Comparison of Automated Brain Segmentation Tools in FTD: Implications for Imaging Biomarker Choice in Clinical Trials

Gordon E1, Neason M1, Warren JD1, Cash D1, Bocchetta M1, Rohrer JD1

1Dementia Research Centre, UCL Queen Square Institute of Neurology, London, United Kingdom

BACKGROUND

Automated measures of volumetric brain loss are valuable markers of disease progression and commonly used as endpoints in neurodegenerative therapeutic trials. However validation of these techniques within the FTD patient population is currently lacking.

METHODS

We compared six widely used automated segmentation tools to extract whole brain volumes: BMAPS1, SPM12, GIFT1, Freesurfer (6.0)12, MALPEM (v1.2)13 and SIENAX14 in 264 individuals who had undergone 2 same-scanner 3DT1 magnetic resonance imaging (MRI). Minimal MRI interval for inclusion was 6 months (Table 1). All techniques were applied using standard published pipelines and in native space, except SIENAX which was applied to images after transformation into MNI125 space (Figure 1). For techniques that produced multi-label or multi-tissue type parcellations, all relevant regions were summed to produce a single whole brain volume. Annualised rates of change were calculated by subtraction of raw segmented volumes, or application of the boundary shift integral (BSI) in both kn-BSI (knBSI) and generalised (gBSI) forms. The longitudinal pipelines for VBM and SIENA were also applied to derive annual percentage brain volume change (PBVC).

RESULTS

Annual rates of change (Table 2) are expressed as percentage loss from baseline volume (mean (standard deviation)). All analyses were performed in STATA 14. Effect sizes with 95% confidence intervals (CI) were calculated using bootstrapping (2000 replications) and applied to all annual rates of change, comparing controls to each of the patient subgroups for all techniques. Sample size calculations to detect a 25% treatment effect with 80% power were derived based on these effect size analyses (Table 3). Key findings: There were clear qualitative and quantitative differences between techniques with BMAPS, GIFT and SPM most often performing best in accurately segmenting the whole brain structures and producing raw segmented brain volume. Over all poor biomarker with substantial improvements gained once either BSI technique was applied. Subgroups with more homogeneous rates of atrophy (PPA groups clinically, MAPT and GRN genetically, tautopaths pathologically) tended to have lower sample sizes.

CONCLUSIONS

Automated whole brain segmentation techniques are applicable across the full FTD spectrum. However performance qualitatively and quantitatively varies, so careful selection of tools will be important in forthcoming FTD therapeutic trials. Application of the BSI substantially improves the variability of measures of volume loss and should be used preferentially to raw volume change as an imaging biomarker. Enrolment targeting more stratified and homogenous groups will also prove helpful in lowering sample sizes required.

Table 1: Patient demographics for all FTD subgroups and healthy controls. Figures are in years for bottom four rows.

Table 2: Mean (standard deviation) annual rates of atrophy as a percentage loss from baseline volumes.

Table 3: Sample size estimates per treatment arm to detect a 25% reduction in volume loss at 80% power with 95% confidence intervals in parentheses.