



Published in final edited form as:

Parkinsonism Relat Disord. 2016 July ; 28: 29–35. doi:10.1016/j.parkreldis.2016.04.006.

Clinical correlates of longitudinal brain atrophy in progressive supranuclear palsy

Richard M. Tsai^{a,*}, Iryna Lobach^b, Jee Bang^a, Jennifer L. Whitwell^c, Matthew L. Senjem^c, Clifford R. Jack Jr.^c, Howard Rosen^a, Bruce Miller^a, and Adam L. Boxer^a on behalf of the AL-108-231 Investigators

^aMemory and Aging Center, University of California at San Francisco, San Francisco, CA, USA

^bDepartment of Epidemiology and Biostatistics, Division of Biostatistics, University of California at San Francisco, San Francisco, CA, USA

^cDepartment of Radiology, Mayo Clinic, Rochester, MN, USA

Abstract

Introduction—There are no effective treatments for progressive supranuclear palsy (PSP). Volumetric MRI (vMRI) may be a useful surrogate outcome measure in PSP clinical trials. The goal of the study was to evaluate the potential of vMRI to correlate with clinical outcomes from an international clinical trial population.

Methods—PSP patients (n = 198) from the AL-108-231 trial who had high quality vMRI and Progressive Supranuclear Palsy Rating Scale (PSPRS), Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), Schwab and England Activities of Daily Living (SEADL), Color Trails Test, Geriatric Depression Screen (GDS) and one year Clinician Global Impression of Change (CGIC) data from the baseline and 52 week visits were included. Linear regression was used to relate baseline values and annual clinical rating scale changes to annual regional vMRI changes (whole brain, ventricular, midbrain and superior cerebellar peduncle volumes).

Results—Effect sizes (Cohen's *d*) measuring disease progression over one year were largest for vMRI (midbrain [1.27] and ventricular volume [1.31]) but similar to PSPRS (1.26). After

*Corresponding author. Memory and Aging Center, Department of Neurology, University of California at San Francisco, 675 Nelson Rising Lane, Suite 190, San Francisco, CA 94115, USA. Richard.Tsai@ucsf.edu (R.M. Tsai).

Authorship roles

Richard M. Tsai, MD, MBA. Corresponding author, drafting and revising manuscript, analysis and interpretation of data.

Iryna Lobach, PhD. Iryna.Lobach@ucsf.edu: Analysis and interpretation of data.

Jee Bang, MD. Jee.Bang@ucsf.edu: Analysis and interpretation of data.

Jennifer L. Whitwell, PhD. Whitwell.Jennifer@mayo.edu: Revising the manuscript for content.

Matthew L. Senjem. Senjem.Matthew1@mayo.edu: Analysis and interpretation of data.

Clifford R. Jack, Jr., MD. Jack.Clifford@mayo.edu: Revising the manuscript for content.

Howard Rosen, MD. Howie.Rosen@ucsf.edu: Revising the manuscript for content.

Bruce Miller, MD. Bruce.Miller@ucsf.edu: Revising the manuscript for content.

Adam L. Boxer, MD, PhD. Adam.Boxer@ucsf.edu: Drafting/revising the manuscript for content, including final approval of the version to be submitted, study concept/design, analysis/interpretation of data.

Conflicts of interest

Drs. Tsai, Lobach, Bang, Whitwell, Jack, Rosen, Miller and Mr. Senjem report no conflicts of interest.

Dr. Boxer receives research support from Avid, Biogen, Bristol Myers Squibb, C2N Diagnostics, Cortice Biosciences, Eli Lilly, Forum Pharmaceuticals, Genentech and TauRx. He has served as a consultant for Asceneuron, Ipiarian, Isis, Janssen and Merck. He has stock/options in Alector and Delos.

multiple comparison adjustment, annual changes in PSPRS, RBANS, SEADL, Color Trails Test, GDS and one year CGIC were modestly correlated with annual vMRI changes ($p < 0.05$). Baseline neuropsychological status on RBANS ($p = 0.019$) and Color Trails ($p < 0.01$) predicted annual midbrain atrophy rates.

Conclusion—Standard vMRI measurements are sensitive to disease progression in large, multicenter PSP clinical trials, but are not well correlated with clinical changes. vMRI changes may be useful as supportive endpoints in PSP trials.

Keywords

Progressive supranuclear palsy; Clinical trials; Biomarkers; Imaging; MRI

1. Introduction

Progressive supranuclear palsy (PSP) is a neurodegenerative disease characterized by early falls, supranuclear ophthalmoplegia, axial rigidity and executive dysfunction [1]. Clinical diagnosis of the most common presentation of PSP, Richardson's syndrome, is highly predictive of abnormal tau protein deposition in neurons and glia at autopsy [2,3]. Potential therapies are entering early clinical trials. To achieve regulatory approval, therapeutics must demonstrate treatment-associated benefits on clinical or functional outcomes in a placebo controlled trial, requiring large numbers and extended longitudinal follow up. Biomarkers such as volumetric magnetic resonance imaging (vMRI) may have advantages over clinical measures from sensitivity to change and precision of measurement, leading to more efficient trials. Moreover, bio-markers may provide data supportive of a disease modifying effect [4]. In multiple sclerosis, incorporation of MRI as a biomarker has allowed more efficient determination of proof of concept for new therapeutics in phase 2 trials [5].

