Published in final edited form as:

Parkinsonism Relat Disord. 2012 March ; 18(3): 252–256. doi:10.1016/j.parkreldis.2011.10.013.

Rates of brain atrophy and clinical decline over 6 and 12-month intervals in PSP: determining sample size for treatment trials

Jennifer L. Whitwell, PhD¹, Jia Xu, MS², Jay Mandrekar, PhD², Jeffrey L. Gunter, PhD³, Clifford R. Jack Jr, MD¹, and Keith A. Josephs, MD, MST, MSc⁴

¹Department of Radiology, Mayo Clinic, Rochester, MN

²Department of Health Sciences Research (Biostatistics), Mayo Clinic, Rochester, MN

³Department of Information Technology, Mayo Clinic, Rochester, MN

⁴Department of Neurology, Mayo Clinic, Rochester, MN

Abstract

Imaging biomarkers are useful outcome measures in treatment trials. We compared sample size estimates for future treatment trials performed over 6 or 12-months in progressive supranuclear palsy using both imaging and clinical measures. We recruited 16 probable progressive supranuclear palsy patients that underwent baseline, 6 and 12 month brain scans, and 16 age-matched controls with serial scans. Disease severity was measured at each time-point using the progressive supranuclear palsy rating scale. Rates of ventricular expansion and rates of atrophy of the whole brain, superior frontal lobe, thalamus, caudate and midbrain were calculated. Rates of atrophy and clinical decline were used to calculate sample sizes required to power placebo-controlled treatment trials over 6 and 12-months. Rates of whole brain, thalamus and midbrain atrophy, and ventricular expansion, were increased over 6 and 12-months in progressive supranuclear palsy compared to controls. The progressive supranuclear palsy rating scale increased by 9 points over 6-months, and 18 points over 12-months. The smallest sample size estimates for treatment trials over 6-months were achieved using rate of midbrain atrophy, followed by rate of whole brain atrophy and ventricular expansion. Sample size estimates were further reduced over 12-month intervals. Sample size estimates for the progressive supranuclear palsy rating scale were worse than imaging measures over 6-months, but comparable over 12-months. Atrophy and clinical decline can be detected over 6-months in progressive supranuclear palsy. Sample size estimates suggest that treatment trials could be performed over this interval, with rate of midbrain atrophy providing the best outcome measure.

Keywords

Progressive supranuclear palsy; atrophy; midbrain; power calculations; short interval

INTRODUCTION

Progressive supranuclear palsy (PSP) is a neurodegenerative disorder characterized by vertical supranuclear gaze palsy, axial rigidity, and postural instability that lead to falls early in the disease course[1]. Imaging studies demonstrate that PSP is associated with atrophy of the brain, particularly the midbrain, premotor cortex, basal ganglia and thalamus[2–5]. Although specific pathological diagnoses underlying PSP can vary, the vast majority have

deposition of the abnormal protein tau, and hence PSP is classified as a tauopathy[6]. In the absence of a biomarker specific for tau, clinically diagnosed PSP is perhaps the best available construct to test therapies aimed at targeting tau. Indeed, clinical trials are already underway in PSP assessing treatments for tau. It is therefore critically important to identify disease biomarkers that can serve as outcome measures in these treatment trials.

Rates of brain atrophy measured over serial MRI have the potential to be one such biomarker. Rates of whole brain atrophy and ventricular expansion are increased in PSP compared to controls[7, 8], and rates of atrophy of regional structures, such as the midbrain and frontal lobe, are also increased in PSP[9]. Rates of brain atrophy correlate with clinical dysfunction and may hence be useful biomarkers of disease progression in PSP[9].

