Original Article

Deep learning-based reconstruction for 3-dimensional heavily T2-weighted fat-saturated magnetic resonance (MR) myelography in epidural fluid detection: image quality and diagnostic performance

Mingyu Kim[1], Jisook Yi[1]^, Ho-Joon Lee[1], Seok Hahn[1], Yedaun Lee[1], Joonsung Lee[2]

[1]Department of Radiology, Inje University College of Medicine, Haeundae Paik Hospital, Busan, Republic of Korea; [2]GE HealthCare Korea, Seoul, Republic of Korea

Contributions: (I) Conception and design: J Yi; (II) Administrative support: J Yi; (III) Provision of study materials or patients: J Yi, HJ Lee; (IV) Collection and assembly of data: J Yi, HJ Lee; (V) Data analysis and interpretation: M Kim, S Hahn, J Yi, HJ Lee; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jisook Yi, MD. Department of Radiology, Inje University College of Medicine, Haeundae Paik Hospital, 875, Haeun-daero, Haeundae-gu, Busan, Republic of Korea. Email: jsyi2010@gmail.com; jsyi2010@paik.ac.kr.

Background: Heavily T2-weighted fat-saturated (HT2W-FS) magnetic resonance myelography (MRM) is useful for diagnosing the cause of intracranial hypotension. Recently, deep learning-based reconstruction (DLR) has been utilized to improve image signal-to-noise ratios and sharpness while reducing artifacts, all without lengthening acquisition times. This study aimed to compare the diagnostic performance and image quality of conventional reconstruction (CR) and DLR of 3-dimensional (3D) HT2W-FS MRM applied to detecting epidural fluid in patients with clinically suspected intracranial hypotension.

Methods: This retrospective study included 21 magnetic resonance myelography examinations using both CR and DLR in 21 patients who experienced orthostatic headache between April 2021 and September 2022. Quantitative image quality evaluation was performed by comparing signal-to-noise ratios at the lower thoracic levels. The image quality and artifacts were graded by three readers. The presence of epidural fluid was reported with a confidence score by two readers, and the area under the receiver operating curve, interobserver agreement, and inter-image-set agreement were evaluated. The conspicuity of the dura mater where the epidural fluid was detected was also investigated.

Results: Quantitative and subjective image quality, and artifacts significantly improved with DLR (all P<0.001). Diagnostic performance of DLR was higher for both readers [reader 1: area under the curve (AUC) of CR =0.929; 95% confidence interval (CI): 0.902–0.950, AUC of DLR =0.965 (95% CI: 0.944–0.979), P=0.007; reader 2: AUC of CR =0.834 (95% CI: 0.798–0.866), AUC of DLR =0.877 (0.844–0.905), P=0.040]. Correlation with standard care of MRM in CR and DLR were both strong in reader 1 (rho =0.868–0.919, P<0.001), but was respectively strong and moderate in reader 2 (rho =0.734–0.805, P<0.001). Interobserver agreement was substantial (κ=0.708–0.762). The inter-image-set agreement was almost perfect for reader 1 (κ=0.907) and was substantial for reader 2 (κ=0.750). Dura mater conspicuity significantly improved with DLR (P<0.014, reader 1; P<0.001, readers 2 and 3).

Conclusions: HT2W-FS magnetic resonance myelography with DLR demonstrates substantial improvements in image quality and may improve confidence in detecting epidural fluid.

^ ORCID: 0000-0002-6633-0526.

Keywords: Intracranial hypotension; myelography; magnetic resonance imaging (MRI); deep learning

Submitted Mar 07, 2024. Accepted for publication Jul 09, 2024. Published online Aug 07, 2024.
doi: 10.21037/qims-24-455
View this article at: https://dx.doi.org/10.21037/qims-24-455

Introduction

Spontaneous cerebrospinal fluid (CSF) leaks are the most commonly identified cause of spontaneous intracranial hypotension, but can also result from other identifiable causes such as trauma or lumbar puncture (1-3). The most common clinical manifestation is headache, generally but not invariably orthostatic; other findings include vertigo, muffled hearing, nausea and emesis, disequilibrium, and cognitive impairment (1). Major causes of spontaneous intracranial hypotension have been identified: (I) posterior vertebral osteophytes/calcified disc herniation causing dural dehiscence, and (II) CSF-venous fistula, a recently described entity, rarely presents with extradural CSF collection (4,5).

Spinal imaging findings in spontaneous intracranial hypotension include epidural fluid collections, dilated epidural veins, a festooned appearance of the epidural venous plexus, dural enhancement, and a focal area of fluid signal intensity between the spinous process of C1–C2 (6). Although the optimal imaging protocol has not been established, the mainstays of spinal imaging include CT myelography with contrast, conventional spine magnetic resonance imaging (MRI), and MR myelography (MRM) with or without contrast (5). Although MRM without contrast administration is a static examination that has limited ability to delineate the dynamic process of CSF egress from the dural sac, it can be used to demonstrate CSF collections, especially when acquired with heavily T2-weighted (HT2W) MRI, a fluid-sensitive sequence (7,8). HT2W sequences without intrathecal contrast were shown to be comparable to CT myelography and noninferior to MRM with intrathecal contrast in detecting extradural CSF collections, and appear to be ubiquitous across multiple studies on MRM (8-11).

Owing to the subtleties of MR imaging findings, achieving adequate spatial resolution is important in suspected cases (5). Recently, deep learning-based reconstruction (DLR) has been utilized to improve image signal-to-noise ratios and sharpness while reducing artifacts, all without lengthening acquisition times (12). It has been applied to imaging of the brain, shoulder, knee, abdomen, and spine (13-17). However, to the best of our knowledge, no studies have evaluated the application of DLR to epidural fluid. This study aimed to compare the diagnostic performance and image quality of conventional reconstruction (CR) and DLR of 3-dimensional (3D) HT2W MRM applied to detecting However, to the best of our knowledge, no studies have evaluated the application of DLR to epidural fluid in patients with clinically suspected intracranial hypotension. We present this article in accordance with the STARD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-24-455/rc).

