Supplementary Online Content


**eAppendix 1. Methods**

**eAppendix 2. Results**

**eReferences**

**eTable.** Effect of Imaging Parameters Tested as Predictors of Disease Progression (Annual Slopes) From the Confirmatory Study

This supplementary material has been provided by the authors to give readers additional information about their work.
eAppendix 1. Methods

Spectroscopy predictors in the preliminary dataset

The acquisitions had a phase-encoding matrix of 12 x 12 x 8 with an FOV of 120 x 120 x 80 mm resulting in a nominal voxel size of 1 cm^3. A point-resolved spectroscopy (PRESS) volume selection was used. Chemical shift misregistration was reduced by prescribing a PRESS box larger than the region of interest (ROI, overpress) by a factor of 1.2 using 4 saturation bands to suppress signals arising from beyond the ROI. Metabolite contributions within each voxel were estimated by adjusting short echo time signals to a model function created from a prior knowledge basis set of metabolite signals according to the semi-parametric approach developed in HR-QUEST accounting for macromolecular contributions, and for NAA and mI T1 relaxation values estimated at 3T (T1 NAA_{WM}=1.37s and T1 NAA_{GM}=1.36s in healthy controls; T1 NAA_{WM}=1.31s and T1 NAA_{GM}=1.44s in MS subjects; T1 mI_{WM}=1.09s and T1 mI_{GM}=1.28s in healthy controls; T1 mI_{WM}=0.98s and T1 mI_{GM}=1.16s in MS subjects) using a similar method described previously. Non-brain regions were removed from the anatomical images using a semi-automated brain extraction tool (FSL, www.fmrib.ox.ac.uk/fsl). T1-weighted three-dimensional inversion recovery spoiled gradient-echo (IR-SPGR) images were segmented into GM and WM compartments using a hidden Markov random field model with expectation maximization. The GM and WM maps were re-gridded to spectroscopy resolution and convolved with the point spread function for spectroscopic imaging to yield the percent GM and WM content within each spectroscopic voxel. Manual lesion segmentation was performed on from 3D T1-weighted gradient-echo images. For patients, the voxels containing the manually segmented lesions were removed from the analysis. Moreover, the spectroscopic voxels were included in the linear fit only if their concentration estimates had estimated Cramer-Rao bounds within threshold values (10% for [NAA], [Cho] and [Cr], 30% for the [mI]).

Statistical Analysis

Mixed-effects models were initially fit with single predictors (each outcome versus each predictor). Subsequently, models with multiple predictors were fitted to determine independent additive value of predictors, e.g., combinations of multiple metabolites in NAWM, NAGM and MWF_{NAWM}. Postulated models were based on clinical relevance and once the model was fitted, variables with no clinical or/statistical significance were dropped. Clinical relevance was primarily based on biological plausibility and was used to define subsets of variables to consider as predictors (e.g. all GM metabolite predictors).

Results from the mixed-effects models are given in terms of estimated annualized changes along with 95% confidence interval (CI).

Many of the mixed-effects models were checked for assumptions of normality (via qq-plots) and they appeared reasonable with no extreme outliers (including metabolite ratios). There were too many models examined to check the assumptions for every one of them. However, because many of the outcomes considered were inter-related, the checks performed on a subset of models provide some insurance against distributional violations in the remaining models.

Although we examine many differences and issues, we report nominal P-values, without adjustment for multiple testing. Adjustment would require that each result detract from the others, but there are clear biological relationships among many of the variables that we examine leading to correlated tests (in contrast to the assumed independence of tests in conventional multiple testing procedures), and these permit coherent sets of findings to reinforce each other rather than detract from one another. Thus, multiple comparison adjustment would do exactly the wrong purpose in this situation. We therefore rely on scientific judgment and study design rather than formal adjustment methods to indicate where caution is warranted despite findings with P < 0.05. Finally, because we are validating all results in an independent dataset the potential for false positive results is greatly diminished.
**eAppendix 2. Results**

**Individual metabolites as predictors of brain volume and clinical outcomes from the confirmatory dataset**

Using single predictor analyses, we did not observe a statistically significant influence of mI or NAA on brain volume loss (eTable 1). However, there were additive longitudinal effects of mI and NAA in NAWM on PBVC over the course of the study [-0.1 annual slope; 95% CI: -0.19 to -0.008; \( P = .03 \) and 0.05%; 95% CI: -0.001 to 0.096; \( P = .06 \) respectively]. mI levels in NAWM predicted longitudinal changes in MSFC [-0.03 point annually; 95% CI: -0.042 to -0.009; \( P = .02 \)]. When combined additively in a multiple predictor analysis, NAA [-0.31; 95% CI: -0.59 to -0.04; \( P = .03 \)] and mI [0.62; 95% CI: 0.11 to 1.15; \( P = .02 \)] from NAWM predicted 12-month sustained EDSS progression.