Previous cross-sectional vMRI studies demonstrated atrophy of the midbrain, premotor cortex, basal ganglia and thalamus in PSP [6,7]. Longitudinal studies have explored rates of brain atrophy, but were typically performed in single centers with relatively small cohorts, potentially limiting their findings' relevance to multicenter clinical trials [8–12].

To examine the utility of vMRI for measuring disease progression in multicenter clinical trials, we used longitudinal vMRI data from the AL-108-231 clinical trial of davunetide for PSP [13]. The goals were: 1) power of standard vMRI measurements to capture disease progression in a multicenter PSP clinical trial, 2) how regional vMRI changes relate to changes in clinical ratings and neuropsychological scores, and 3) value of baseline clinical measures in predicting atrophy rate.

2. Methods

2.1. Participants

The AL-108-231 clinical trial was a phase 2/3 double-blind, placebo-controlled trial, completed in 2012, conducted at 48 centers on three continents that enrolled 313 patients with PSP randomized 1:1 to davunetide (a neurotrophic peptide) or placebo for one year. 198 individuals with baseline and week 52 MRI data, complete clinical data including the

primary outcome measures: Progressive Supranuclear Palsy Rating Scale (PSPRS) [14] and Schwab and England Activities of Daily Living scales (SEADL) [15] were included. All participants met criteria for PSP (Richardson's syndrome), modified from the National Neuroprotection and Natural History in Parkinson Plus Syndromes (NNIPPS) study: at least a 12-month history of postural stability or falls during the first 3 years of symptoms, reduced downward saccade velocity or supranuclear ophthalmoplegia, and an akinetic-rigid syndrome with axial rigidity [16]. Additional details were in the original trial manuscript [13]. Ethics approval was obtained at each site from the local ethics committee and all participants gave written informed consent at recruitment.

2.2. Clinical and neuropsychological assessments

The primary outcomes were the PSPRS and SEADL. The PSPRS comprises 28 items in six categories: history, mentation, bulbar, limb, gait, oculomotor function and separate scores are generated for each domain [14]. Neuropsychological outcome was measured via the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). RBANS yields a total score and sub-scores for five cognitive domains: immediate memory, delayed memory, attention, visual spatial and language, with higher scores reflecting better performance [17]. Other clinical outcomes included Color Trails Test [18], Clinical Global Impression of Change (CGIC) [19], Clinical Global Impression of Disease Severity (CGIDs) [20] and the Geriatric Depression Screen (GDS) to assess mood [21].

2.3. MRI

T1 MRI scans were performed on 1.5 or 3T scanners and all sites were qualified by the Mayo Clinic Aging and Dementia Imaging Research (ADIR) laboratory (Rochester MN USA). Scans were performed at baseline (-6 to 0 weeks) and 52 weeks per the clinical trial protocol. Regional brain volume analyses methods included in this study were pre-specified as a priori secondary or exploratory outcome measures under special agreement with the FDA prior to trial start. Whole brain and ventricular volume changes were generated using the boundary shift integral technique [13,22]. Midbrain and superior cerebellar peduncle (SCP) volumes were measured using label propagation in SPM5 [13,23].

2.4. Analysis

Data from placebo and treatment groups were combined since there were no group differences in any endpoints, including exploratory voxel based morphometry, from the original study [13]. Differences in regional brain volumes between the baseline and 52 week MRI were expressed as percent change from baseline scan. To determine whether vMRI changes were limited by floor effect, the bottom quartile volumes of each vMRI measure at screening were compared to week 52 volumes using a two-sample paired *t*-test.

Linear regression was used to investigate the relationship between baseline or annual changes in PSPRS, RBANS, Color Trails Test, SEADL, GDS scores, CGIC scores and annual changes in whole brain, ventricular volume, midbrain and SCP volumes while controlling for total intracranial volume. Nonparametric variables were transformed via Box-Cox transformation [24]. We presented the relationship as R^2 to illustrate the proportion of inter-patient variability of vMRI that is explained by the variability in clinical outcomes.

Demographic and clinical covariates of interest included in the models were age, sex, coenzyme Q10 use during study, tau haplotype, treatment group assignment and disease duration (greater than or less than or equal to 5 years). Due to floor effects, data from participants with the maximum time for Color Trails Test at baseline were excluded from analyses involving this measurement.