Methods

Participants

We searched for consecutive adult patients who underwent 3-T HT2W MRM due to orthostatic headache between April 2021 and September 2022 in a single center (Haeundae Paik Hospital); a total of 31 examinations in 29 patients were initially included. The exclusion criteria were as follows: (I) 3D images not acquired in the sagittal plane (n=1), (II) loss of k-space data for DLR (n=4), (III) insufficient data resulting in insufficient coverage of the spine (n=3), and (IV) poor image quality (n=2). A total of 21 examinations with both CR and DLR in 21 patients were included, resulting in 483 intervertebral disc levels from C2–C3 to L5–S1 in the final study cohort (*Figure 1*). Archived records were manually reviewed to collect the following data: sex, age, brain MRI obtained within 1 month of MRM, prior history suggesting violation of the dura mater, mean duration of headache, history of epidural blood patch, and clinical outcomes after epidural blood patch. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of Inje University Haeundae Paik Hospital (No. HPIRB 2022-10-010) and individual consent for this retrospective analysis was waived.

MRI protocol

MRM was performed at a single institution using a 3-T MR

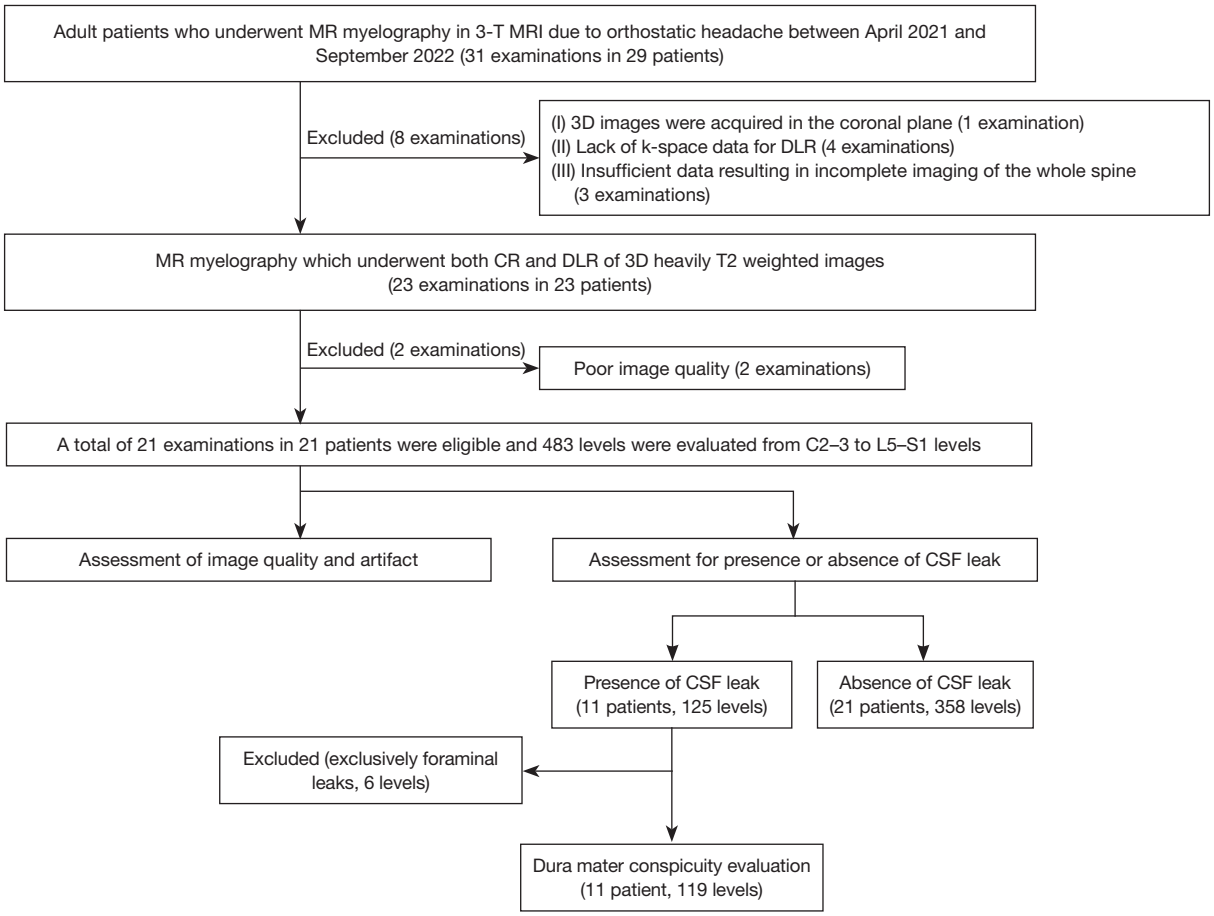

Figure 1 Flowchart of patient inclusion and image evaluation. MR, magnetic resonance; MRI, magnetic resonance imaging; 3D, three-dimensional; DLR, deep learning-based reconstruction; CR, conventional reconstruction; CSF, cerebrospinal fluid.

system (Signa Architect, GE Healthcare) with a 16-channel head and neck coil for the cervical spine and two 30-channel anterior array coils for the posterior aspect of the spine. Two-dimensional single-shot fast spin-echo heavily T2-weighted fat-saturated (HT2W-FS) images were obtained in the axial plane with 4–8 stacks and 3D fast spin-echo HT2W-FS images were obtained in the sagittal plane with 3 stacks. The detailed MR acquisition parameters are listed in *Table 1*.

3D HT2W-FS MRM underwent both CR and DLR using a commercially available DLR pipeline. A vendor-supplied prototype of a DLR pipeline (AIR^{TM} Recon DL 3D (DLR); GE HealthCare, Waukesha, WI, USA) was employed for 3D image reconstruction, utilizing an artificial intelligence-enhanced image reconstruction pipeline (17,18). This reconstruction process enables denoising, reduces ringing artifacts, and performs interpolation in three-directions, resulting in a sharp image volume by eliminating truncation artifacts (12). The DLR pipeline incorporates a deep convolutional neural network (CNN) that was trained using a supervised learning approach with pairs of pristine and typical datasets. The DLR algorithm has a modifiable noise reduction level (low, medium, and high), and a high reduction level was used in our institution. The two-dimensional HT2W-FS MRM also underwent DLR with a high noise reduction level. Both CR and DLR images of the 3D HT2W-FS MRM were reconstructed in the axial plane with a slice thickness of 2 mm and a slice interval of 1 mm using commercially available software (Aquarius iNtuition, version 4.4.11, TeraRecon).

