Precision and Bias in ADC Measurements on Pre-Clinical MRIs Using a Standardized DWI Phantom and Procedure

Thomas L. Chenevert* on behalf of CIRP IADP DWI Phantom Round-Robin Project Participants

* Department of Radiology
University of Michigan
tlchenev@med.umich.edu
Background & Motivation

- Prior work indicates reasonable repeatability ($\approx 3\%$) and reproducibility ($\approx 6\%$) of ADC measurements on pre-clinical MRIs, though inferior relative to clinical MRIs.
- Pre-clinical MRIs exhibited significant absolute bias ($\approx 15\%$).


Objectives

- Investigate apparent discrepancy of ADC measurement on clinical vs pre-clinical systems
- Assess ADC repeatability and reproducibility on pre-clinical MRIs
- Measure absolute bias, spatial uniformity, and SNR$_{DWI}$
- Generate multi-vendor DWI/ADC data in vendor-native format
- Assess sites’ ITK-compatible format of same multi-vendor data
- Compare site- vs central-lab ADC measurements on common datasets
Experiment Design

- Round-robin of ice water-based DWI phantom
- Detailed phantom preparation instructions
- Standardized (simple) 3 bvalue DWI protocol
- Acquire test-retest data for short- & long-term repeatability
CIRP / IADP Participation Summary

- 10 systems
- 7 sites
- 3 vendors
- 6 field strengths (3 - 14T)

<table>
<thead>
<tr>
<th>Sys</th>
<th>Site</th>
<th>Vendor</th>
<th>Magnetic Field (T)</th>
<th>SW version</th>
<th>Day1</th>
<th>Day2</th>
<th>Site ROIs</th>
<th>Central ROIs</th>
<th>Vendor Format</th>
<th>ITK Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>UMICH</td>
<td>Bruker</td>
<td>7</td>
<td>PV7.0.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>MHD</td>
</tr>
<tr>
<td>2</td>
<td>UPENN</td>
<td>Bruker</td>
<td>9.4</td>
<td>PV6.0.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>MHD &amp; DICOM</td>
</tr>
<tr>
<td>3</td>
<td>MDACC</td>
<td>Bruker</td>
<td>7</td>
<td>PV6.0.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>4</td>
<td>WUSTL</td>
<td>Bruker</td>
<td>9.4</td>
<td>PV360 v2.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>5</td>
<td>WUSTL</td>
<td>Agilent</td>
<td>11.74</td>
<td>VnmrJ4. 2 revA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>6</td>
<td>UCSF</td>
<td>Bruker</td>
<td>3</td>
<td>PV6.0.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>7</td>
<td>BAYLOR</td>
<td>Bruker</td>
<td>9.4</td>
<td>PV360 v3.0</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>NIFTI</td>
</tr>
<tr>
<td>8</td>
<td>UWASH</td>
<td>Bruker</td>
<td>4.7</td>
<td>PV6.0.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>9</td>
<td>UWASH</td>
<td>Bruker</td>
<td>14</td>
<td>PV5.1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>DICOM</td>
</tr>
<tr>
<td>10</td>
<td>UMICH</td>
<td>MR Solutions</td>
<td>3</td>
<td>v4.0.2.4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>MHD &amp; DICOM</td>
</tr>
</tbody>
</table>
Data Processing Workflow

DICOM Format DWI

Native Vendor Format DWI (b=0, 1k, 2k) (2dseq; fdf; sur)

Matlab Scripts

Matlab Scripts ‘lscov’ least-sq linear fit of log(S) vs b-value

MHD format:
- Low-b DWI
- High-b DWI
- ADC map

Export to CSV / excel

Independent stats each slice

Graphical summary each system

ROI each slice

4mm φ ROI
Results: Individual System Bias vs at Magnet Isocenter

± error bars are +max and –min of all scans on each system

±5%
Individual System Bias vs at Magnet Isocenter

CIRP – IADP Study

Prior Study
Bias & Cross-System Reproducibility vs Z-location

Bias & Reproducibility All Systems Included (N=10)

Bias & Reproducibility Excluding Sys9 & Sys10 (N=8)
Bias & Inter-Exam (long-term) Repeatability vs Z-location

Bias & Inter-Exam Repeatability Including All Systems (N=10)

Bias & Inter-Exam Repeatability Excluding Sys9 & Sys10 (N=8)
Bias & Intra-Exam (short-term) Repeatability vs Z-location
System SNR Estimation

- Two-scan method: pixelwise average and difference of identical scans

\[ M = \frac{[scan_1 + scan_2]}{2}; \quad D = [scan_1 - scan_2] \]

\[ SNR_{ROI} = \sqrt{2} \frac{ROI \text{ Mean} (M)}{ROI \text{ Stdev} (D)} \]

- Identical receiver gain confirmed on only 3 of 10 systems
- Background noise method: estimate noise from background (slice1)

\[ SNR_{ROI} = \sqrt{2 - \frac{\pi}{2}} \quad \frac{ROI \text{ Mean} (scan_i)}{Stdev (background_i)} \]
Based on estimated SNR & simulations, noise should not contribute to bias on any of these systems (iso-center ± 15mm)
Site vs Central Lab ADC Measurement

Sys2 Day1 Scan1 Site Measurement

Sys6 Day1 Scan1 Site Measurement

Sys2 Day1 Scan1 Central Lab Measurement

Sys6 Day1 Scan1 Central Lab Measurement
Summary & Conclusions

• Main objectives met
• ADC reproducibility, repeatability AND bias of pre-clinical MRIs is comparable to clinical MRIs at isocenter - two outlier systems identified: ave bias < 5% at isocenter; excl outliers ave bias < 2%
• Increased bias and poorer reproducibility / repeatability with distance from isocenter
• SNR estimates indicate noise is not a contributor to bias
• Spatial pattern of bias is consistent with gradient non-linearity
• Sources of site- vs central-lab ADC measurement discrepancies:
  - ADC fit routine
  - improper interpretation of DICOM intensity scaling
Thank You!