Correlations between MRI volume changes and clinical scales may be influenced by differences in scanner field strength. Subjects scanned on either 1.5T or 3T scanners were determined by site availability and no patients switched scanners during the duration of the trial. To investigate this possibility, images from 1.5T or 3T scanners were analyzed separately for significant correlations between vMRI and the primary outcome measures PSPRS, SEADL. The original trial also stratified for disease duration due to concern that longer disease duration may affect disease progression. Subjects with longer disease duration (greater than five years) were analyzed as a subgroup and compared to the remainder to determine if correlations are affected by disease duration. Sample size estimates were performed to estimate minimum sample sizes required to detect 10%, 25%, 37.5%, 50% change in progression of brain volume atrophy, changes in PSPRS, RBANS and SEADL using a two-tailed, two-sample *t*-test with an α of 0.05, and a power (β) of 0.90. We derived numbers of subjects per group incorporating the 23% attrition rate that was observed in the AL-108-231 study. Effect sizes reported as Cohen's *d*.

Statistically significant relationships were defined by $p < 0.05$ corrected for multiple comparisons using False Discovery Rate [25]. *p*-values reported for regression analysis are False Discovery Rate adjusted unless otherwise specified. Analysis was performed using Stata version 13.

3. Results

3.1. Demographics

Patient characteristics are shown in Table 1. The baseline demographic and clinical scale values of this group were similar to the overall study population [13]. In addition, the primary outcomes at week 52 for this study were no different compared to overall study population at the end of trial (data not shown) [13]. Clinical progression over 1 year was measured by an increase of PSPRS 9.6 ± 0.6 points and SEADL declined $16 \pm 1\%$. Cognitive function worsened over one year while mood measured by GDS remained stable. Table 1 also shows the 1 year changes from baseline in clinical and imaging outcome measures as well as its corresponding effect size. The mean volume of the bottom quartiles of vMRI outcomes at baseline then week 52 showed changes in the expected direction, suggesting no floor effect was present ($p < 0.001$ uncorrected) (data not shown). Midbrain, ventricular volume, PSPRS and SEADL changes provided the largest effect size and were not significantly different. Consistent with the a priori hypothesis, subjects with longer disease duration had similar brain atrophy rates but significantly higher baseline PSPRS score (47.4 ± 2 vs 38.3 ± 0.8 $p < 0.01$) and slower disease progression by annual PSPRS change compared to those with disease duration less than or equal to 5 years (2.4 ± 1.7 vs 10.4 ± 0.6 , $p < 0.01$).

3.2. Relationship between annual changes in clinical scales and vMRI measures

Whole brain atrophy rates correlated most strongly with clinical measures including PSPRS ($R^2 = 0.168$, $p = 0.002$), SEADL ($R^2 = 0.095$, $p = 0.025$), CGIC ($R^2 = 0.161$, $p = 0.003$), RBANS ($R^2 = 0.016$, $p = 0.016$), GDS ($R^2 = 0.055$, $p = 0.020$). Ventricular expansion correlated with fewer clinical measures, and not with PSPRS. Midbrain atrophy rates correlated with three clinical measures, including PSPRS ($R^2 = 0.159$, $p = 0.036$) (Fig. 1A) (Table 2). Superior cerebellar peduncle atrophy rates did not correlate with any clinical measures.

We explored whether the modest correlations between MRI volume changes and clinical scales might have been influenced by differences in scanner field strength at each study center. 96 and 102 subjects were scanned on 1.5T and 3T scanners respectively. 3T whole brain, midbrain to PSPRS ($R^2 = 0.215$ $p = 0.061$, $R^2 = 0.234$ $p = 0.024$) was compared to 1.5T whole brain, midbrain to PSPRS ($R^2 = 0.244$ $p = 0.024$, $R^2 = 0.221$ $p = 0.095$). Similar results were seen comparing 3T whole brain, ventricular volume ($R^2 = 0.133$ $p = 0.036$, $R^2 = 0.151$ $p = 0.024$) to 1.5T whole brain, ventricular volume ($R^2 = 0.102$ $p = 0.131$, $R^2 = 0.169$ $p = 0.012$) regressed to SEADL, suggesting the ability to predict variance did not favor a particular scanner field strength. We also determined that the relationship between vMRI changes and clinical scale changes was not modified by field strength in a regression model (data not shown).

Subjects with longer disease duration were analyzed separately to assess if different disease progression may affect vMRI correlations. Subjects with longer disease duration did not show significant relationships between annual brain volume changes and clinical scale changes. The remainder cohort of subjects with short disease duration retained but did not improve upon the significant correlations between vMRI and clinical scales shown by the overall population (data not shown).

3.3. Relationship between baseline status and annual vMRI changes

Annual vMRI changes were regressed to baseline outcome measurements to assess if prediction of atrophy rates was possible. Baseline neuropsychological status (RBANS [$R^2 = 0.107$, $p = 0.019$] and Color Trails 1 [$R^2 = 0.111$, $p = 0.004$] and 2 [$R^2 = 0.141$, $p = 0.006$]) predicted annual midbrain atrophy rates (Fig. 1B). No other baseline clinical scales or MRI volumes predicted annual changes in vMRI measures.