In this study, we aimed to determine whether rates of whole brain and regional atrophy could be useful outcome measures for treatment trials in PSP. We were particularly interested in determining whether sample size estimates would be reasonable for treatment trials performed over a short interval of only 6-months, compared to the more typical interval of 12-months. Shorter clinical treatment trials would serve to reduce patient burden, which could be particularly important for patients that progress rapidly[10], and would significantly shorten the time taken to develop successful treatments. In addition, we aimed to determine how sample size estimates for these imaging outcome measures compare to those calculated using a clinical rating scale developed to assess disease severity in PSP; the PSP rating scale (PSPRS)[11]. The PSPRS assesses all important signs and symptoms associated with PSP, including behavioral change, ocular motor deficits, gait, and balance impairment.

METHODS

Subjects

Between August 1st 2009 and July 8th 2010, we prospectively recruited all consecutive PSP subjects that were referred to the Department of Neurology at the Mayo Clinic, Rochester, Minnesota. All subjects were evaluated by a neurodegenerative specialist and PSP expert (KAJ), received a clinical diagnosis of PSP, and underwent standardized testing that included a volumetric MRI. Patients were followed and received follow-up MRI and clinical assessments at 6 and 12-month intervals.

Inclusion criteria—Subjects must have met the NINDS-SPSP criteria for probable PSP[1]. Specifically, subjects must have been age 40 or older with symptoms being gradually progressive, and presence of vertical (upward or downward gaze) supranuclear palsy and prominent postural instability with falls in the first year of onset. In addition, there could be no evidence of another disease to explain the symptoms and patients would have to meet mandatory exclusion criteria[1]. All subjects must have completed a usable volumetric MRI at baseline, 6 and 12-months.

Exclusion criteria—Subjects were excluded from the study if they met NINDS-SPSP criteria for possible PSP[1] or did not have usable MRI at all time-points.

Over this time period, 27 patients were evaluated that met NINDS-SPSP criteria for possible or probable PSP[1]. Seven were excluded because they only met criteria for possible PSP and four were excluded because they did not return for the 12-month MRI (one patient died, one underwent a shunt and developed a subdural hematoma, and two patients dropped out). Therefore, 16 patients were included in the study. Informed consent was obtained from all subjects for participation in the studies, which were approved by the Mayo IRB. All subjects underwent detailed clinical and neurological examination with standardized and valid tests,

including the PSPRS[11]. The total PSPRS, and mentation, ocular-motor and gait/midline subscores were outputted at all time-points. We selected these specific subscores since they assess cardinal features of PSP. Eleven subjects (69%) were started on treatments for parkinsonism during the first few months of the study, although all but one were subsequently titrated off medication within 3 months due to lack of improvement in symptoms. The other five subjects were not on medications for parkinsonism during the course of the study. The mean (standard deviation) age in the 16 patients at baseline was 69 (7) years with nine (56%) female patients. The PSP patients had an average disease duration of 3.1 (1.2) years. The 16 PSP patients were matched by age and gender to a cohort of healthy control subjects (age 71 (5), nine (56%) female) that had serial volumetric MRI with a mean scan interval of 14.4 (1.3) months.

MRI analysis

All subjects had volumetric MRI performed at 3T using a standardized protocol. A 3D magnetization prepared rapid acquisition gradient echo (MPRAGE) was performed at each time-point using the following parameters: TR/TE/T1, 2300/3/900 ms; flip angle 8°, 26-cm FOV; 256 × 256 in-plane matrix with a phase FOV of 0.94, slice thickness of 1.2 mm, in-plane resolution 1. All MPRAGE images underwent pre-processing correction for gradient non-linearity[12] and intensity non-uniformity[13].

Rates of whole brain atrophy and ventricular expansion were calculated using the boundary shift integral (BSI)[14, 15]. All serial MRI were registered to baseline using 9 degrees-of-freedom registration. Change in whole brain and ventricle volume was then calculated from these registered scan-pairs using the BSI.