Image analysis

A representative axial image of 3D H2TW-FS MRM with CR and DLR was selected for each intervertebral disc level

Table 1 MRI parameters

Parameters	2D HT2W-FS SSFSE (axial)	3D HT2W-FS FSE (sagittal)
Repetition time/echo time (ms)	4,000/551.5	2,500/736.6
FOV (mm^2)	180×180	280×98
FOV method	–	Outer volume suppression
Section thickness/gap (mm)	6/0	1/0
Acquisition matrix	300×240	280×98
Echo train length	N/A	160
Number of excitation	1	1
Parallel imaging factor	–	2×1
Number of stacks	4–8	3
Noise reduction factor	High	High
Acquisition time/stack	1 minute 18 seconds	1 minute 12 seconds

MRI, magnetic resonance imaging; 2D, two-dimensional; HT2W-FS, heavily T2-weighted fat-saturated image; SSFSE, single-shot fast spin-echo; 3D, three-dimensional; FSE, fast spin-echo; FOV, field of view; N/A, not applicable.

from C2–C3 to L5–S1, resulting in a total of 23 disc levels per image set. The diagnostic confidence of CSF leaks and subjective image quality/artifacts were evaluated based on the pre-selected image for each level.

CSF leak detection

The presence of CSF leaks was established by consensus between two experienced radiologists (a musculoskeletal radiologist with 8 years of subspecialty experience, and a neuroradiologist with 8 years of subspecialty experience) based on three image sets (two-dimensional HT2W-FS MRM with DLR, 3D HT2W-FS MRM with DLR, and 3D HT2W-FS MRM with CR) and they were blinded to clinical information and radiologic reports. Evaluation of the diagnostic performance of 3D HT2W-FS MRM with CR and DLR was done by two readers (reader 1, a musculoskeletal radiologist with 9 years of subspecialty experience, and reader 2, a 1st-year resident with prior training in spinal imaging). The readers independently evaluated both image sets for the presence of epidural fluid at all intervertebral disc levels from C2–C3 to L5–S1 without any clinical information and radiologic reports. For each disc level, five slices, including the preselected representative axial slice and two slices cranial and caudal to the preselected slice, were evaluated. A 2-week rest period was implemented between the interpretation of 3D HT2W-FS MRM images with CR and with DLR. Epidural fluid was defined as (I) a fluid collection of presumed epidural location, often with a linear low signal intensity structure representing the dura, (II) fuzzy fluid extension at the neural foramen along the nerve root, or (III) a combination of the above. The presence or absence of an epidural fluid was reported with its diagnostic confidence using a 5-point scale; (I) definitely absent, (II) probably absent, (III) equivocal, (IV) probably present, and (V) definitely present. Diagnostic confidence scores were dichotomized with a cut-off for the presence of epidural fluid between confidence scores of 4 and above (present) and 3 and below (absent).

Image quality assessment

For the quantitative analysis of image quality, circular regions of interest (ROIs) were drawn in the bone marrow of the lower thoracic vertebrae, thoracic spinal cord, and CSF on both CR and DLR images of 3D HT2W-FS MRM (Figure S1). Based on these measurements, the signal-to-noise ratio (SNR) was calculated for both image sets using the following formula (19,20):

$$\text{SNR} = \frac{\text{mean signal intensity of CSF}}{\text{mean standard deviation of bone marrow signal intensity}} \quad [1]$$

The subjective image quality and artifacts of both image sets were evaluated by three readers (readers 1, 2, and 3; reader 3, a musculoskeletal radiologist with 8 years of subspecialty experience). Prior to evaluation, the readers were given standardized instructions and trained on image sets from five patients not included in this study (Figure S2). Readers independently evaluated the preselected axial images in both CR and DLR image sets from 3D HT2W-FS MRM, with a 2-week rest period between the CR and DLR series to reduce recall bias. Overall image quality was subjectively graded using a 5-point Likert scale as follows: (I) unacceptable diagnostic quality, (II) subdiagnostic, (III) acceptable diagnostic quality, (IV) good, (V) excellent (21). The presence of artifacts such as Gibbs, motion, CSF pulsation, and incomplete fat saturation was also subjectively graded using a 5-point scale: (I) extreme, resulting in unreadable images; (II) severe artifacts affecting the diagnosis of epidural fluid; (III) moderate artifacts slightly affecting the

diagnosis of epidural fluid; (IV) mild artifacts not affecting the diagnosis of epidural fluid; and (V) no artifacts noted.

For each level with epidural fluid according to the standard care, dura mater conspicuity was additionally evaluated by three readers (readers 1, 2, and 3) using a 5-point scale; 1, not visible or barely visible; 2, partially delineated with blurring; 3, partially but sharply delineated; 4, fully delineated with blurring; and 5, fully and sharply delineated. Levels with foraminal CSF leaks and without other epidural fluid collection were excluded from the evaluation of dura mater conspicuity.

Statistical analysis

Patient characteristics were presented as numbers, means, and standard deviations. The normality of the data was assessed using the Shapiro-Wilk test. The sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of epidural fluid detection were calculated for 3D H2TW-FS MRM with CR and DLR. McNemar's test was performed to compare the concordance of two image sets compared to standard care (22). Receiver operating characteristic curves and the area under the receiver operating characteristic curves of CR and DLR images were compared for both readers to compare the discriminatory power of the two image reconstruction algorithms (23). Spearman rank correlation was employed to evaluate the correlation with standard of care in MRM usage for detecting epidural fluid in 3D HT2W-FS MRM with CR and 3D MRM with DLR at both the individual spinal level and the patient level. The Spearman rank correlation was interpreted as none, weak, moderate, and strong for coefficients of less than 0.2, 0.21–0.50, 0.51–0.80, >0.81 (24). Interobserver agreement and inter-image-set agreement in identifying CSF leaks was assessed with the Fleiss-Cohen unweighted kappa and were interpreted as slight, fair, moderate, substantial, or excellent for kappa coefficients of <0.20, 0.21–0.40, 0.41–0.60, 0.61–0.80, and 0.81–1.0, respectively (25). Subjective scores of image quality and artifacts were compared between CR and DLR for all readers using Mann-Whitney U-tests. Visualization of the dura mater was compared between CR and DLR for all readers using the Wilcoxon signed-rank test.