3.4. Sample size estimates for volumetric and clinical measures

We estimated the number of patients per arm necessary to detect different treatment effects on brain volume and clinical measures, accounting for the 23% observed attrition rate in the AL-108-231 study (Table 3). Imaging biomarkers (midbrain, ventricular volume) and clinical outcomes (PSPRS) provided comparable and smallest sample size estimates. Sample size estimates using whole brain, SCP, neuropsychology testing via RBANS were large.

4. Discussion

Using data from a large, multicenter, international clinical trial, we show that longitudinal vMRI measurements of regional brain atrophy in PSP are feasible and sensitive to disease progression over one year, but only modestly correlate with change measured by clinical scales and neuropsychology testing. Whole brain, midbrain atrophy and ventricular volume changes were associated with clinical disease progression measured by PSPRS, SEADL, GDS and CGIC, and cognitive decline measured by RBANS and Color Trails Test. SCP volume changes were marginally correlated with clinical measures. Of the vMRI measures examined, midbrain atrophy and ventricular volume expansion had the largest effect sizes leading to the smallest sample sizes if used as outcome measures in future trials, but were similar to PSPRS. Baseline neuropsychological status was most predictive of annual midbrain volume changes. These analyses suggest that vMRI provides complementary information to clinical rating scales in PSP trials.

4.1. vMRI measurements are a feasible outcome in international, multicenter trials

The rates of whole brain and ventricular volume changes per annum from this multicenter clinical trial were similar to those reported in single center natural history studies [8-12]. The 3.5% annual rate of midbrain volume loss in this study was smaller than the annual 10% reduction reported in two previous studies. The differences may arise from different imaging measurements or differences in study populations [9-11]. Our annual midbrain atrophy rate was similar to that reported in another case series [12]. In addition, a recent PSP trial utilizing tideglusib, a glycogen synthase kinase 3 inhibitor, included a smaller MRI sub study (n = 37) which showed reduced treatment group whole brain, parietal-occipital atrophy rate without clinical benefits [26]. The whole brain atrophy rate in the treatment group and midbrain atrophy rate in both groups were similar to our results. The placebo group's 3.1% whole brain atrophy rate was larger than this study. This may have reflected different volumetric measurement methods used.

Sample size calculations suggest midbrain and ventricular volume are the most sensitive outcome measures, with largest effect sizes. Of the clinical measures, PSPRS was close in terms of sample size efficiency and effect size significance. These results are in line with previous sample size estimates, where one study found midbrain atrophy (84 subjects per arm with 90% power for 30% effect) and ventricular expansion over one year offered the smallest sample size estimates, followed by PSPRS and all three were superior to whole brain volume change.¹⁴ Another study also found sample size estimates using midbrain atrophy to be more efficient than whole brain and SCP volume changes [27]. Our estimates may be more generalizable to future trials since we used data from a 48 site, international clinical trial as compared to previous single center studies [9-27].

4.2. vMRI is sensitive to but only modestly correlated with clinical change

Clinical progression over 1 year measured by PSPRS and SEADL was similar to previous reports and the results of the AL-108-231 trial, suggesting that our population with complete MRI data is representative of a typical PSP population [13-14]. The correlation between annual brain volume changes and clinical measures changes provide support for use of

vMRI as an outcome measure in PSP clinical trials. Our data suggest that vMRI changes relate to changes in motor, cognitive, functional outcomes and clinicians' disease state impression. It is interesting to note that change in RBANS visual construction sub-score had a strong relationship with all four regions of brain volume change. This may reflect the saccade control network's role (connections from the cerebellum via the SCP to midbrain, subcortical and cortical regions) in visual construction.

The US FDA currently allows accelerated drug approval for serious, life-threatening illnesses based on the effect of a drug on a surrogate marker and not a clinical outcome under subpart H of 21 CFR 314.500 (Code of Federal Regulations). A surrogate endpoint is defined as “a substitute for a clinically meaningful endpoint that is a direct measure of how a patient feels, functions or survives and is expected to predict the effect of a therapy [28].”

In our study, although the significant relationships between brain volume changes and clinical changes support use of imaging as a surrogate outcome measure, the correlation coefficients from these models suggest only 20–30% of the variance in clinical changes was associated with brain volume changes. Between subjects, there was considerable variability in the relationship between vMRI and clinical scales even after controlling for potential confounders. Floor effects did not confound our vMRI measurements. Thus although vMRI reflects PSP disease progression, the magnitude of the clinical correlation suggests that the vMRI measurements we used may not directly capture clinically relevant changes necessary for regulatory approval of new therapeutics. Thus it is unlikely that vMRI will be able to serve as a surrogate outcome measure in future trials. However, due to its sensitivity to change, vMRI outcomes may capture physiologic changes that elude clinical scales. Therefore, vMRI has potential use as an outcome measure to detect disease modifying effect, and may be helpful for small, proof of concept studies.