Rates of atrophy of specific regional structures were also assessed, including superior frontal lobe, thalamus, caudate nucleus, and midbrain. We selected these regions based on the results of previous imaging and pathological studies in PSP[2, 5, 16]. The longitudinal freesurfer software pipeline using 6 degrees-of-freedom registrations and version 4.5.0[17, 18] (<http://surfer.nmr.mgh.harvard.edu/>) was utilized to generate grey matter volumes of the superior frontal lobe, thalamus, caudate nucleus, and total intracranial volume (TIV) at each time-point. Briefly, all images underwent skull-stripping and affine and non-linear transformations into Talaraich space, followed by subcortical segmentation, identification of the grey/white boundary, automated topology correction, and surface deformation. The resulting cortical models were registered to a spherical atlas, utilizing individual cortical folding patterns to match cortical geometry across subjects[18]. The cerebral cortex was parcellated into regions based on gyral and sulcal structure[19]. Area of the midbrain was measured manually by one rater (JLW) using Analyze software (Biomedical Imaging Resource, Mayo Clinic, Rochester, MN) based on previously published criteria[5]. Measurements were performed on the mid-sagittal image with the inferior boundary defined by a line parallel to the conjunction between the genu and splenium of the corpus callosum. Midbrain areas were measured on baseline scans and on serial scans that had been registered to baseline. All measurements were performed blinded to clinical diagnosis and the temporal order of the MRI. To evaluate intra-rater reproducibility, the same rater blindly repeated ten midbrain measurements (five PSP and five controls), several weeks later. The intraclass correlation coefficient for midbrain area was 0.997 (95% confidence interval 0.994, 0.999).

All rates of atrophy were calculated as a change in volume expressed as a percentage of the baseline volume (i.e. repeat volume-baseline volume/baseline volume*100). Baseline and repeat volumes of the superior frontal lobe, thalamus, and caudate nucleus were normalized to TIV estimates to account for any scaling changes that may have occurred over time[20]. TIV correction was not necessary for the BSI or midbrain measurements since they were

performed on scans that had been registered using 9 degrees-of-freedom including scaling factors.

Statistical analysis

Baseline measurements were compared between PSP and controls using Mann Whitney U test. To analyze longitudinal changes in PSP, we compared the MRI and PSPRS measurements at 6 and 12-months follow-up, to baseline measures, using Wilcoxon Signed Rank test. Rates of atrophy (percentage change from baseline) were adjusted for scan interval to provide rates over 6 and 12-months. Rates of atrophy were then compared between PSP and controls using Mann Whitney U test. All of the above tests were assessed at a significance level of 5% using SAS v. 9.0 (Cary, NC).

Annualized and 6-month adjusted sample size requirements were estimated based on differences in the rates of atrophy or change in PSPRS over time between PSP and controls. We accounted for the rates of atrophy observed in the controls in our sample size estimates. Therefore, a reduction of 20% means a 20% reduction in the difference between the PSP rate and the control rate. We performed the calculations with nQuery Advisor 7.0 (Statistical Solutions, Saugus, MA) to have a power of 80% or 90% to detect the specified treatment effect at a 2-sided significance level of 5%. Since the degree of effect for future treatments is unclear we calculated sample size estimates for a range of different treatment effects (10–40%). Sample size estimates were only calculated for ROI's that showed significant decline over time.

RESULTS

At baseline, PSP showed significantly reduced volume of the thalamus and area of the midbrain, and increased volume of the ventricles, compared to controls (Table 1). No differences were observed across PSP and controls in volumes of the whole brain, superior frontal lobe or caudate nucleus.

The PSP patients showed significantly reduced volumes of the whole brain and thalamus, and area of the midbrain, and increased volume of the ventricles, at both 6 and 12-months compared to the baseline assessment (Table 1). The rates of atrophy calculated over 6 and 12-months for these regions were also significantly greater than controls (Table 2). The superior frontal lobe volumes only showed significant decline, and increased rates of atrophy compared to controls, over the 12-month interval (Table 1 and 2). Caudate volume did not significantly decline over time in PSP. The PSPRS total score, as well as the oculomotor and gait/midline subscores, were significantly increased in PSP at both 6 and 12 months compared to baseline (Table 1).

