

Standardizing Quality Assurance of Deep-Learning Diffusion-Weighted Imaging Using Qiba Profile and Phantom

Louis-Bernard St-Cyr¹, Philippe Dionne¹, Louis Archambault¹, Wei Liu², Éric Poulin¹.

¹ CHU de Québec-Université Laval, Québec, Canada | ² Siemens Healthineers AG, Forchheim, Germany

INTRODUCTION

Background:

- DWI and ADC are key quantitative MRI biomarkers in oncology, requiring stable, accurate, reproducible acquisition and reconstruction [1]

Gap:

- DL MR reconstruction (FDA/Health-Canada approved, clinically deployed) improves image quality [5,7], but its effect on quantitative ADC is poorly characterized and not routinely tracked by QA programs (SNR, spatial resolution)

Objective:

- Evaluate commercial DL algorithm [4] on quantitative DWI using a Caliber MRI Model 128 phantom and the QIBA DWI Profile [2,3], for clinical SSEPI and prototype rsEPI sequences
- Report ADC, SNR, and MTF to characterize the noise-reduction vs. spatial-resolution trade-off relevant to QA

METHODS AND MATERIALS

- Phantom:** Caliber MRI Model 128 diffusion phantom (13 vials, six PVP-water concentrations 0–50%) [9]; temperature-matched NIST ADC as ground truth [8]; vial #8 (50%) excluded (suspected defect)
- Acquisition:** 1.5 T MAGNETOM Sola (Siemens); clinical SSEPI and a work in progress rsEPI; $b = 50, 500, 1200$ s/mm²; QIBA DWI Profile conformance protocol [3]
- DL reconstruction:** Deep Resolve — three conditions: No DL, DRB (Boost = denoising), DRBS (Boost + Sharp = super-resolution) [4]
- Metrics & stats:** ADC bias and SNR (QIBA-recommended [2], via qCal + in-house Python); spatial resolution via edge-based MTF (MTF₅₀, MTF₁₀); SNR gains tested with paired one-sample t-test ($\alpha = 0.05$)

RESULTS

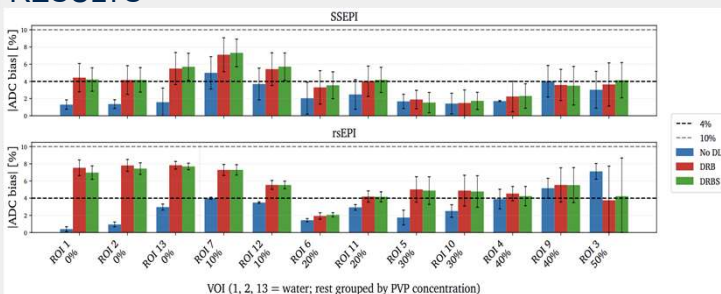


Figure 1. Absolute ADC bias per VOI for the five reconstruction conditions, for SSEPI (top) and rsEPI (bottom) at the QIBA-aligned baseline. The dashed line separates the water vials (1, 2, 13) from the PVP vials. The PVP concentration is specified below the ROI label. Error bars indicate inter-session standard deviation where multiple same-condition sessions are available.

Seq.	Recon vs no DL	b	n	Δ SNR	t	p
SSEPI	DRB	50	3	14.20	3.13	0.089
SSEPI	DRBS	50	3	6.38	0.98	0.430
SSEPI	DRB	500	3	9.60	2.71	0.113
SSEPI	DRBS	500	3	0.96	0.16	0.888
SSEPI	DRB	1200	3	0.00	0.00	0.998
SSEPI	DRBS	1200	3	-4.25	-1.75	0.222
rsEPI	DRB	50	3	17.58	4.09	0.055
rsEPI	DRBS	50	3	13.52	3.62	0.068
rsEPI	DRB	500	3	16.49	13.57	0.005
rsEPI	DRBS	500	3	12.20	8.53	0.013
rsEPI	DRB	1200	3	7.13	7.57	0.017
rsEPI	DRBS	1200	3	4.56	5.26	0.034