4.3. vMRI may be useful in special populations

Longer disease duration in PSP may reflect lower oligodendroglial tau burden [29]. We found subjects with longer disease duration had decreased rate of PSPRS progression but similar rate of brain atrophy compared to those with less than 5 years of disease. Longer disease duration subjects had no correlations between annual brain atrophy rates and clinical scale changes. The number of subjects with longer disease duration was small ($n = 18$) and had limited power to detect an effect. It is also possible these individuals reached a ceiling on PSPRS since their baseline mean PSPRS score was higher than the remainder of the population. Thus vMRI may be particularly useful in monitoring disease progression in clinically advanced PSP patients.

4.4. Limitations

A limitation of this study is that not all patients from the original cohort completed annual MRI, and clinical evaluations. Discontinuation reasons included death and adverse events leading to drug termination. This could have introduced a bias to our analysis since more impaired patients may have dropped out prior to follow up MRI scans. Some end of study scans were of insufficient quality for analysis. However, the primary outcomes at start and week 52 in this study were no different compared to the overall population at start and end

of the original trial, suggesting our cohort is representative of the original trial population. Another limitation is that study participants were not autopsy confirmed PSP, though this represents a real world clinical trial, and historically, clinically diagnosed PSP patients have a high likelihood of PSP pathology [2,16]. Furthermore, the study included only PSP-Richardson's presentation, and atypical PSP subtypes may have differential vMRI changes and clinical correlations that cannot be explored here. Finally, use of baseline image as reference produces systematic asymmetric risk that may introduce bias to imaging measurements [30], and our study did not examine regional brain volume changes, which could be more sensitive to change.

Newer vMRI analysis methods might strengthen the relationship between brain atrophy and clinical changes. The wide range of atrophy rates here may be a result of scanner differences or technical limitations of the volumetric methods. Other regions of interest, updated longitudinal protocols, or volumetric measurement tools may produce less variable results. Scanner differences such as software, manufacturer might also affect image outcomes, though it may not be realistic for future multi-center trials to insist on uniform scanners. Structural imaging modalities such as diffusion tensor imaging or functional imaging and their use as biomarkers should be investigated.

4.5. Conclusion

This study confirms that volumetric MRI is a useful biomarker for measuring disease progression in PSP clinical trials. While it is unlikely to sufficiently correlate with clinical measures to serve as a true surrogate outcome measure of disease progression, vMRI changes could provide supportive evidence for a disease modifying effect of an experimental therapy in future studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

This work was supported by the National Institute of Health/National Institute of Aging (grant numbers K24 AG045333-01, R01 AG032306).

Study funding

R01AG038791, U54NS092089, Tau Consortium and Allon Therapeutics.

References

- [1]. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, Goetz CG, Golbe LI, Grafman J, Growdon JH, Hallett M, Jankovic J, Quinn NP, Tolosa E, Zee DS. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology*. 1996; 47:1–9. [PubMed: 8710059]
- [2]. Hughes AJ, Daniel SE, Ben-Shlomo Y, Lees AJ. The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain*. 2002; 125:861–870. [PubMed: 11912118]
- [3]. Respondek G, Stamelou M, Kurz C, Ferguson LW, Raiput A, Chiu WZ, van Swieten JC, Troakes C, Al Sarraj S, Gelpi E, Gaig C, Tolosa E, Oertel WH, Giese A, Roerber S, Arzberger T,