Sample size estimates per treatment arm for a placebo-controlled treatment trial are shown in Figure 1 and Supplemental Tables 1 and 2. Over an interval of 6-months, rate of midbrain atrophy provided the smallest sample size estimates of all the biomarkers, followed by whole brain and ventricular volume. Sample size estimates were reduced over an interval of 12-months when, once again, midbrain area provided the smallest sample size estimates, followed by rate of ventricular expansion. Over 12-months, the sample size estimates for the PSPRS total score were smaller than the sample size estimates for rate of whole brain atrophy. Sample size estimates using the PSPRS subscores were large.

DISCUSSION

We have shown excellent sample size estimates based on data collected over only 6-months in a cohort of probable PSP patients, suggesting that short interval clinical treatment trials

are a realistic possibility in PSP. Rate of midbrain atrophy appears to be the most useful outcome measure.

Increased rates of whole brain, midbrain, and thalamic atrophy, and ventricular expansion, were detected over 6-months in our cohort of probable PSP patients. No previous studies have assessed 6-month intervals in PSP, although studies in pathologically confirmed PSP have observed similar rates of whole brain atrophy and ventricular expansion over 12 months[7, 8]. Midbrain atrophy has been shown to be a diagnostic feature of PSP, with a classic “hummingbird sign” observed in the brainstem on MRI[5, 21, 22]. Midbrain atrophy has also been similarly observed over longer intervals in PSP[9]. The thalamus lies on the dentatorubrothalamic pathway which is heavily affected in PSP[16, 23]. Reduced thalamic integrity[24, 25], volume[16, 26] and functional connectivity[16] have been previously observed cross-sectionally, and we now demonstrate that the thalamus undergoes progressive atrophy over time in PSP. Reduced frontal lobe volumes have also been reported in PSP, and been shown to particularly focus in the lateral premotor cortices and supplemental motor area[2, 16]. While we did not observe increased rates of atrophy of our superior frontal lobe ROI over 6-months, we did observe significant changes over 12-months showing that it does undergo progressive atrophy. Our superior frontal lobe ROI includes premotor regions, but also part of the prefrontal cortex, and hence may not be sensitive enough to detect focal atrophic changes over 6-months. Contrary to some previous studies[27], we did not observe atrophy of the caudate in our group of PSP patients. Some non-significant volume loss was observed at baseline but no atrophy was observed over time, suggesting that caudate nucleus volume would not be a good biomarker of disease progression in PSP.

We observed significant increases in the PSPRS total score over both 6 and 12-month intervals in PSP. An average increase of 9 points was observed over the first 6-months and 18 points over 12-months, showing an even rate of progression over 12 months. These rates of progression are higher than the rate of between 10–12 points per year that has been previously reported[11], although the previous study used different clinical diagnostic criteria and calculated rates using multiple visits including assessments before patients fulfilled criteria for PSP. Significant progression was observed in the oculomotor and gait/midline subscores of the PSPRS, demonstrating worsening of these symptoms that are typical for PSP. No change over time was observed in the mentation subscore showing that this is not a good measure of disease progression in PSP.

Detecting change over intervals as short as 6 months could have particularly important consequences for clinical treatment trials in PSP, allowing for shorter patient follow-up. When assessing the imaging biomarkers, our power calculations showed that the number of subjects required to detect treatment effects in a 6-month trial was the smallest when using rate of midbrain atrophy as the outcome measure, closely followed by rate of whole brain atrophy and ventricular expansion. A much larger number of subjects would be required with rate of thalamic atrophy. Rate of midbrain atrophy most likely provides the best sample size estimates since this region is heavily affected in PSP; we observed a 44% reduction in midbrain area at baseline in PSP compared to controls. Whole brain and ventricular volumes may be less sensitive because PSP is not associated with widespread cortical atrophy. However, a potential drawback of the midbrain measurement is that it involves increased rater time. Rates of whole brain and ventricular change were calculated using the BSI which is a completely automated measure and so they may prove to be more valuable for clinical trials that involve large numbers of subjects. Sample size estimates calculated over 6 months with the PSPRS total score were larger than those calculated using rate of midbrain, whole brain and ventricular change, and were only marginally better than those calculated using thalamus. Sample size estimates using the PSPRS subscores were also poor. Imaging

biomarkers therefore appear to be more sensitive than the PSPRS to the subtle changes that occur over this short interval.

