2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Diffusion Weighted MRI Measures on a Commercially Available Low-Field (0.55T), Wide-Bore MRI Simulation System Using Deep Learning Reconstruction

Ellen Park, MS[1]; Joshua P. Kim, PhD[1]; Karen C. Snyder, MS[1]; Ryan Bosca-Harasim, PhD[2]
[1]Henry Ford Health – Department of Radiation Oncology, [2]Henry Ford Health – Department of Radiology

ABSTRACT

Purpose: Commercial availability of low-field (0.55 T), wide-bore (80 cm) MRI systems offer unique opportunities for treatment planning and response assessment in radiation oncology, especially when coupled with deep learning (DL) reconstruction algorithms. Understanding the bias and variance of quantitative diffusion measurements on these systems is important for implementing advanced planning protocols. In this study, diffusion weighted imaging (DWI) was performed with and without DL reconstructions to assess bias and variance of the apparent diffusion coefficient (ADC) on a low-field system at two time points.

Methods and Materials: ADC measurements were calculated from DW images of a NIST/RSNA diffusion phantom (Caliber MRI, Boulder, CO) consisting of 13 compartments with ADC values ranging from approx. 0.238 to 1.111x10^{-3} mm^2/s at 0°C. A commercially available 0.55T (Siemens Free.Max, Germany) MRI system was used to acquire multiple b-value (N=4) DWI images at two time points (baseline and 2 weeks) with traditional and DL-based (DeepResolve; strength=8) reconstructions. Correlation and Bland-Altman repeatability (same reconstruction method) and agreement (different reconstruction method) measures were computed.

Results: Correlations of the traditional and DL-based ADC measures were significant for the two reconstruction algorithms (R^2>0.9997; slopes ranging from 1.0073 to 1.0185). Mean ADC differences (agreement) and limits of agreement for the analysis of reconstruction method agreement ranged from -11.38 to -5.78x10^{-3} mm^2/s and -27.84 to 7.53x10^{-3} mm^2/s, respectively. BA analysis results for repeatability ranged from -38.66 to -29.16x10^{-3} mm^2/s with limits of agreement ranging from -111.73 to 34.41 x10^{-3} mm^2/s. No trends were observed in the BA analyses for either repeatability or reconstruction method agreement.

Conclusion: The clinical low-field, wide-bore MRI scanner evaluated in this study demonstrates minimal bias/variance of ADC measurements using traditional and DL-based reconstruction methods. Future work will focus on developing/optimizing quantitative DWI protocols on this system for in vivo assessment and prediction of treatment response.

CONTACT

Ryan Bosca-Harasim
Henry Ford Health
2799 W. Grand Blvd; Detroit, MI
rbosca1@rad.hfh.edu

INTRODUCTION

Quantitative imaging biomarkers (QIB) derived from diffusion weighted imaging (DWI) are important in diagnosing/staging cancers and assessing response to cancer treatments.[1] Commercial availability of a low-field (0.55 T), wide-bore (80 cm) MRI system offers unique opportunities for treatment planning and response assessment in radiation oncology. Deep learning (DL) based reconstruction algorithms can be used to improve signal characteristics, overcoming inherently lower signal at 0.55T. The purpose of this work was to evaluate bias and variance of ADC estimates on this MR system using a commercially available DL-based reconstruction method.

METHODS AND MATERIALS

A NIST/RSNA diffusion phantom (Caliber MRI, Boulder, CO) consisting of 13 compartments with theoretical ADC values ranging from approx. 0.238 to 1.111x10^{-3} mm^2/s at 0°C was used to acquire multiple b-value axial DWI images on a Free.Max (Siemens, Germany) MRI system. Representative ADC maps for both reconstruction methods are shown in Figure 1.

DWI acquisitions were performed using a traditional reconstruction and DL reconstructions (DLR) at baseline and 2 weeks. The Quantitative Imaging Biomarker Alliance (QIBA) recommend protocol was used for the traditional reconstruction acquisition and adapted for the DRL acquisition. Common acquisition parameters were: matrix=192x192; half-scan=6/8; slice th./sp.=4/1mm; accel. factor=2; min TE~=110ms; TR=10,0000ms; b-values=0, 500, 900, 2000 mm^2/s; 3-scan trace; Head/Neck coil. The number of signal averages (NSA) for the traditional reconstruction acquisition were acquired with 1, 8, 8, 8 for each respective b-values; DLR acquisitions used NSA=1 for all b-values and DeepResolve strength=8. Images for both methods were acquired at 2 time points: baseline and at 2 weeks. An additional acquisition was performed at baseline using DeepResolve strength=4.