- Wagenpfeil S, Höglinger GU. Movement Disorder Society-endorsed PSP Study Group, The phenotypic spectrum of progressive supranuclear palsy: a retrospective multicenter study of 100 definite cases. *Mov. Disord.* 2014; 29:1758–1766. [PubMed: 25370486]
- [4]. Blennow K. Biomarkers in Alzheimer's disease drug development. *Nat. Med.* 2010; 16:1218–1222. [PubMed: 21052077]
- [5]. Barkhof F, Simon JH, Fazekas F, Rovaris M, Kappos L, de Stefano N, Polman CH, Petkau J, Radue EW, Sormani MP, Li DK, O'Connor P, Montalban X, Miller DH, Filippi M. MRI monitoring of immunomodulation in relapse-onset multiple sclerosis trials. *Nat. Rev. Neurol.* 2011; 8:13–21. [PubMed: 22143362]
- [6]. Josephs KA, Whitwell JL, Dickson DW, Boeve BF, Knopman DS, Petersen RC, Parisi JE, Jack CR Jr. Voxel-based morphometry in autopsy proven PSP and CBD. *Neurobiol. Aging.* 2008; 29:280–289. [PubMed: 17097770]
- [7]. Boxer AL, Geschwind MD, Belfor N, Gorno-Tempini ML, Schauer GF, Miller BL, Weiner MW, Rosen HJ. Patterns of brain atrophy differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Arch. Neurol.* 2006; 63:81–86. [PubMed: 16401739]
- [8]. Whitwell JL, Jack CR Jr. Parisi JE, Knopman DS, Boeve BF, Petersen RC, Ferman TJ, Dickson DW, Josephs KA. Rates of cerebral atrophy differ in different degenerative pathologies. *Brain.* 2007; 130:1145–1158.
- [9]. Whitwell JL, Xu J, Mandrekar JN, Gunter JL, Jack CR Jr. Josephs KA. Rates of brain atrophy and clinical decline over 6 and 12-month intervals in PSP: determining sample size for treatments trials. *Park. Relat. Disord.* 2012; 18:252–256.
- [10]. Josephs KA, Whitwell JL, Boeve BF, Shiung MM, Gunter JL, Parisi JE, Dickson DW, Jack CR. Rates of cerebral atrophy in autopsy-confirmed progressive supranuclear palsy. *Ann. Neurol.* 2006; 59:200–203. [PubMed: 16278856]
- [11]. Josephs KA, Xia R, Mandrekar J, Gunter JL, Senjem ML, Jack CR Jr. Whitwell JL. Modeling trajectories of regional volume loss in progressive supranuclear palsy. *Mov. Disord.* 2013; 28:1117–1124. [PubMed: 23568852]
- [12]. Paviour DC, Price SL, Jahanshahi M, Lees AJ, Fox NC. Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy. *Brain.* 2006; 129:1040–1049. [PubMed: 16455792]
- [13]. Boxer AL, Lang AE, Grossman M, Knopman DS, Miller BL, Schneider LS, Doody RS, Lees A, Golbe LI, Williams DR, Corvol JC, Ludolph A, Burn D, Lorenzl S, Litvan I, Roberson ED, Höglinger GU, Koestler M, Jack CR Jr. Van Deerlin V, Randolph C, Lobach IV, Heuer HW, Gozes I, Parker L, Whitaker S, Hirman J, Stewart AJ, Gold M, Morimoto BH. AL-108-231 Investigators. Davunetide in patients with progressive supranuclear palsy: a randomized, double-blind, placebo-controlled phase 2/3 trial. *Lancet Neurol.* 2014; 13:676–685. [PubMed: 24873720]
- [14]. Golbe LI, Ohman-Stricklan PA. A clinical rating scale for progressive supranuclear palsy. *Brain.* 2007; 130:1552–1565. [PubMed: 17405767]
- [15]. Schwab, R.; England, A.; Gillingham, F.; Donaldson, M. Projection technique for evaluating surgery in Parkinson's disease. Third Symposium on Parkinson's Disease Research, ES Livingstone; Edinburgh, Scotland. 1969.
- [16]. Bensimon G, Ludolph A, Agid Y, Vidailhet M, Payan C, Leigh PN, NNIPPS Study Group. Riluzole treatment, survival and diagnostic criteria in Parkinson plus disorders: the NNIPPS study. *Brain.* 2009; 132:156–171. [PubMed: 19029129]
- [17]. Randolph C, Tierney MC, Mohr E, Chase TN. The repeatable battery for the assessment of neuropsychological status (RBANS): preliminary clinical validity. *J. Clin. Exp. Neuropsychol.* 1998; 20:310–319. [PubMed: 9845158]
- [18]. Dugbartey AT, Townes BD, Mahurin RK. Equivalence of the color trails test and trail making test in nonnative English-speakers. *Arch. Clin. Neuropsychol.* 2000; 15:425–431. [PubMed: 14590218]
- [19]. Schneider LS, Olin JT, Doody RS, Clark CM, Morris JC, Reisberg B, Schmitt FA, Grundman M, Thomas RG, Ferris SH. Validity and reliability of the Alzheimer's disease cooperative study-clinical global impression of change. The Alzheimer's disease cooperative study. *Alzheimer Dis. Assoc. Disord.* 1997; 11:S22–S32. [PubMed: 9236949]

- [20]. Payan CA, Viallet F, Landwehrmeyer BG, Bonnet AM, Borg M, Durr F, Lacomblez L, Verny M, Fermanian J, Agid Y, Ludolph AC, Leigh PN, Bensimon G, NNIPPS Study Group. Disease severity and progression in progressive supranuclear palsy and multiple system atrophy: validation of the NNIPPS-Parkinson Plus Scale. *PLoS One*. 2011; 6:e22293. [PubMed: 21829612]
- [21]. Yesavage JA, Brink TL, Role TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression scale: a preliminary report. *J. Psychiatr. Res.* 1983; 17:37–49. [PubMed: 7183759]
- [22]. Gunter JL, Shiung MM, Manduca A, Jack CR Jr. Methodological considerations for measuring rates of brain atrophy. *J. Magn. Reson Imaging*. 2003; 18:16–24. [PubMed: 12815635]
- [23]. Botha H, Whitwell JL, Madhavan A, Senjem ML, Lowe V, Josephs KA. The pimple sign of progressive supranuclear palsy syndrome. *Park. Relat. Disord.* 2014; 20:180–185.
- [24]. Weisberg, S., editor. *Applied Linear Regression*. fourth ed. Wiley Wiley; 2014. Box-Cox transformation, skew power family.
- [25]. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B.* 1995; 57:289–300.
- [26]. Höglinger GU, Huppertz HJ, Wagenpfeil S, Andrés MV, Belloch VV, León T, Del Ser T. TAUROS MRI Investigators, Tideglusib reduces progression of brain atrophy in progressive supranuclear palsy in a randomized trial. *Mov. Disord.* 2014; 29:479–487. [PubMed: 24488721]
- [27]. Paviour DC, Prie SL, Lees AJ, Fox NC. MRI derived brain atrophy in PSP and MSA-P. Determining sample size to detect treatment effects. *J Neurol.* 2007; 254:478–481. [PubMed: 17401522]
- [28]. Katz R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx*. 2004; 1:189–195. [PubMed: 15717019]
- [29]. Josephs KA, Mandrekar JN, Dickson DW. The relationship between histo-pathological features of progressive supranuclear palsy and disease duration. *Park. Relat. Disord.* 2006; 12:109–112.
- [30]. Leung KK, Ridgway GR, Ourselin S, Fox NC. Alzheimer’s Disease Neuro-imaging Initiative, Consistent multi-time-point brain atrophy estimation from the boundary shift integral. *Neuroimage*. 2012; 50:3995–4005. [PubMed: 22056457]