All statistical analyses were performed using SPSS software version 25 (IBM Corp., Armonk, NY, USA) and MedCalc Software Version 20.210 (Ostend, Belgium), and a P value <0.05 in a two-tailed test was deemed statistically significant.

Results

Demographics and patient characteristics

Ten men (47.6%) and 11 women (52.4%) with a mean age of 39.9 years (range, 25–82 years) were included. Among them, 20 patients underwent brain MRI within one month of MRM, and three of the 20 patients (15.0%) showed intracranial imaging findings of intracranial hypotension according to Bern score: subdural fluid collections, enhancement of the pachymeninges/dura mater, cerebellar tonsillar herniation, venous engorgement, suprasellar, prepontine, and mamillopontine distance. The detailed characteristics of the enrolled patients are shown in *Table 2*.

Diagnostic performance of epidural fluid detection

Dichotomized analyses of epidural fluid presence on 3D HT2W-FS MRM with CR and DLR by both readers 1 and 2 were compared to the standard care of MRM usage. According to the standard case of MRM usage, epidural fluid was detected in 11 subjects, with a mean of 11.4 levels per patient with CSF leaks. On 3D H2TW-FS MRM with CR, reader 1 reported epidural fluid in 12 patients, with a mean of 9.5 levels per patient with epidural fluid, and reader 2 reported epidural fluid in 13 patients, with a mean of 7.3 levels per patient with epidural fluid. On 3D H2TW-FS MRM with DLR, reader 1 reported epidural fluid in 11 patients, with a mean of 11.3 levels per patient with epidural fluid, and reader 2 reported epidural fluid in 16 patients, with a mean of 16 levels per patient. Sensitivity, specificity, and diagnostic accuracy were higher for DLR than for CR for both readers (*Table 3*). Positive and negative predictive values also improved with DLR compared with CR for both readers. The areas under the receiver operating curve for diagnosing epidural fluid are shown in *Table 4* and *Figure 2*. The areas for the DLR images were significantly higher than those for the CR images for both readers (P=0.007, 0.040). The diagnostic performance of reader 1 (P=0.049) was significantly different between the two image sets of 3D HT2W-FS MRM, whereas that of reader 2 was not (P=0.27).

Significant and strong positive correlation with standard of care in MRM usage were observed in 3D HT2W-FS MRM with CR and 3D HT2W-FS MRM with DLR for reader 1 at both the individual spinal level (CR, rho =0.868,

Table 2 Patient characteristics

Characteristics	Value (n=21)
Sex	
Male	10 (47.6)
Female	11 (52.4)
Age (years)	39.9±14.39
Brain MRI obtained within 1 month of MR myelography	20 (95.2)
Evidence of intracranial hypotension on brain MRI	3/20 (15.0)
History prior to current symptoms suggestive of dura mater violation	11 (52.4)
Nerve blocks in thoracolumbar spine	5 (23.8)
Spinal "injections"	2 (9.5)
Spinal operation for spinal stenosis	1 (4.8)
Percutaneous epidural neuroplasty*	1 (4.8)
Spinal trauma	1 (4.8)
Lumbar puncture	2 (9.5)
Duration of headache (days)	51.86±106.91
Patients who underwent EBP for current symptoms	12 (57.1)
Patients with remote history of EBP	3 (14.3)
Clinical outcome after EBP for current symptoms†	n=12
No improvement of symptoms	2 (9.5)
Partial improvement of symptoms	5 (23.8)
Complete improvement of symptoms	4 (19.0)
Failed procedure	1 (4.8)

Data are presented as n (%) or mean ± standard deviation. *, a single patient had prior history of both nerve block and percutaneous epidural neuroplasty; †, for patients who underwent multiple series of EBP, the final outcome was documented in the above table. MRI, magnetic resonance imaging; MR, magnetic resonance; EBP, epidural blood patch.

P<0.001; DLR, rho =0.919, P<0.001) and the patient level (CR, rho =0.908, P<0.001; DLR, rho =1, P<0.001). Also, there was significant and strong positive correlation between standard of care in MRM usage and 3D HT2W-FS MRM with DLR for reader 2 at the individual spinal level (DLR, rho =0.805, P<0.001). There was moderate significant correlation between standard of care in MRM usage and 3D HT2W-FS MRM with CR for reader 2 at the individual spinal level (CR, rho =0.734, P<0.001). But, there was no statistically significant correlation between standard of care MRM and 3D HT2W-FS MRM with both image reconstruction methods for reader 2 at the patient level (P=0.052–0.106). The correlations to the epidural fluid diagnosis according to "standard of care in MRM usage" is shown in *Table 5*.

Interobserver and inter-image-set agreement

Interobserver agreement between the two readers was substantial in both CR and DLR images (CR, κ=0.708; DLR, κ=0.762). Inter-image-set agreement between CR and DLR was excellent for reader 1 (κ=0.907) and substantial for reader 2 (κ=0.750).

Image quality assessment

The SNR of 3D HT2W-FS MRM with DLR (mean SNR =143.68, standard deviation =55.86) was significantly higher than with CR (mean SNR =90.27, standard deviation =31.57; P<0.001). The subjective scores of the qualitative assessments for each reader are listed in the *Table 6*. The mean subjective scores for image quality improved with DLR compared with CR for all three readers, with statistical significance (P<0.001; *Figure 3*). The mean subjective scores for artifacts also improved with DLR compared with CR for all three readers, with statistical significance (P<0.001). The conspicuity of the dura mater in segments with epidural fluid also improved with DLR compared to CR for all three readers (reader 1, P=0.014; readers 2 and 3, P<0.001; *Table 6*, *Figure 4*).

Discussion

In the current study, we compared DLR and CR images of 3D HT2W-FS MRM in terms of image quality, artifacts, and diagnostic performance in detecting epidural fluid in patients with clinically suspected intracranial hypotension. 3D HT2W-FS MRM with DLR demonstrated improved diagnostic performance in detecting epidural fluid compared with CR. Interobserver agreement was substantial, and the inter-image-set agreement was substantial to excellent. 3D HT2W-FS MRM with DLR showed an improvement over CR in both quantitative and subjective grading of image quality and artifacts as well as in dural conspicuity at levels at which epidural fluid were present.

