. Analysis of a deep learning-based superresolution algorithm tailored to partial Fourier gradient Echo sequences of the abdomen at 1.5 T: reduction of breath-hold time and improvement of image quality. Investig Radiol 2022;57(3):157–62. https://doi.org/10.1097/RLI.0000000000000825.

16 Wessling D, Herrmann J, Afat S, et al. Application of a deep learning algorithm for combined super-resolution and partial Fourier reconstruction including time reduction in T1-weighted precontrast and postcontrast gradient echo imaging of abdominopelvic MR imaging. Diagnostics (Basel) 2022;12(10):2370. https://doi.org/10.3390/diagnostics12102370.

17 Paszke A, Gross S, Massa F, et al. Pytorch: An Imperative Style, High-performance Deep Learning Library. ArXiv, abs/191201703. 2019.

18 ONNX. Runtime developers. ONNX runtime 2021.

19 Almansour H, Gassenmaier S, Nickel D, et al. Deep learning-based superresolution reconstruction for upper abdominal magnetic resonance imaging: an analysis of image quality, diagnostic confidence, and lesion conspicuity. Investig Radiol 2021;56(8):509–16. https://doi.org/10.1097/RLI.0000000000000769.

20 Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. Br J Math Stat Psychol 2008;61(Pt 1):29–48. https://doi.org/10.1348/000711006X126600.

21 A DG. Practical Statistics for Medical Research. 1st ed. Chapman and Hall/CRC; 1990.

22 Takayama Y, Sato K, Tanaka S, Murayama R, Jingu R, Yoshimitsu K. Effectiveness of deep learning-based reconstruction for improvement of image quality and liver tumor detectability in the hepatobiliary phase of gadoxetic acid-enhanced magnetic resonance imaging. Abdom Radiol (NY) 2024;49(10):3450–63. https://doi.org/10.1007/s00261-024-04374-w.

23 Yun SM, Hong SB, Lee NK, et al. Deep learning-based image reconstruction for the multi-arterial phase images: improvement of the image quality to assess the small hypervascular hepatic tumor on gadoxetic acid-enhanced liver MRI. Abdom Radiol (NY) 2024;49(6):1861–9. https://doi.org/10.1007/s00261-024-04236-5.