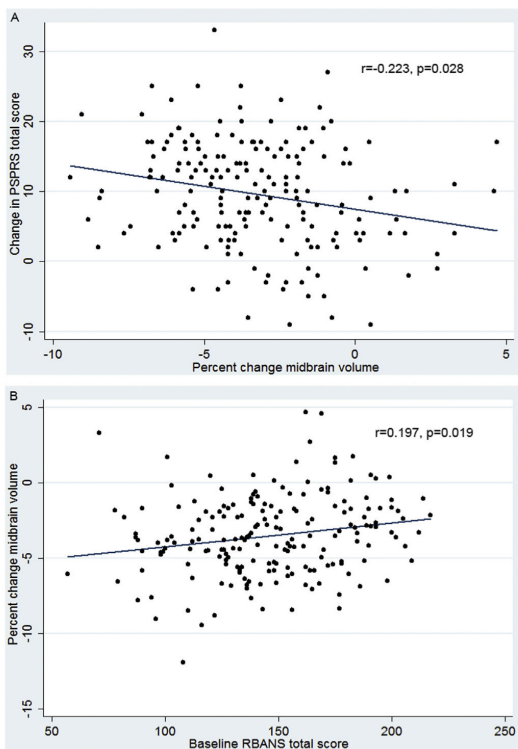


Fig. 1. A) Annual midbrain volume change is correlated to annual PSPRS change, B) Baseline RBANS score predicts 1 year midbrain volume changes. A) Worsening PSPRS relates to increased midbrain atrophy. 1 year correlation by Pearson (r) between changes in PSPRS and changes in midbrain volume, p-value corrected for multiple comparisons using false discovery rate. PSPRS = progressive supranuclear palsy rating scale. B) Poor performance on RBANS at baseline predicts increased midbrain atrophy over one year, Pearson correlation (r) shown, p-value corrected for multiple comparison using false discovery rate. RBANS = repeatable battery for the assessment of neuropsychological disease severity.

Table 1

Baseline demographic and 1 year change in imaging, clinical measures of study population.

N = 198	Baseline	Week 52	Mean change	Range	Mean % change	Effect size
Demographics						
Age (years)	68 (0.4)	–	–	–	–	–
Sex, female	93 (47%)	–	–	–	–	–
Coenzyme Q10 use	41 (21%)	–	–	–	–	–
Treatment Group	105 (53%)	–	–	–	–	–
Disease Duration >5 years	18 (9.1%)	–	–	–	–	–
Tau Haplotype						
H1/H1	149 (75%)	–	–	–	–	–
H1/H2	42 (21%)	–	–	–	–	–
Missing	7 (4%)	–	–	–	–	–
MMSE	26.6 (0.2)	–	–	–	–	–
Clinical outcomes						
PSPRS	38.9 (0.8)	48.8 (1.0)	9.6 (0.6)	–9 to 33	24.7	1.26
SEADL	54% (1.5)	38% (1.4)	–16% (1)	–60 to 20%	–30	1.10
CGIC	–	5.39 (0.1)	–	–	–	–
CGIDs	3.87 (0.1)	4.74 (0.1)	0.9 (0.1)	–1 to 6	23.2	1.01
RBANS total raw	146.3 (2.4)	131.5 (2.9)	–18.4 (1.5)	–91 to 16	–12.5	0.94
RBANS total scaled	74.9 (0.9)	71.0 (1.1)	–5.1 (0.6)	–29 to 14	–6.8	0.65
Color Trails Test 1	157.4 (5.1)	167.8 (5.7)	43.7 (5.0)	–120 to 202	27.8	0.76
Color Trails Test 2	231.9 (5.4)	238.2 (6.2)	46.2 (5.9)	–138 to 220	20.0	0.73
GDS	12.7 (0.5)	13.1 (0.6)	0.4 (0.4)	–13 to 14	3.1	0.08
Imaging measures						
Whole brain volume	1.3 × 10 ⁶ (9107)	1.3 × 10 ⁶ (9056)	–1.1 × 10 ⁴ (930)	–4.6 × 10 ⁴ to 2.7 × 10 ⁴	–0.8	0.80
Ventricular volume	4.73 × 10 ⁴ (1545)	5.14 × 10 ⁴ (1630)	4139 (258)	–3207 to 2.1 × 10 ⁴	9.1	1.31
Midbrain volume	6888 (63.5)	6649 (62.7)	–239 (13.5)	–813 to 320	–3.5	1.27
SCP volume	388 (7.6)	362 (7.7)	–26.5 (2.7)	–157 to 88	–6.9	0.72