ROIs were manually defined on ADC images (syngo MR XA60). Bland-Altman (BA) analysis[2] was performed and correlations calculated on the ROI-averaged ADC values for both reconstruction methods at baseline and 2 weeks (agreement) and between the same reconstruction methods at baseline and 2 weeks (repeatability).

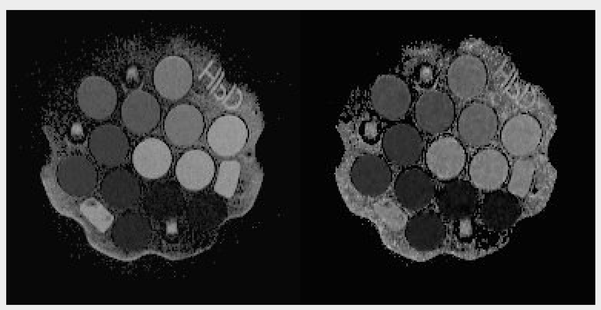


Figure 1. Central slice of the axial ADC maps calculated from the traditional (left) and DL-based (right) reconstruction used to estimate ROI-averaged ADC values from the NIST/RSNA diffusion phantom in week 2.

RESULTS

Correlation between ADC values estimated from traditional and DL-based reconstruction methods were significant (R^2>0.9997), with slopes range from 1.0073 to 1.185. Representative plots are shown in Figure 2.

For reconstruction method agreement, mean differences ranged from -11.38 to -5.78x10^{-3} mm^2/s and limits of agreement ranged from -27.84 to 7.53x10^{-3} mm^2/s. BA plots are shown in Figure 3. Repeatability of the ADC values ranged from -38.66 to -29.16x10^{-3} mm^2/s with limits of agreement ranging from -111.73 to 34.41 x10^{-3} mm^2/s. BA plots are shown in Figure 3.

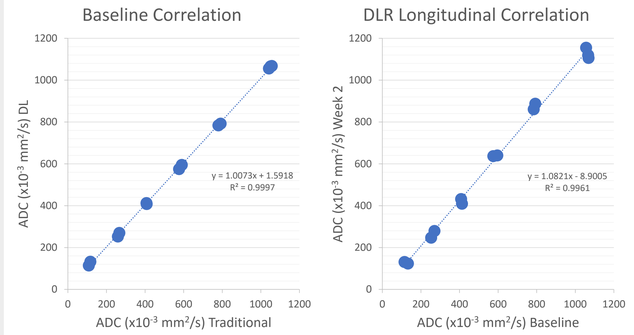


Figure 2. Correlation plots for ADC values estimated from traditional and DL-based reconstructions. *Left*: baseline ADC value correlation between traditional and DL reconstructions (strength 8); *Right*: ADC value correlation with DL reconstruction (strength 8) between baseline and week2.

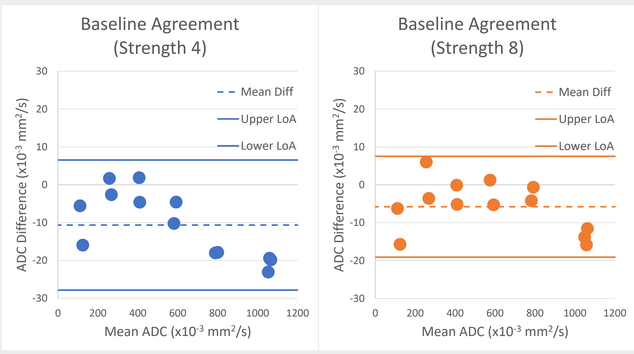
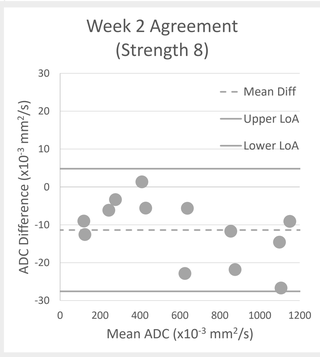


Figure 3. Bland-Altman agreement plots for ADC values estimated from traditional and DL-based reconstructions. The mean of the differences (agreement) and upper/lower limits of agreement are displayed. *Top left*: baseline agreement with DL strength 4; *Top right*: baseline agreement with DL strength 8; *Bottom right*: week 2 agreement with DL strength 8.



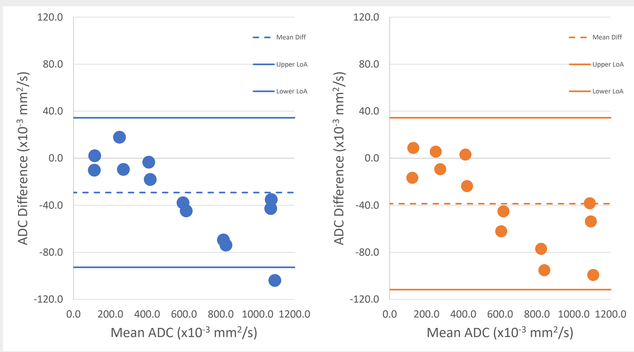


Figure 4. Bland-Altman repeatability plots for ADC values estimated from traditional and DL-based reconstructions. *Left*: repeatability of ADC values estimated from traditional reconstruction; *Right*: repeatability of ADC values estimated from DL-based reconstructions with DL strength 8.

DISCUSSION

ADC values estimated from traditional and DL-based reconstruction methods were in good agreement (i.e., minimal bias and variance) at both time points. Percent differences were ~5% or less, except for values estimated from ROI 7 (lowest ADC value) which was ~12%. Repeatability demonstrated larger variance for both reconstruction methods with minimal bias. Percent differences were less than ~11% for the traditional reconstruction and less than ~13% for the DL-based reconstruction. The consistency in increased variance of repeatability with consistent agreement at baseline and 2 weeks between methods suggests sources of variation other than the reconstruction method.

While ROI-averaged ADC values were in good agreement between traditional and DL-based reconstructions, increased noise and small regional variations are visually apparent in ADC maps created using DLR. These variations are likely to affect more granular analyses (e.g., voxelwise response).

For both methods, bias and variance for agreement and repeatability were below the proposed repeatability coefficients published by QIBA[3] for tumors of the brain, liver, prostate, and breast, suggesting that ADC values derived from the DL-based reconstructions could be used as QIBs for assessing treatment response. Comparable changes in ADC following treatment have also been reported for head/neck and cervical cancers.[4,5]

CONCLUSIONS

The NIST/RSNA diffusion phantom was used to evaluate agreement between traditional and DL-based reconstruction methods and to assess repeatability of ADC measurements. Bias and variance introduced by DL-based image processing and reconstruction methods should be characterized using reference test objects, especially when such methods are used to derive QIBs.

Future work will focus on developing and optimizing quantitative, organ specific DWI protocols on this system – protocols that can be used to assess tumor response to cancer treatment. Bias and variance in ADC estimates derived from these protocols will need to be characterized with phantom studies and in vivo imaging.

REFERENCES

1. Padhani, Anwar R., et al. "Diffusion-weighted magnetic resonance imaging as a cancer biomarker: consensus and recommendations." Neoplasia 11.2 (2009): 102-125.
2. Bland, J. Martin, and Altman, Douglas G. "Statistical methods for assessing agreement between two methods of clinical measurement." The lancet 327.8476 (1986): 307-310.
3. Boss, Michael A., et al. "The QIBA profile for diffusion-weighted MRI: apparent diffusion coefficient as a quantitative imaging biomarker." Radiology 313.1 (2024): e233055.
4. Wong, Kee H., et al. "Changes in multimodality functional imaging parameters early during chemoradiation predict treatment response in patients with locally advanced head and neck cancer." European Journal of Nuclear Medicine and Molecular Imaging 45.5 (2018): 759-767.
5. Vasić, Jelena, et al. "The apparent diffusion coefficient as a biomarker in the diagnosis of cervical cancer and the assessment of therapeutic response to chemoradiation therapy." Frontiers in Oncology 15 (2025): 1610090.