Sample size estimates were further reduced when the inter-scan interval was increased to 12-months and were once again the smallest for rate of midbrain atrophy. In fact, for a treatment trial with 80% power, only 63 subjects would be needed per treatment arm to detect treatment effects of 30% using rate of midbrain atrophy. This reduction of sample size at 12-months is likely because of reduced measurement variability over longer intervals. Low sample size estimates were also observed using rate of whole brain atrophy and ventricular expansion, with slightly lower numbers required using ventricular volume. Ventricular volume has previously been shown to provide superior sample size estimates to whole brain volume in a study assessing Alzheimer's disease [28], likely since ventricular measurements can be performed with high accuracy and tend to be less sensitive to intensity inhomogeneity. Interestingly, over a 12 month interval, sample size estimates from the PSPRS total score were in fact better than those calculated using rate of whole brain atrophy, and were not much larger than those calculated using rate of midbrain atrophy. This shows that although the PSPRS is not a good measure of progression over 6 months, it could be a valuable outcome measure for 12 month trials, and may be particularly useful for those trials that do not include MRI assessments.

Our finding that rates of midbrain atrophy provide better sample size estimates than rates of whole brain atrophy or ventricular expansion concurs with a previous study that assessed sample size estimates over 12-month intervals in PSP[29]. However, our sample size estimates were much smaller than those previously reported (for example, sample size estimates to detect a 30% treatment effect at 90% power were 84 for rate of midbrain atrophy and 117 for rate of whole brain atrophy in our study, compared to 147 and 399 in the previous study). This difference may be due to the fact that we only included patients with a clinical diagnosis of probable PSP according to the Litvan et al. 1996 criteria[1], whereas the previous study also included patients with a clinical diagnosis of possible PSP[1] that have a lower probability of having PSP pathology[30]. Indeed, our rates of midbrain atrophy were over four times larger than those previously reported, and showed less variability[29]. The method used for measuring midbrain did differ across studies, although the BSI were calculated using almost the identical procedure, and therefore differences are unlikely to be due to methodological technique.

The strength of our study is the fact that our PSP patients all fulfilled probable clinical criteria for PSP[1] and were followed at both 6 and 12-month intervals allowing a comparison of sample size. A limitation of the study was that we did not have an ROI that specifically assessed the premotor cortex, which is focally involved in PSP[2, 16]. Sample size estimates may be improved using a more focal premotor ROI although it is unlikely that it would be superior to midbrain estimates. Regardless, the results of this study are important for future clinical treatment trials in PSP, particularly early phase studies assessing treatment efficacy. Our data demonstrate that 6-month trials are a feasible possibility in PSP in terms of sample size, although sample size estimates can be reduced further by increasing the follow-up time to 12-months. Future studies will be needed to determine whether treatment effects can be detected over 6-month intervals in PSP.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Dana Foundation; National Institutes of Health (grant numbers R01-DC010367, R01-AG037491, R21-AG38736, and R01-AG11378) and the Alexander Family Alzheimer's Disease Research Professorship of the Mayo Foundation. We would like to acknowledge Drs. Scott Eggers, J. Eric Ahlskog and Joseph Matsumoto for referring patients for this study.