Table 1. Paired one-sample t -test of mean Δ SNR = SNR_{DL} - SNR_{No DL} for the central water vial (VOI 1), pooled across same-condition sessions of the QIBA-aligned baselines (SSEPI: body coil, $b_{\max} = 1200$ s/mm², avg = 12; rsEPI: body coil, $b_{\max} = 1200$ s/mm², avg = 6, DS = 0). n is the number of paired sessions contributing to the test. Values with $p < 0.05$ are in bold.

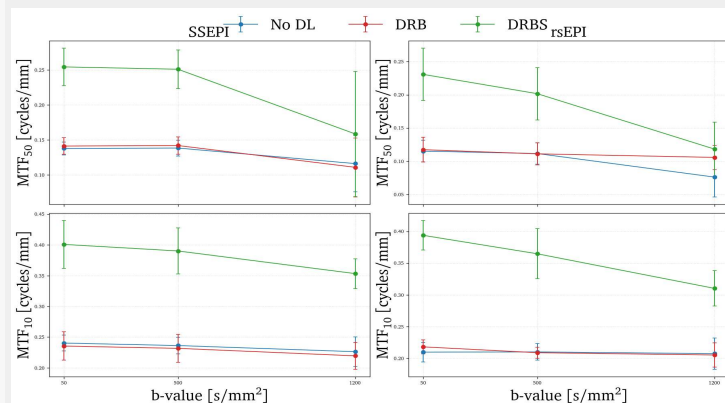


Figure 2. Horizontal MTF₅₀ (top) and MTF₁₀ (bottom) as a function of b -value, for SSEPI (left) and rsEPI (right) at the QIBA-aligned baseline. Error bars are pass-level standard deviations across the four passes of a baseline session. Vertical-MTF results follow the same trends and are omitted for clarity.

DISCUSSION

SNR

- rsEPI central water vial gains up to ~95% at $b = 500$, statistically significant at $b = 500$ and 1200 (all DL conditions, $p < 0.05$)
- SSEPI gains are smaller and non-significant (more averages already acquired)

Spatial resolution (MTF)

- Sharp shifts MTF toward higher frequencies (super-resolution)
- Boost-only does not degrade MTF₅₀ or MTF₁₀ — key for denoising in quantitative DWI

ADC bias (the trade-off)

- DL adds a systematic positive ADC bias (~5% SSEPI, ~7% rsEPI) in the central water vial, decreasing for high-PVP (low-diffusion) vials
- ADC bias is reproducible and low without DL → caused by DL reconstruction [5,6]

Other metrics

- DL increases b -value dependence for both sequences
- Repeatability and precision are unchanged by DL, for both sequences

CONCLUSION

Image quality

- DL improves SNR and combined with super-resolution the MTF: Boost alone does not impact the resolution.

Drawback

- A systematic, reproducible positive ADC bias at low PVP is present.
- The ADC bias is caused by the DL networks, conventional reconstruction (No DL) is unbiased.

QA implication

- QA programs for quantitative DWI should track ADC bias after DL reconstruction: not only SNR and spatial resolution.

REFERENCES

- Padhani AR, et al. Neoplasia. 2009;11:102–125.
- Shukla-Dave A, et al. J Magn Reson Imaging. 2019;49:e101–e121.
- Boss M, et al. Radiological Society of North America. 2024
- Hammernik K, et al. Magn Reson Med. 2018;79:3055–307.
- Hanamoto S, et al. Diagn Interv Radiol. 2023;232149.
- Lemaître T, et al. Magn Reson Imaging. 2024.
- Zompas-Petridis K, et al. Radiol Artif Intell. 2021;3:e200273.
- Keenan KE, et al. Magn Reson Med. 2020;84:1552–1563.
- B. V. Robinson (Lewis Publishers, Chelsea, MI, 1990).