Data are mean (SE) or number (%).

MMSE = mini-mental State Examination, PSPRS = progressive supranuclear palsy rating scale, SEADL = Schwab and England activities of daily living, CGIC = clinical global impression of change, CGIDs = clinical global impression of disease severity, RBANS = repeatable battery for the assessment of neuropsychological disease severity, GDS = geriatric depression screen.

Brain volume reported in mm³.

Table 2

Coefficient of determination (R^2) from linear regression model describing association between 1 year change in clinical outcomes and annualized brain volume change.

	Unadjusted model				Adjusted model ^a			
	Whole brain	Ventricular volume	Midbrain	SCP	Whole brain	Ventricular volume	Midbrain	SCP
PSPRS total	0.065 ^b	0.036	0.049 ^b	0.010	0.168 ^b	0.141	0.159 ^b	0.125
PSPRS bulbar	<0.001	<0.001	0.004	0.015	0.034	0.034	0.040	0.052
PSPRS gait	0.026	0.051 ^b	0.022	0.007	0.139	0.157	0.140	0.129
PSPRS history	0.057 ^b	0.023	0.019	0.009	0.143 ^b	0.114	0.113	0.102
PSPRS limb	0.036 ^b	0.003	0.020	0.003	0.077	0.047	0.062	0.050
PSPRS mentation	0.014	<0.001	0.012	<0.001	0.043	0.029	0.039	0.028
PSPRS ocular	0.003	0.014	0.013	0.007	0.050	0.057	0.058	0.056
SEADL	0.046 ^b	0.090 ^b	0.047 ^b	0.017	0.095 ^b	0.131 ^b	0.087	0.069
CGIC	0.064 ^b	0.032 ^b	0.082 ^b	0.033	0.161 ^b	0.129 ^b	0.167 ^b	0.122
RBANS total	0.037 ^b	0.078 ^b	0.023	0.017	0.107 ^b	0.145 ^b	0.079	0.074
RBANS attention	0.012	<0.001	0.007	0.015	0.081	0.077	0.090	0.100
RBANS delayed memory	0.001	0.008	0.010	0.001	0.090	0.107	0.099	0.090
RBANS immediate memory	0.034	0.049 ^b	<0.001	0.001	0.079 ^b	0.097 ^b	0.034	0.034
RBANS language	0.007	0.029	0.001	0.001	0.043	0.076	0.036	0.037
RBANS visual spatial	0.065 ^b	0.089 ^b	0.050 ^b	0.045 ^b	0.144 ^b	0.179 ^b	0.114 ^b	0.110 ^b
Color Trails 1	0.018	0.056 ^b	0.056 ^b	0.011	0.054	0.084 ^b	0.086 ^b	0.050
Color Trails 2	0.036	0.045	0.087 ^b	0.004	0.093	0.103	0.115	0.081
GDS	0.041 ^b	0.029	0.008	0.004	0.055 ^b	0.042	0.020	0.017

SCP = superior cerebellar peduncle, PSPRS = progressive supranuclear palsy rating scale, SEADL = Schwab and England activities of daily living, CGIC = clinical global impression of change, RBANS = repeatable battery for the assessment of neuropsychological disease severity, GDS = geriatric depression screen.

^aRegression model adjusted for age, sex, treatment group, tau haplotype, disease duration, use of coenzyme Q10 and total intracranial volume.

^bSignificant ($p < 0.05$) after multiple comparison adjustment using False Discovery Rate.

Table 3

Estimated sample size required per treatment arm in a one year parallel group clinical trial for imaging and clinical outcomes, incorporating 23% attrition rate.

Effect size	Imaging outcome measures				Clinical outcome measures		
	Whole brain	Ventricular volume	Midbrain	SCP	PSPRS	RBANS	SEADL
10%	2222	788	870	2606	868	1532	1132
25%	358	130	139	417	142	249	184
37.5%	151	60	64	188	65	112	83
50%	87	35	36	108	38	65	48

SCP = superior cerebellar peduncle, PSPRS = progressive supranuclear palsy rating scale, RBANS = repeatable battery for the assessment of neuropsychological disease severity, SEADL = Schwab and England activities of daily living.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript