A Meta-Analysis on repeatability of Magnetic Resonance Elastography of Liver.

Suraj D. Serai, Sudhakar K. Venkatesh, Claude B. Sirlin, Nancy Obuchowski, Patricia Cole, Edward Ashton, Frank Miller, Richard L. Ehman
Diagnosis of Liver Fibrosis

• Currently, liver biopsy is the “gold” standard
• However, it’s an imperfect “gold” standard

• Painful
• Complications (~0.5%) ¹
• Sampling error ~15-25% (only ~1/5,000th of liver mass obtained - <0.02% of the 1.5 kg liver)
• Sedation/GA
• Typically 24 hr. inpatient observation
• Substantial variability in staging of fibrosis (discordance in up to 33% of cases) ²
• ~20% of specimens understaged ³

² Regev A, et al. AJG 2002;97:2614-2618
Diagnosis of Liver Fibrosis

- Currently, liver biopsy is the “gold” standard
- However, it’s an imperfect “gold” standard

Hence:

Non-invasive techniques for assessment and quantification of liver fibrosis are critical for clinical surveillance and validation.

2 Regev A, et al. AJG 2002;97:2614-2618
MR Elastography
Quantitative Imaging of Tissue Stiffness

- Main application: Assessing liver fibrosis (Diagnosis, Surveillance and Therapeutic monitoring).
- Available on both – 1.5T and 3T platforms.
- Acquisition time: ~ 1 minute GRE based; 15 seconds SE-EPI based
- Installed clinical base: > 900 systems worldwide.

![MRI images showing different grades of liver fibrosis](image)
Measurement Accuracy of MRE

Average ± Standard Deviation of Shear Modulus (kPa)
Calculated from Dynamic Mechanical Analysis and MRE

<table>
<thead>
<tr>
<th>%Agar</th>
<th>DMA</th>
<th>MRE manual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>14.28 ± 1.34</td>
<td>17.42 ± 1.12</td>
</tr>
<tr>
<td>2</td>
<td>25.02 ± 0.21</td>
<td>33.54 ± 1.56</td>
</tr>
<tr>
<td>2.5</td>
<td>49.26 ± 1.38</td>
<td>52.28 ± 2.28</td>
</tr>
<tr>
<td>3</td>
<td>80.81 ± 1.57</td>
<td>86.07 ± 4.02</td>
</tr>
<tr>
<td>3.5</td>
<td>108.30 ± 4.35</td>
<td>108.10 ± 8.89</td>
</tr>
</tbody>
</table>

Ringleb et al. MRM (2005)

Plot of DMA vs MRE shear stiffness. The dotted line is the line of unity. ICC = 0.99 (95% CI = 0.97-0.99)

Arunachalam et al. MRM (2016)
Purpose of this study: MRE Profile Longitudinal Claim Test – Retest Repeatability

For a given measured percentage change in the magnitude of the complex shear modulus, a plausible range for the true change is the measured change, with 95% confidence.

(Assuming no change in hardware and software platform and analysis method.)

Example: 15 year old patient with NASH

Pre-Treatment  (6 months)  Post-Treatment

3.0 kPa

2.3 kPa
Methods: MRE repeatability search

Inclusion criteria:
1. studies that reported measurements of change in liver stiffness measured at two or more timepoints under similar conditions;
2. studies that reported MRE-based stiffness values as absolute value of shear modulus;
3. studies that reported the time between repeat measurements, mean liver stiffness and the coefficient of variation.

Exclusion criteria:
1. duplicate publication (based on the same primary study);
2. non-original research; and
3. studies not published in English.
Methods: MRE repeatability search

From the 12 studies, the following data were extracted:

(1) Author, journal and year of publication;
(2) number of subjects;
(3) number of readers;
(4) within-subject Coefficient of variation (wCV);
(5) notes on method used to calculate the wCV.

Library 1: 309 articles found through database search.
Duplicate articles removed and list combined in single endnote. 450 articles collected.

Library 2: 350 articles found through database search.

Articles were independently screened by two observers for repeatability and reproducibility studies.