Table 3 Sensitivity, specificity, positive and negative predictive values and accuracy of 3D HT2W-FS with CR or with DLR in the detection of epidural fluid for both readers

Parameters	Reader 1		Reader 2	
	3D HT2W-FS with CR	3D HT2W-FS with DLR	3D HT2W-FS with CR	3D HT2W-FS with DLR
Sensitivity (95% CI)	0.84 (0.75–0.90)	0.95 (0.90–0.98)	0.70 (0.61–0.78)	0.78 (0.69–0.85)
Specificity (95% CI)	0.94 (0.91–0.96)	0.98 (0.96–0.99)	0.97 (0.95–0.99)	0.98 (0.96–0.99)
PPV (95% CI)	0.77 (0.69–0.83)	0.94 (0.88–0.97)	0.90 (0.82–0.94)	0.92 (0.86–0.96)
NPV (95% CI)	0.96 (0.94–0.97)	0.98 (0.96–0.99)	0.90 (0.88–0.92)	0.93 (0.90–0.95)
Accuracy (95% CI)	0.92 (0.89–0.94)	0.97 (0.95–0.98)	0.90 (0.87–0.93)	0.93 (0.90–0.95)

3D, three-dimensional; HT2W-FS, heavily T2-weighted fat-saturated image; CR, conventional reconstruction; DLR, deep learning-based reconstruction; PPV, positive predictive value; NPV, negative predictive value.

Table 4 Comparison of the AUC values of 3D HT2W-FS with CR or DLR for both readers

Readers	3D HT2W-FS with CR	3D HT2W-FS with DLR	Difference*	P value
Reader 1 (95% CI)	0.929 (0.902–0.950)	0.965 (0.944–0.979)	0.036 (0.010–0.062)	0.007
Reader 2 (95% CI)	0.834 (0.798–0.866)	0.877 (0.844–0.905)	0.043 (0.002–0.084)	0.040

*, data are mean differences in AUC between the 3D HT2W-FS with DLR group and the 3D HT2W-FS with CR group. P values for the comparison of AUC were calculated using DeLong's method. AUC, area under the curve; 3D, three-dimensional; HT2W-FS, heavily T2 weighted fat-saturated image; CR, conventional reconstruction; DLR, deep learning-based reconstruction.

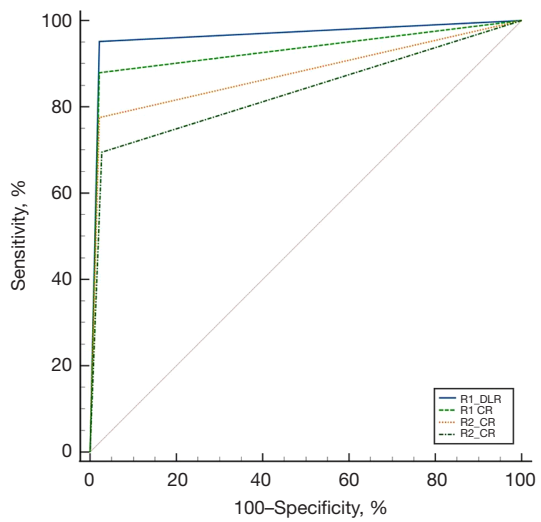

Figure 2 Receiver operating characteristic curves of CR and DLR for both readers. R1, reader 1; R2, reader 2. DLR, deep learning-based reconstruction; CR, conventional reconstruction.

3D MRI shows a high through-plane spatial resolution and reduces of partial volume artifacts compared to two-dimensional sequences. Consequently, subtle or thin lesions can be more readily detected on 3D than with two-dimensional pulse sequences. Chokshi *et al.* (26) reported improved anatomic evaluation of the cervical spine using 3D pulse sequences compared to two-dimensional sequences. Moreover, markedly reduced CSF pulsation artifacts were noted on the 3D sequence; these artifacts can be misinterpreted as a flapped dura mater due to a dural tear and subdural fluid collection or as peri-medullary vascular flow (9,27,28).

HT2W-FS MRM has advantages over CT myelography or MRM with intrathecal gadolinium in that it avoids radiation exposure, adverse effects of contrast media, and lumbar punctures (10). Previous investigations using 3D fat-saturated T2WI or HT2W MRM showed accurate epidural fluid detection in patients with intracranial hypotension compared with CT myelography with intrathecal contrast (95% detection rate) or radionuclide cisternography (81.5%) (9,27). In our study, 3D HT2W-FS MRM was utilized as a mainstay sequence in the diagnosis and evaluation of epidural fluid in patients with clinically suspected intracranial hypotension. Similar to previous studies using HT2W-FS MRM, we found that the diagnostic accuracy of both CR and DLR of 3D HT2W-FS MRM in detecting CSF leaks was 90–97% (9,27).

However, 3D HT2W-FS MRM is not suitable for

Table 5 Correlation of conventional reconstruction and deep learning reconstruction with standard of care in MRM usage

Parameters	Reader 1		Reader 2	
	rho	P value	rho	P value
Per individual spinal level (n=482)				
Conventional reconstruction	0.868	<0.001	0.734	<0.001
Deep learning-based reconstruction	0.919	<0.001	0.805	<0.001
Per patient (n=21)				
Conventional reconstruction	0.908	<0.001	0.430	0.052
Deep learning-based reconstruction	1	<0.001	0.362	0.106

MRM, magnetic resonance myelography.

Table 6 Comparison of qualitative assessment between CR and DLR images

Parameters	Reader 1			Reader 2			Reader 3		
	3D HT2W-FS with CR	3D HT2W-FS with DLR	P value	3D HT2W-FS with CR	3D HT2W-FS with DLR	P value	3D HT2W-FS with CR	3D HT2W-FS with DLR	P value
Image quality	3.72 (0.645)	4.02 (0.632)	<0.001[a]	3.26 (0.640)	4.01 (0.613)	<0.001[a]	3.66 (0.679)	4.28 (0.723)	<0.001[a]
Artifact	4.50 (0.609)	4.73 (0.513)	<0.001[a]	3.58 (0.681)	4.20 (0.658)	<0.001[a]	3.60 (0.543)	3.93 (0.520)	<0.001[a]
Dura mater conspicuity*	2.57 (1.266)	2.80 (1.183)	0.014[b]	2.18 (1.255)	2.77 (1.168)	<0.001[b]	1.87 (1.318)	2.37 (1.327)	<0.001[b]

Data are presented as mean (standard deviation). *, dura mater conspicuity was evaluated in levels that showed evidence of epidural fluid according to reference standard, except for levels with foraminal CSF leakage (total of 119 levels). [a], using Mann-Whitney U-tests; [b], using Wilcoxon signed rank tests. CR, conventional reconstruction; DLR, deep learning-based reconstruction; 3D, three-dimensional; HT2W-FS, heavily T2-weighted fat-saturated image.

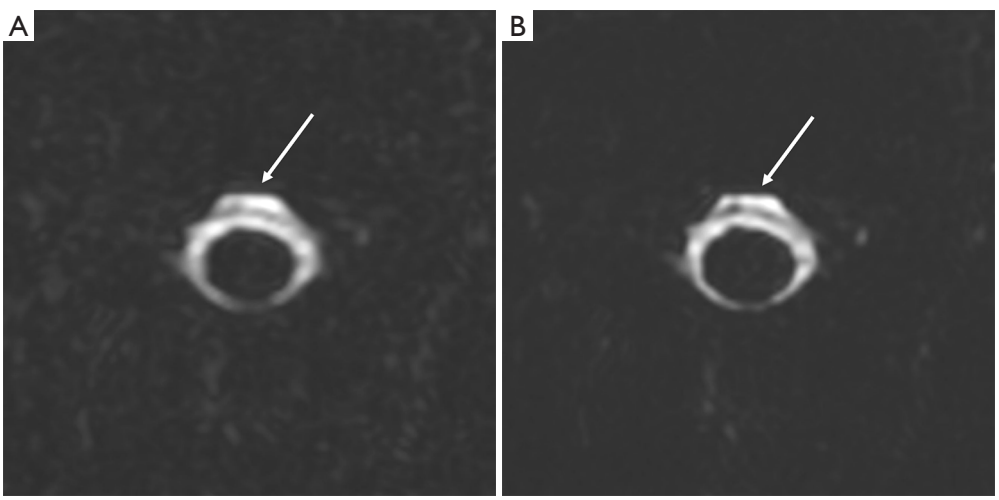

Figure 3 A 25-year-old female with headaches for 2 days. Axial images of 3-dimensional heavily T2-weighted fat-saturated magnetic resonance myelography with CR (A) and DLR (B) at the level of C2–C3 show a ventral epidural fluid (white arrows). All three readers reported the DLR image (B) as having a greater image quality and equal or less artifacts than CR image (A) (image quality scores for CR *vs.* DLR, reader 1: 3 *vs.* 4, reader 2: 3 *vs.* 4, reader 3: 4 *vs.* 5; artifact scores for CR *vs.* DLR, reader 1: 4 *vs.* 4, reader 2: 3 *vs.* 5, reader 3: 4 *vs.* 5). CR, conventional reconstruction; DLR, deep learning-based reconstruction.

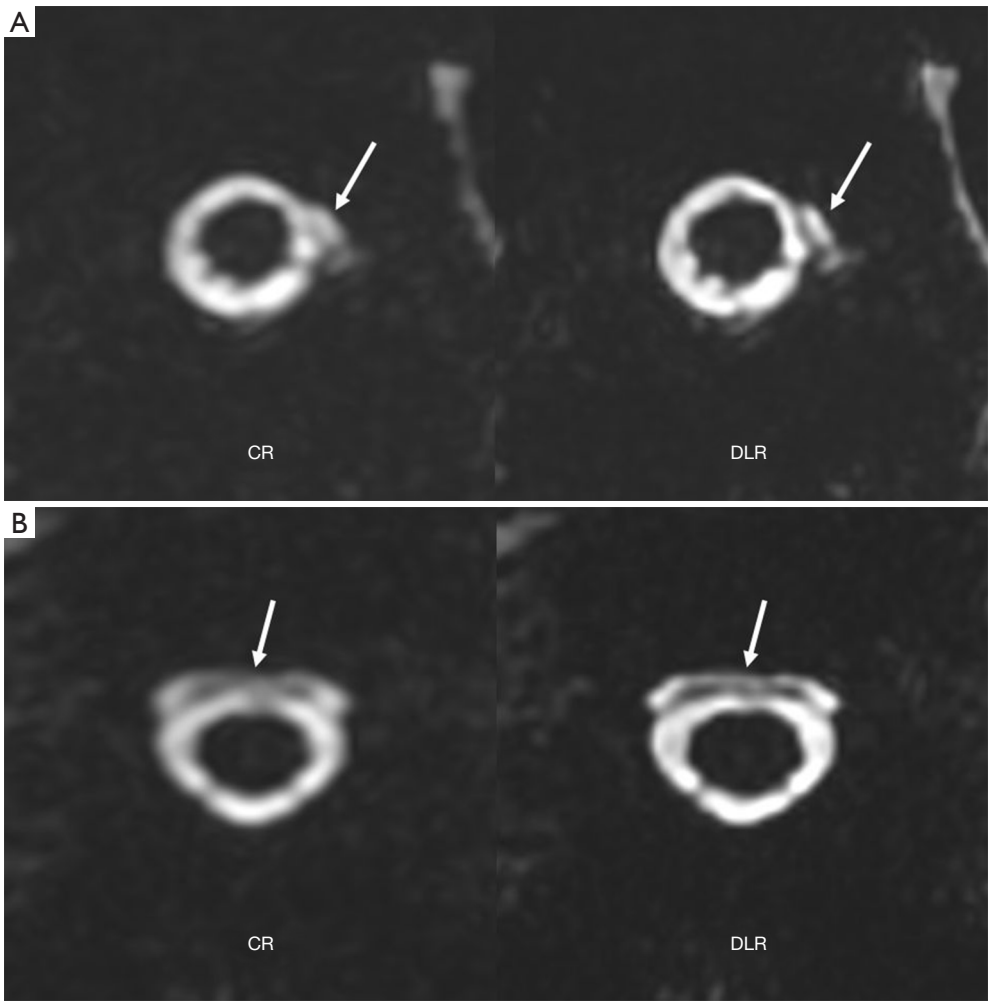

Figure 4 Comparison of dura mater conspicuity of MR myelography. (A) A 30-year-old female with headaches for 4 days. Axial images of 3-dimensional heavily T2-weighted fat-saturated MR myelography with CR and DLR at the level of T8–9 shows an epidural fluid (white arrows) in the left aspect of the dural sac. (B) A 36-year-old female with headaches for 2 months. Axial images of 3-dimensional heavily T2-weighted fat-saturated MR myelography with CR and DLR at the level of T1–2 shows an epidural fluid (white arrows) in the ventral aspect of the dural sac. Although the dura mater separating the epidural fluid and subdural space is visible in both CR and DLR images, all readers reported an improvement in dura mater conspicuity with the DLR images. (Conspicuity scores for CR *vs.* DLR, in patient (A), reader 1, 1 *vs.* 2, reader 2, 1 *vs.* 4, reader 3, 3 *vs.* 4; in patient (B), reader 1, 3 *vs.* 4, reader 2, 3 *vs.* 4, reader 3, 3 *vs.* 4). CR, conventional reconstruction; DLR, deep learning-based reconstruction; MR, magnetic resonance.

detecting the exact location of dural tears because of the static nature of the modality (5). Conventional myelography with or without postmyelography CT, or dynamic CT myelography under CT fluoroscopic guidance, is recommended for evaluating the exact location of dural tears in patients with failed, nontargeted epidural blood patch or with planned surgical exploration. Although we did not evaluate the possible site of the dural tear, we assessed the conspicuity of the dura mater where the epidural fluid was noted. The conspicuity of the dura mater was significantly higher in the DLR group than in the CR group for all readers (P=0.014, P<0.001, and P<0.001, respectively). Further research should focus on the application of DLR to high-resolution MRM for dural tear localization.

This study had several limitations. First, only a small sample size was retrospectively evaluated. However, it's important to note that intracranial hypotension is an underdiagnosed disease, with an annual incidence of 5 in

100,000 per year (29). We believe it is valuable to report our initial application of DLR for 3D MRM to assess both the presence of epidural fluid and image quality. Further prospective studies with larger sample sizes and multi-readers are inevitably required to generalize the results of this study. Second, two-dimensional HT2W-FS MRM and 3D HT2W-FS MRM which represent the standard of care in MRM usage, were employed to assess the presence or absence of epidural fluid. Furthermore, MRM employing heavily T2-weighted FSE sequences remains valuable in clinical practice for detecting CSF leaks and confirming the presence of meningeal diverticula. While MRM is sensitive, the detection of minimal CSF leaks, particularly lateral leaks, can present challenges. Nevertheless, it is imperative to acknowledge that, in comparison to dynamic CT myelography, which is considered reference standard for localizing the CSF leak site while dynamic CT myelography excels in this regard due to its superior temporal resolution (5,9,11). However, the purpose of this study was to detect the presence of extradural epidural fluid rather than the exact site of leakage. By applying DLR, the conspicuity of the dura mater was improved compared to that of CR. In future, high-resolution 3D MRM and DLR might be helpful in identifying the exact location of dural tears and might also be helpful prior to surgical exploration. Third, CSF-venous fistula, another cause of intracranial hypotension, is difficult to detect with MRM without intrathecal contrast administration because of its static nature (5). Patients who are negative on 3D HT2W-FS MRM but highly suspected to have intracranial hypotension should undergo another study, such as lateral decubitus CT myelography or digital subtraction myelography, to determine the exact site of leakage or CSF-venous fistulae. Fourth, the measurement of SNR in the current study was conducted using a conventional approach. Optimized methods of SNR measurement are available after applying specific reconstruction filters, multichannel reconstruction, or parallel imaging (30). However, it is challenging to perform multiple acquisitions on actual patient image due to factors such as patient motion or physiological signal variations, which are necessary to evaluate the true SNR accurately. Several recent researches that also employed DLR for rectal cancer and lumbar spine imaging used conventional SNR approach (19,20). Our investigation revealed no significant difference between subjective visual image quality analysis and the conventional SNR evaluation method. Nevertheless, there remains a need to develop an optimized approach for evaluating SNR in the context of deep learning-based MRI.

Conclusions

In conclusion, HT2W-FS MRM with DLR showed improved image quality, dura mater conspicuity, and diagnostic performance in the detection of epidural fluid compared to CR. Consequently, DLR may improve diagnostic confidence for the detection of epidural fluid in patients with intracranial hypotension.

Acknowledgments

Funding: This study was supported by the 2022 Inje University Haeundae Paik Hospital research grant (No. 2022-10-010). The funder had no role in the design of the study; collection, analysis, and interpretation of the data; or writing the manuscript.

Footnote

Reporting Checklist: The authors have completed the STARD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-24-455/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-24-455/coif). J.L. is a current employee of GE HealthCare Korea. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by institutional review board of Inje University Haeundae Paik Hospital (No. HPIRB 2022-10-010) and individual consent for this retrospective analysis was waived.

References

1. Schievink WI. Spontaneous Intracranial Hypotension. N Engl J Med 2021;385:2173-8.
2. Wang YF, Fuh JL, Lirng JF, Chen SP, Hseu SS, Wu JC, Wang SJ. Cerebrospinal fluid leakage and headache after lumbar puncture: a prospective non-invasive imaging study. Brain 2015;138:1492-8.
3. Oh JW, Kim SH, Whang K. Traumatic Cerebrospinal Fluid Leak: Diagnosis and Management. Korean J Neurotrauma 2017;13:63-7.
4. Amrhein TJ, Kranz PG. Spontaneous Intracranial Hypotension: Imaging in Diagnosis and Treatment. Radiol Clin North Am 2019;57:439-51.
5. Kranz PG, Gray L, Malinzak MD, Amrhein TJ. Spontaneous Intracranial Hypotension: Pathogenesis, Diagnosis, and Treatment. Neuroimaging Clin N Am 2019;29:581-94.
6. Medina JH, Abrams K, Falcone S, Bhatia RG. Spinal imaging findings in spontaneous intracranial hypotension. AJR Am J Roentgenol 2010;195:459-64.
7. Tsai PH, Fuh JL, Lirng JF, Wang SJ. Heavily T2-weighted MR myelography in patients with spontaneous intracranial hypotension: a case-control study. Cephalalgia 2007;27:929-34.
8. Wang YF, Lirng JF, Fuh JL, Hseu SS, Wang SJ. Heavily T2-weighted MR myelography vs CT myelography in spontaneous intracranial hypotension. Neurology 2009;73:1892-8.
9. Dobrocky T, Winklehner A, Breiding PS, Grunder L, Peschi G, Häni L, Mosimann PJ, Branca M, Kaesmacher J, Mordasini P, Raabe A, Ulrich CT, Beck J, Gralla J, Piechowiak EI. Spine MRI in Spontaneous Intracranial Hypotension for CSF Leak Detection: Nonsuperiority of Intrathecal Gadolinium to Heavily T2-Weighted Fat-Saturated Sequences. AJNR Am J Neuroradiol 2020;41:1309-15.
10. Lee SJ, Kim D, Suh CH, Heo H, Shim WH, Kim SJ. Diagnostic yield of MR myelography in patients with newly diagnosed spontaneous intracranial hypotension: a systematic review and meta-analysis. Eur Radiol 2022;32:7843-53.
11. Dobrocky T, Mosimann PJ, Zibold F, Mordasini P, Raabe A, Ulrich CT, Gralla J, Beck J, Piechowiak EI. Cryptogenic Cerebrospinal Fluid Leaks in Spontaneous Intracranial Hypotension: Role of Dynamic CT Myelography. Radiology 2018;289:766-72.
12. Lebel RM. Performance characterization of a novel deep learning-based MR image reconstruction pipeline. arXiv: 2008.06559, 2020.
13. Lee DH, Park JE, Nam YK, Lee J, Kim S, Kim YH, Kim HS. Deep learning-based thin-section MRI reconstruction improves tumour detection and delineation in pre- and post-treatment pituitary adenoma. Sci Rep 2021;11:21302.
14. Hahn S, Yi J, Lee HJ, Lee Y, Lim YJ, Bang JY, Kim H, Lee J. Image Quality and Diagnostic Performance of Accelerated Shoulder MRI With Deep Learning-Based Reconstruction. AJR Am J Roentgenol 2022;218:506-16.
15. Recht MP, Zbontar J, Sodickson DK, Knoll F, Yakubova N, Sriram A, et al. Using Deep Learning to Accelerate Knee MRI at 3 T: Results of an Interchangeability Study. AJR Am J Roentgenol 2020;215:1421-9.
16. Wang X, Ma J, Bhosale P, Ibarra Rovira JJ, Qayyum A, Sun J, Bayram E, Szklaruk J. Novel deep learning-based noise reduction technique for prostate magnetic resonance imaging. Abdom Radiol (NY) 2021;46:3378-86.
17. Sun S, Tan ET, Mintz DN, Sahr M, Endo Y, Nguyen J, Lebel RM, Carrino JA, Sneag DB. Evaluation of deep learning reconstructed high-resolution 3D lumbar spine MRI. Eur Radiol 2022;32:6167-77.
18. Son JH, Lee Y, Lee HJ, Lee J, Kim H, Lebel MR. LAVA HyperSense and deep-learning reconstruction for near-isotropic (3D) enhanced magnetic resonance enterography in patients with Crohn's disease: utility in noise reduction and image quality improvement. Diagn Interv Radiol 2023;29:437-49.
19. Yoo H, Yoo RE, Choi SH, Hwang I, Lee JY, Seo JY, Koh SY, Choi KS, Kang KM, Yun TJ. Deep learning-based reconstruction for acceleration of lumbar spine MRI: a prospective comparison with standard MRI. Eur Radiol 2023;33:8656-68.
20. Kim B, Lee CM, Jang JK, Kim J, Lim SB, Kim AY. Deep learning-based imaging reconstruction for MRI after neoadjuvant chemoradiotherapy for rectal cancer: effects on image quality and assessment of treatment response. Abdom Radiol (NY) 2023;48:201-10.
21. Choi J, Kim B, Choi Y, Shin NY, Jang J, Choi HS, Jung SL, Ahn KJ. Image Quality of Low-Dose Cerebral Angiography and Effectiveness of Clinical Implementation on Diagnostic and Neurointerventional Procedures for Intracranial Aneurysms. AJNR Am J Neuroradiol 2019;40:827-33.
22. McNEMAR Q. Note on the sampling error of the

difference between correlated proportions or percentages. Psychometrika 1947;12:153-7.
23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837-45.
24. Zou KH, Tuncali K, Silverman SG. Correlation and simple linear regression. Radiology 2003;227:617-22.
25. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
26. Chokshi FH, Sadigh G, Carpenter W, Allen JW. Diagnostic Quality of 3D T2-SPACE Compared with T2-FSE in the Evaluation of Cervical Spine MRI Anatomy. AJNR Am J Neuroradiol 2017;38:846-50.
27. Tomoda Y, Korogi Y, Aoki T, Morioka T, Takahashi H, Ohno M, Takeshita I. Detection of cerebrospinal fluid leakage: initial experience with three-dimensional fast spin-echo magnetic resonance myelography. Acta Radiol 2008;49:197-203.
28. Morris JM, Kaufmann TJ, Campeau NG, Cloft HJ, Lanzino G. Volumetric myelographic magnetic resonance imaging to localize difficult-to-find spinal dural arteriovenous fistulas. J Neurosurg Spine 2011;14:398-404.
29. Urbach H, Fung C, Dovi-Akue P, Lützen N, Beck J. Spontaneous Intracranial Hypotension. Dtsch Arztebl Int 2020;117:480-7.
30. Dietrich O, Raya JG, Reeder SB, Reiser MF, Schoenberg SO. Measurement of signal-to-noise ratios in MR images: influence of multichannel coils, parallel imaging, and reconstruction filters. J Magn Reson Imaging 2007;26:375-85.

Cite this article as: Kim M, Yi J, Lee HJ, Hahn S, Lee Y, Lee J. Deep learning-based reconstruction for 3-dimensional heavily T2-weighted fat-saturated magnetic resonance (MR) myelography in epidural fluid detection: image quality and diagnostic performance. Quant Imaging Med Surg 2024;14(9):6531-6542. doi: 10.21037/qims-24-455