**Lesion volume correlation with mI/NAANAWM ratio**

We have explored the following correlations using Spearman’s correlation: mI/NAANAWM ratio was statistically and significantly correlated with baseline lesion volume in both (1) the preliminary (estimated correlation 0.28, 95% CI: 0.02 to 0.51, \( P = .04 \)) and (2) confirmatory (estimated correlation 0.22, 95% CI: 0.10 to 0.35, \( P = .0009 \)) datasets.

**Lesion volume as predictor of disease evolution**

Baseline lesion volume was able to predict the evolution of PBVC in the confirmatory dataset [-0.006 annual slope; 95% CI: -0.011 to -0.001; \( P = .03 \)] but not in the preliminary data set [-0.008 annual slope; 95% CI: -0.019 to 0.003; \( P = .14 \)].

Baseline lesion volume was not able to predict the evolution of EDSS and zMSFC in either dataset. EDSS evolution: -0.00034; 95% CI: -0.000013 to 0.000006; \( P = .50 \) and zMSFC evolution: -0.000001; 95% CI: -0.000004 to 0.000002; \( P = .54 \), in the preliminary study. EDSS evolution: 0.0002; 95% CI: -0.0029 to 0.0033; \( P = .89 \) and zMSFC evolution: 0.0002; 95% CI: -0.0008 to 0.0012, \( P = .67 \), in the confirmatory study.

**eReferences**

**eTable.** Effect of Imaging Parameters Tested as Predictors of Disease Progression (Annual Slopes) From the Confirmatory Study

<table>
<thead>
<tr>
<th>Predictors</th>
<th>PBVC (%/y)</th>
<th>EDSS (pts/y)</th>
<th>zMSFC (pts/y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAA$_{NAWM}$</td>
<td>0.02 (-0.02, -0.06), $P$=.35</td>
<td>-0.008 (-0.04, 0.02), $P$=.56</td>
<td>-0.004 (-0.01, 0.004), $P$=.32</td>
</tr>
<tr>
<td>NAA$_{NAGM}$</td>
<td>0.02 (-0.02, 0.05), $P$=.34</td>
<td>-0.002 (-0.02, 0.02), $P$=.89</td>
<td>-0.003 (-0.01, 0.004), $P$=.40</td>
</tr>
<tr>
<td>ml$_{NAWM}$</td>
<td>-0.06 (-0.14, 0.02), $P$=.16</td>
<td>0.04 (-0.01, 0.09), $P$=.14</td>
<td>-0.03 (-0.042, -0.009), $P$=.002</td>
</tr>
<tr>
<td>ml$_{NAGM}$</td>
<td>0.02 (-0.03, 0.08), $P$=.45</td>
<td>-0.015 (-0.05, 0.02), $P$=.40</td>
<td>0.0006 (-0.01, 0.01), $P$=.92</td>
</tr>
<tr>
<td>ml/NAA$_{NAGM}$ ratio</td>
<td>-0.13 (-0.75, 0.49), $P$=.69</td>
<td>-0.25 (-0.64, 0.15), $P$=.22</td>
<td>0.07 (-0.06, 0.20), $P$=.29</td>
</tr>
</tbody>
</table>

Mean and 95% confidence intervals of longitudinal effect of metabolite predictors on longitudinal change in PBVC, EDSS and zMSFC expressed in annual slope per year are shown, based on single predictor models analysis (longitudinal mixed-effects models). PBVC results are described in terms of annualized percent (%) change, whereas EDSS and zMSFC are given as annualized change in points (pts) score. Results are estimates from linear mixed effects models. Cells with statistically significant results at $\alpha = 0.05$ are highlighted in bold.

Abbreviations: PBVC, percentage of brain volume change; ml, myo-inositol; NAA, N-acetyl aspartate; NAGM, MS normal-appearing grey matter; NAWM, MS normal-appearing white matter; EDSS, Expanded Disability Status Scale; MSFC, Multiple Sclerosis Functional Composite.