12 articles met the inclusion criteria and were included for analysis. All authors confirmed and verified reciprocally.
Results: 12 studies comprising of 274 patients met the inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Study</th>
<th>Year</th>
<th>Study design</th>
<th>Sample size</th>
<th>Age range</th>
<th>Male %</th>
<th>MR scanner</th>
<th>Field strength</th>
<th>Frequency (Hz)</th>
<th>Property measured</th>
<th>Subjects</th>
<th>No. of readers</th>
<th>Time interval</th>
<th>COV reported(%)</th>
<th>fasting vs feeding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wang et al</td>
<td>2011</td>
<td>Prospective</td>
<td>5</td>
<td>NA</td>
<td>NA</td>
<td>Siemens Espree</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>2</td>
<td>2 weeks</td>
<td>9 - 12%</td>
<td>No information</td>
</tr>
<tr>
<td>Venkatsh et al</td>
<td>2014</td>
<td>Prospective</td>
<td>41</td>
<td>23 - 63</td>
<td>44</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>2</td>
<td>4 - 6 weeks</td>
<td>8.4</td>
<td>4 - 6 hrs. fasting</td>
</tr>
<tr>
<td>Shire et Al</td>
<td>2011</td>
<td>Prospective</td>
<td>9</td>
<td>20 - 57</td>
<td>44</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy, 4 patients</td>
<td>3</td>
<td>1 - 2 weeks</td>
<td>6 - 11%</td>
<td>8 hrs. fasting</td>
</tr>
<tr>
<td>Shingawala et Al</td>
<td>2014</td>
<td>Prospective</td>
<td>10</td>
<td>27 - 63</td>
<td>90</td>
<td>GE 750W</td>
<td>3.0 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>1</td>
<td>1 week</td>
<td>NA</td>
<td>No information</td>
</tr>
<tr>
<td>Shin et Al</td>
<td>2014</td>
<td>Retrospective</td>
<td>15</td>
<td>57 (mean)</td>
<td>NA</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>patients</td>
<td>2</td>
<td>2 weeks</td>
<td>NA</td>
<td>No information</td>
</tr>
<tr>
<td>Shi et al</td>
<td>2014</td>
<td>Prospective</td>
<td>22</td>
<td>18 - 56</td>
<td>41</td>
<td>GE HDx</td>
<td>3.0 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>2</td>
<td>1 week - short term; 27 - 30 wks - long term</td>
<td>5.75</td>
<td>8 hrs fasting</td>
</tr>
<tr>
<td>Lee et Al</td>
<td>2014</td>
<td>Retrospective</td>
<td>47</td>
<td>27 - 82</td>
<td>68</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>patients</td>
<td>2</td>
<td>8 - 10 mins.</td>
<td>13</td>
<td>No information</td>
</tr>
<tr>
<td>Jajamovich et al</td>
<td>2014</td>
<td>Prospective</td>
<td>30</td>
<td>55.8 (mean)</td>
<td>77</td>
<td>GE 750</td>
<td>3.0 T</td>
<td>60</td>
<td>Magnitude</td>
<td>11 healthy, 19 patients</td>
<td>2</td>
<td>20 minutes</td>
<td>3.8</td>
<td>6 hrs fasting and then repeated after feeding</td>
</tr>
<tr>
<td>Hines et al</td>
<td>2010</td>
<td>Prospective</td>
<td>30</td>
<td>21 - 68</td>
<td>53</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>20 volunteers, 10 patients</td>
<td>2</td>
<td>2 - 4 weeks</td>
<td>17.4</td>
<td>No information</td>
</tr>
<tr>
<td>Hines et al</td>
<td>2011</td>
<td>Prospective</td>
<td>11</td>
<td>23 - 39</td>
<td>75</td>
<td>GE HDx</td>
<td>1.5 T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>1</td>
<td>5 weeks</td>
<td>8.5</td>
<td>Feeding between scans</td>
</tr>
<tr>
<td>Bohte et al</td>
<td>2013</td>
<td>Prospective</td>
<td>30</td>
<td>19 - 59</td>
<td>50</td>
<td>Philips</td>
<td>3.0 T</td>
<td>50</td>
<td>Magnitude, Propagation velocity</td>
<td>16 volunteers, 14 patients</td>
<td>1</td>
<td>1 - 4 weeks</td>
<td>10.1</td>
<td>No information</td>
</tr>
<tr>
<td>Trout et al</td>
<td>2016</td>
<td>Prospective</td>
<td>24</td>
<td>22 - 55</td>
<td>21</td>
<td>GE &amp; Philips</td>
<td>1.5T &amp; 3.0T</td>
<td>60</td>
<td>Magnitude</td>
<td>Healthy subjects</td>
<td>1</td>
<td>same day</td>
<td>10.7</td>
<td>6 - 8 hrs. fasting</td>
</tr>
</tbody>
</table>