References

1. Litvan I, Agid Y, Calne D, Campbell G, Dubois B, Duvoisin RC, et al. Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop. *Neurology*. 1996 Jul; 47(1):1–9. [PubMed: 8710059]
2. Josephs KA, Whitwell JL, Dickson DW, Boeve BF, Knopman DS, Petersen RC, et al. Voxel-based morphometry in autopsy proven PSP and CBD. *Neurobiol Aging*. 2008 Feb; 29(2):280–9. [PubMed: 17097770]
3. Agosta F, Kostic VS, Galantucci S, Mesaros S, Svetel M, Pagani E, et al. The in vivo distribution of brain tissue loss in Richardson's syndrome and PSP-parkinsonism: a VBM-DARTEL study. *The European journal of neuroscience*. 2010 Aug; 32(4):640–7. [PubMed: 20597976]
4. Boxer AL, Geschwind MD, Belfor N, Gorno-Tempini ML, Schauer GF, Miller BL, et al. Patterns of brain atrophy that differentiate corticobasal degeneration syndrome from progressive supranuclear palsy. *Archives of neurology*. 2006 Jan; 63(1):81–6. [PubMed: 16401739]
5. Cosottini M, Ceravolo R, Faggioni L, Lazzarotti G, Michelassi MC, Bonuccelli U, et al. Assessment of midbrain atrophy in patients with progressive supranuclear palsy with routine magnetic resonance imaging. *Acta neurologica Scandinavica*. 2007 Jul; 116(1):37–42. [PubMed: 17587253]
6. Josephs KA, Hodges JR, Snowden J, Mackenzie IR, Neumann M, Mann D, et al. Neuropathological background of phenotypical variability in frontotemporal dementia. *Acta Neuropathol*. 2011; 122(2):137–53. [PubMed: 21614463]
7. Josephs KA, Whitwell JL, Boeve BF, Shiung MM, Gunter JL, Parisi JE, et al. Rates of cerebral atrophy in autopsy-confirmed progressive supranuclear palsy. *Annals of neurology*. 2006 Jan; 59(1):200–3. [PubMed: 16278856]
8. Whitwell JL, Jack CR Jr, Parisi JE, Knopman DS, Boeve BF, Petersen RC, et al. Rates of cerebral atrophy in different degenerative pathologies. *Brain*. 2007 Apr; 130(Pt 4):1148–58. [PubMed: 17347250]
9. 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 Apr; 129(Pt 4):1040–9. [PubMed: 16455792]
10. Josephs KA, Ahlskog JE, Parisi JE, Boeve BF, Crum BA, Giannini C, et al. Rapidly progressive neurodegenerative dementias. *Archives of neurology*. 2009 Feb; 66(2):201–7. [PubMed: 19204156]
11. Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. *Brain*. 2007 Jun; 130(Pt 6):1552–65. [PubMed: 17405767]
12. Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. *Neuroimage*. 2006 Apr 1; 30(2):436–43. [PubMed: 16300968]
13. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE transactions on medical imaging*. 1998 Feb; 17(1):87–97. [PubMed: 9617910]
14. Freeborough PA, Fox NC. The boundary shift integral: an accurate and robust measure of cerebral volume changes from registered repeat MRI. *IEEE transactions on medical imaging*. 1997 Oct; 16(5):623–9. [PubMed: 9368118]
15. Gunter JL, Shiung MM, Manduca A, Jack CR Jr. Methodological considerations for measuring rates of brain atrophy. *J Magn Reson Imaging*. 2003 Jul; 18(1):16–24. [PubMed: 12815635]
16. Whitwell JL, Avula R, Master A, Vemuri P, Senjem ML, Jones DT, et al. Disrupted thalamocortical connectivity in PSP: a resting state fMRI, DTI, and VBM study. *Parkinsonism & related disorders*. 2011 Jun 10. [Epub].

17. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002 Jan 31; 33(3):341–55. [PubMed: 11832223]
18. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*. 1999 Feb; 9(2):195–207. [PubMed: 9931269]
19. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*. 2006 Jul 1; 31(3):968–80. [PubMed: 16530430]
20. Whitwell JL, Crum WR, Watt HC, Fox NC. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. *Ajnr*. 2001 Sep; 22(8):1483–9. [PubMed: 11559495]
21. Oba H, Yagishita A, Terada H, Barkovich AJ, Kutomi K, Yamauchi T, et al. New and reliable MRI diagnosis for progressive supranuclear palsy. *Neurology*. 2005 Jun 28; 64(12):2050–5. [PubMed: 15985570]
22. Quattrone A, Nicoletti G, Messina D, Fera F, Condino F, Pugliese P, et al. MR imaging index for differentiation of progressive supranuclear palsy from Parkinson disease and the Parkinson variant of multiple system atrophy. *Radiology*. 2008 Jan; 246(1):214–21. [PubMed: 17991785]
23. Dickson DW. Neuropathologic differentiation of progressive supranuclear palsy and corticobasal degeneration. *Journal of neurology*. 1999 Sep; 246(Suppl 2):II6–15. [PubMed: 10525997]
24. Padovani A, Borroni B, Brambati SM, Agosti C, Broli M, Alonso R, et al. Diffusion tensor imaging and voxel based morphometry study in early progressive supranuclear palsy. *J Neurol Neurosurg Psychiatry*. 2006 Apr; 77(4):457–63. [PubMed: 16306152]
25. Erbetta A, Mandelli ML, Savoirdo M, Grisoli M, Bizzi A, Soliveri P, et al. Diffusion tensor imaging shows different topographic involvement of the thalamus in progressive supranuclear palsy and corticobasal degeneration. *Ajnr*. 2009 Sep; 30(8):1482–7. [PubMed: 19589886]
26. Messina D, Cerasa A, Condino F, Arabia G, Novellino F, Nicoletti G, et al. Patterns of brain atrophy in Parkinson's disease, progressive supranuclear palsy and multiple system atrophy. *Parkinsonism & related disorders*. 2011 Mar; 17(3):172–6. [PubMed: 21236720]
27. Schulz JB, Skalej M, Wedekind D, Luft AR, Abele M, Voigt K, et al. Magnetic resonance imaging-based volumetry differentiates idiopathic Parkinson's syndrome from multiple system atrophy and progressive supranuclear palsy. *Annals of neurology*. 1999 Jan; 45(1):65–74. [PubMed: 9894879]
28. Schott JM, Price SL, Frost C, Whitwell JL, Rossor MN, Fox NC. Measuring atrophy in Alzheimer disease: a serial MRI study over 6 and 12 months. *Neurology*. 2005 Jul 12; 65(1):119–24. [PubMed: 16009896]
29. Paviour DC, Price SL, Lees AJ, Fox NC. MRI derived brain atrophy in PSP and MSA-P. Determining sample size to detect treatment effects. *Journal of neurology*. 2007 Apr; 254(4):478–81. [PubMed: 17401522]
30. Lopez OL, Litvan I, Catt KE, Stowe R, Klunk W, Kaufer DI, et al. Accuracy of four clinical diagnostic criteria for the diagnosis of neurodegenerative dementias. *Neurology*. 1999 Oct 12; 53(6):1292–9. [PubMed: 10522887]

Figure 1.

Estimated sample sizes required per treatment arm in a clinical trial to detect a reduction in each outcome variable in PSP. Plots are shown for a 6-month scan interval at 80% (A) and 90% (B) power, and for a 12-month scan interval at 80% (C) and 90% (D) power. Estimates are given for potential 20%, 30% and 40% reductions in each outcome variable. Y axis scales have been limited to 1000 subjects for 6-month plots and 500 subjects for 12-month plots. * Exact sample size estimates for outcome variables that exceed these limits can be seen in Supplemental Tables 1 and 2.

Table 1

Region-of-interest data for controls and PSP patients at each time point

	Controls			PSP			P values		
	Time-point		P value	Time-point		12 months	Base v 6 months	Base v 12 months	6 v 12 months
	Baseline	12 months		Baseline	6 months				
Whole brain volume	88.30 (2.65)	87.97 (2.58)	0.0386	87.53 (2.45)	86.84 (2.54)	86.31 (2.27)	0.0004	<0.0001	0.0021
Ventricular volume	2.33 (1.11)	2.40 (1.13)	0.0063	3.21 (1.20)*	3.38 (1.22)	3.58 (1.24)	0.0003	<0.0001	0.0002
Superior frontal lobe volume	2.43 (0.11)	2.43 (0.13)	0.7436	2.37 (0.21)	2.35 (0.20)	2.31 (0.20)	0.3484	0.0027	0.0507
Thalamus volume	0.80 (0.03)	0.79 (0.04)	0.5619	0.75 (0.05)*	0.74 (0.05)	0.73 (0.05)	0.0131	0.0003	0.0335
Caudate volume	0.49 (0.07)	0.49 (0.07)	0.9799	0.47 (0.08)	0.47 (0.08)	0.47 (0.09)	0.1591	0.9799	0.8603
Midbrain area	0.01011 (0.00173)	0.01018 (0.00199)	0.9399	0.00571 (0.00120)*	0.00538 (0.00120)	0.00507 (0.00096)	0.0017	<0.0001	0.0027
PSPRS total score	NA	NA	NA	30.94 (11.72)	39.56 (10.40)	48.44 (10.44)	0.0003	<0.0001	<0.0001
PSPRS-Mentation	NA	NA	NA	2.81 (2.51)	2.50 (2.03)	3.13 (2.60)	0.7148	0.3208	0.2986
PSPRS-Oculomotor	NA	NA	NA	7.88 (2.92)	9.63 (1.93)	10.81 (2.64)	0.0177	0.0004	0.0381
PSPRS-Gait/Midline	NA	NA	NA	8.88 (5.07)	11.75 (3.64)	14.81 (3.02)	0.0034	<0.0001	<0.0001

All volumes are shown as a percentage of TIV; PSPRS = PSP rating scale; NA = Not applicable

* Significant differences observed between baseline measurements for PSP compared to controls using Wilcoxon Rank Sum test

Table 2
Percentage tissue lost across ROIs, and PSPRS increase, for controls and PSP patients

	Controls	PSP		P values	
	Adjusted 12 month	Adjusted 6 month	Adjusted 12 month	Controls v PSP 6 month adjusted	Controls v PSP 12 month adjusted
Whole brain volume	-0.30 (0.54)	-0.69 (0.61)	-1.34 (0.73)	0.0001	0.0001
Ventricular volume	2.54 (2.73)	3.96 (3.03)	10.17 (4.83)	<0.0001	<0.0001
Superior frontal lobe volume	-0.23 (2.36)	-0.57 (3.56)	-2.11 (2.38)	0.3860	0.0418
Thalamus volume	-0.37 (1.75)	-1.42 (1.80)	-2.75 (2.20)	0.0288	0.0037
Caudate volume	-0.02 (1.69)	-0.51 (2.99)	-0.38 (4.11)	0.0830	0.9399
Midbrain area	0.35 (3.52)	-5.98 (6.51)	-10.34 (6.35)	0.0014	<0.0001
PSPRS total score	NA	8.63 (7.82)	17.50 (6.91)	NA	NA
PSPRS-Mentation	NA	-0.31 (2.02)	0.31 (2.21)	NA	NA
PSPRS-Occulomotor	NA	1.75 (2.49)	2.94 (2.67)	NA	NA
PSPRS-Gait/Midline	NA	2.88 (3.40)	5.94 (3.30)	NA	NA

Change is calculated as (repeat-baseline)*100/baseline. Data has been adjusted to represent % change over either 6 or 12 months. P-values are calculated using the adjusted data and Wilcoxon Rank Sum test. PSPRS = PSP rating scale; NA = Not applicable