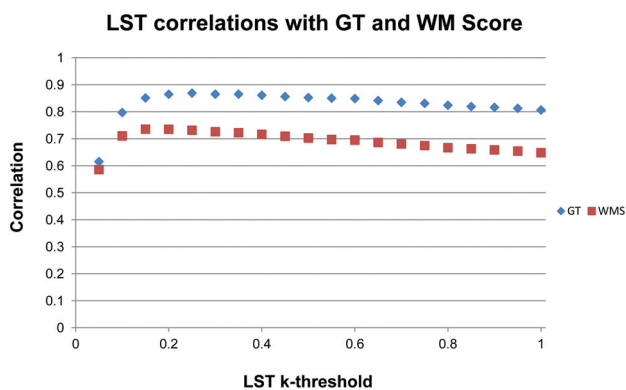**WM Lesion Score Histogram**

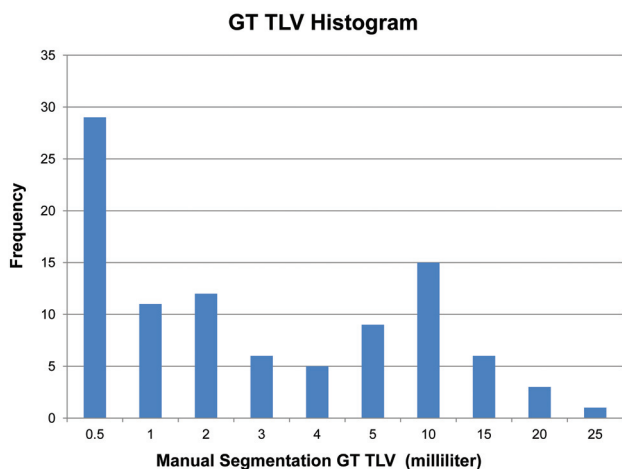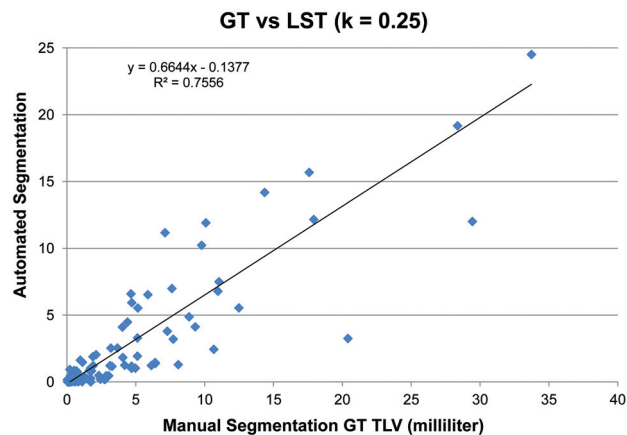**ON-LINE FIG 1.** Histogram plot of average subjective WM lesion scores. Scores were skewed toward the lower end of the scale.
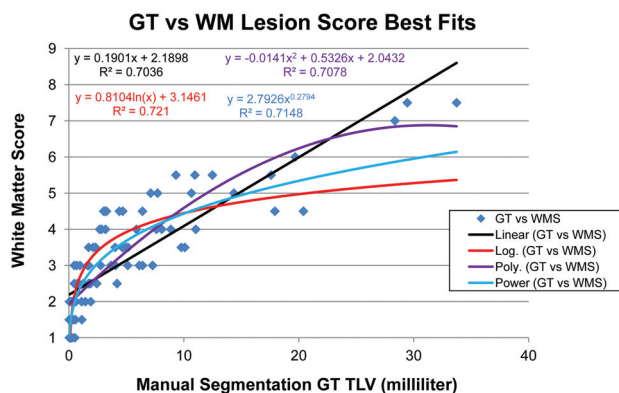

LST correlations with GT and WM Score

ON-LINE FIG 2. Lesion Segmentation Toolbox correlations with ground-truth and WM scores. LST k-threshold is on the x-axis, correlation on the y-axis. Maximum correlation at $k = 0.25$ for GT (blue) and $k = 0.15$ for average WM scores (red).

GT TLV Histogram

ON-LINE FIG 3. Ground truth total lesion volume histogram. Total lesion volume is skewed toward <1 mL.

GT vs LST (k = 0.25)

$y = 0.6644x - 0.1377$
$R^2 = 0.7556$

ON-LINE FIG 4. Ground-truth total lesion volume versus automated Lesion Segmentation Toolbox TLV at $k = 0.25$. A strong linear relationship around unity is demonstrated, with $R^2 = 0.76$.

GT vs WM Lesion Score Best Fits

$y = 0.1901x + 2.1898$
$R^2 = 0.7036$

$y = -0.0141x^2 + 0.5326x + 2.0432$
$R^2 = 0.7078$

$y = 0.8104\ln(x) + 3.1461$
$R^2 = 0.721$

$y = 2.7926x^{0.2794}$
$R^2 = 0.7148$

ON-LINE FIG 5. Ground-truth total lesion volume (milliliters) versus WM scores (1–10) demonstrating a nonlinear relationship. Better fits for log (red), polynomial (purple), and power law (blue) than linear (black) relationship. Best fit for log(GT TLV), $R^2 = 0.72$.