Table – Baseline characteristics of included studies.
QUADAS-2 tool

Risk of Bias
1. Patient Selection
2. Index Test
3. Reference Standard
4. Flow and Timing

Applicability Concerns
5. Patient Selection
6. Index Test
7. Reference Standard

QUADAS – Quality Assessment tool for Diagnostic Accuracy Studies.
Repeatability Coefficient was calculated:

\[
\%RC = 1.96 \times \sqrt{2} \times \%wCV^2
\]

where \(\%wCV\) is the within-subject coefficient of variation.

The 95% confidence interval (CI) for the RC for each study was calculated as:

\[
2.77 \times \sqrt{M \times \%wCV^2 / \chi^2_{M, \alpha}}
\]

where \(\chi^2_{M, \alpha}\) is the \(\alpha\)th percentile of the chi square distribution with \(M\) degrees of freedom. For the lower bound, \(\alpha\) is 0.975, and for the upper bound, \(\alpha\) is 0.025.
Results

Forest Plot from 12 studies: Summary $\text{RC} = 22\%$ [16.1 – 28.2]
## Results

<table>
<thead>
<tr>
<th>Condition</th>
<th># studies</th>
<th>Summary RC</th>
<th>95% bootstrap CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trained operator to draw ROI</td>
<td>10</td>
<td>18.4</td>
<td>[14.2, 22.2]</td>
</tr>
<tr>
<td>&lt;1 week between scans</td>
<td>5</td>
<td>17.5</td>
<td>[11.6, 23.4]</td>
</tr>
<tr>
<td>&gt; 1 week between scans</td>
<td>5</td>
<td>19.3</td>
<td>[15.6, 21.8]</td>
</tr>
<tr>
<td>Untrained operator to draw ROI</td>
<td>2</td>
<td>34.5</td>
<td>--</td>
</tr>
<tr>
<td>1.5 field strength</td>
<td>8</td>
<td>25.2</td>
<td>[17.4, 31.9]</td>
</tr>
<tr>
<td>&lt;1 week between scans</td>
<td>2</td>
<td>21.7</td>
<td>--</td>
</tr>
<tr>
<td>&gt; 1 week between scans</td>
<td>6</td>
<td>26.0</td>
<td>[16.7, 34.2]</td>
</tr>
<tr>
<td>3.0 field strength</td>
<td>4</td>
<td>15.6</td>
<td>[10.5, 20.8]</td>
</tr>
<tr>
<td>&lt;1 week between scans</td>
<td>3</td>
<td>12.7</td>
<td>[10.0, 15.9]</td>
</tr>
<tr>
<td>&gt; 1 week between scans</td>
<td>1</td>
<td>22.2</td>
<td>--</td>
</tr>
<tr>
<td>All 12 studies</td>
<td>12</td>
<td>22.0</td>
<td>[16.1, 28.2]</td>
</tr>
</tbody>
</table>

![Magnitude Image](image1.png)  ![Wave Image](image2.png)  ![Elastogram](image3.png)
Results

Funnel Plot was generated to address publication bias.

Funnel plot of the RC estimates from each study (on the y-axis) versus the effective sample size (on the x-axis). The funnel plot shows that studies with the large sample size fall near the summary value of 22%, and smaller studies fall fairly symmetrically on either side towards the bottom of the plot.
Conclusion

• MRE is a repeatable, non-invasive method for detecting and staging liver fibrosis.

• Our estimated meta-analysis summary: Repeatability coefficient: 22% with a 95% CI of [16.1 – 28.2]

• Assuming no change in MRE hardware and software, a change of 22% or larger can be considered a true change.
Acknowledgements:
RSNA QIBA MRE Biomarker Committee

- Richard L. Ehman, MD, PhD (Co-Chair)
- Patricia Cole, MD, PhD (Co-Chair)
- Sudhakar K. Venkatesh, MD
- Claude B. Sirlin, MD
- Suraj D. Serai, PhD
- Edward Ashton, PhD
- Frank Miller, MD
- Nancy Obuchowski, PhD
- Mark Palmeri, MD, PhD
- Edward Jackson, PhD