UCSF
UC San Francisco Previously Published Works

Title
Clinical correlates of longitudinal brain atrophy in progressive supranuclear palsy

Permalink
https://escholarship.org/uc/item/1v89z6zw

Authors
Tsai, Richard M
Lobach, Iryna
Bang, Jee
et al.

Publication Date
2016-07-01

DOI
10.1016/j.parkreldis.2016.04.006

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Parkinsonism and Related Disorders xxx (2016) 1–8

Contents lists available at ScienceDirect

Parkinsonism and Related Disorders

journal homepage: www.elsevier.com/locate/parkreldis

Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials

Jee Bang [a, 1], Iryna V. Lobach [b, 1], Anthony E. Lang [c], Murray Grossman [d], David S. Knopman [e], Bruce L. Miller [a], Lon S. Schneider [f, g], Rachelle S. Doody [h], Andrew Lees [i], Michael Gold [j], Bruce H. Morimoto [k], Adam L. Boxer [a, *], on behalf of AL-108-231 Investigators

[a] Memory and Aging Center, Department of Neurology, University of California San Francisco, 675 Nelson Rising Lane, Suite 190, San Francisco, CA, 94158, USA
[b] Department of Epidemiology and Biostatistics, Division of Biostatistics, University of California San Francisco, 550 16th Street, San Francisco, CA, 94158, USA
[c] Department of Neurology, University of Toronto, 399 Bathurst St, McLaughlin 7-418, Toronto, ON, M5T 2S8, Canada
[d] Department of Neurology, University of Pennsylvania, 3400 Spruce St, Philadelphia, PA, 19104, USA
[e] Department of Neurology, Mayo Clinic, 200 1st St SW, Rochester, MN, 55905, USA
[f] Department of Psychiatry and the Behavioral Sciences, Keck School of Medicine of the University of Southern California, Los Angeles, CA, USA
[g] Department of Neurology, Keck School of Medicine of the University of Southern California, Los Angeles, CA, USA
[h] Department of Neurology, Baylor College of Medicine, 6550 Fannin St # 1801, Houston, TX, 77030, USA
[i] Institute of Neurology, University College of London, 1 Wakefield Street, London, WC1N 1PJ, UK
[j] UCB BioSciences, Research Triangle, NC, USA
[k] Celerion, 621 Rose St, Lincoln, NE, 68502, USA

ARTICLE INFO

Article history:
Received 11 January 2016
Received in revised form
22 March 2016
Accepted 16 April 2016

Keywords:
Progressive supranuclear palsy
Clinical trial methodology

ABSTRACT

Introduction: Clinical and MRI measurements can track disease progression in PSP, but many have not been extensively evaluated in multicenter clinical trials. We identified optimal measures to capture clinical decline and predict disease progression in multicenter PSP trials.

Methods: Longitudinal clinical rating scales, neuropsychological test scores, and volumetric MRI data from an international, phase 2/3 clinical trial of davunetide for PSP (intent to treat population, n = 303) were used to identify measurements with largest effect size, strongest correlation with clinical change, and best ability to predict dropout or clinical decline over one year as measured by PSP Rating Scale (PSPRS).

Results: Baseline cognition as measured by Repeatable Battery for Assessing Neuropsychological Status (RBANS) was associated with attrition, but had only a small effect. PSPRS and Clinical Global Impression (CGI) had the largest effect size for measuring change. Annual change in CGI, RBANS, color trails, and MRI midbrain and ventricular volumes were most strongly correlated with annual PSPRS and had the largest effect sizes for detecting annual change. At baseline, shorter disease duration, more severe depression, and lower performance on RBANS and executive function tests were associated with faster worsening of the PSPRS in completers. With dropouts included, SEADL, RBANS, and executive function tests had significant effect on PSPRS trajectory of change.

Conclusion: Baseline cognitive status and mood influence the rate of disease progression in PSP. Multiple clinical, neuropsychological, and volumetric MRI measurements are sensitive to change over one year in PSP and appropriate for use in multicenter clinical trials.

© 2016 Elsevier Ltd. All rights reserved.

* Corresponding author.
E-mail addresses: jee.bang@ucsf.edu (J. Bang), iryna.lobach@ucsf.edu (I.V. Lobach), lang@uhnresearch.ca (A.E. Lang), mgrossma@mail.med.upenn.edu (M. Grossman), knopman@mayo.edu (D.S. Knopman), bruce.miller@ucsf.edu (B.L. Miller), lschneid@usc.edu (L.S. Schneider), rdoody@bcm.edu (R.S. Doody), alees@ion.ucl.ac.uk (A. Lees), michael.gold.md@gmail.com (M. Gold), bhmorimoto@gmail.com (B.H. Morimoto), adam.boxer@ucsf.edu (A.L. Boxer).
[1] These authors contributed equally to the manuscript.

http://dx.doi.org/10.1016/j.parkreldis.2016.04.014
1353-8020/© 2016 Elsevier Ltd. All rights reserved.

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

1. Introduction

Progressive supranuclear palsy (PSP) is a fatal neurodegenerative disease characterized by the aggregation of predominantly 4 microtubule binding domain repeat (4R) tau in neurons and glia [1]. There are several clinical presentations of PSP [2,3] and Richardson's syndrome is the most recognizable and rapidly progressive phenotype, characterized by early and severe gait instability with falls, slowed eye movements progressing to supranuclear ophthalmoplegia, axial rigidity, and variable neuropsychiatric symptoms. There are no effective therapies for PSP; however, a variety of new potential treatments targeting tau are entering clinical trials [4].

The feasibility of conducting pivotal clinical trials in PSP was recently demonstrated in three large, international studies [5–7]. A variety of clinical rating scales (such as PSP Rating Scale); PSPRS [8] and volumetric MRI measurements have been developed and validated for use in PSP based on small, single center studies and then applied to large, international clinical trials with little evidence to support their utility in multicenter settings. We examined data from the 48 center, randomized, placebo controlled phase 2/3 clinical trial of davunetide for PSP (AL-108-231) [6] to identify the best baseline clinical and biomarker outcome measures that: 1) capture clinical decline and 2) predict attrition or disease progression over one year.

2. Methods

2.1. Source of data

The data for this study were taken from the previously reported AL-108-231 (clinicaltrials.gov, NCT01110720) international, randomized, double-blind, placebo-controlled, phase 2/3 trial of davunetide for PSP [6]. The study enrolled 313 patients with PSP (Richardson's syndrome) at 48 centers in Australia, Canada, France, Germany, the United Kingdom and the United States. The intent to treat population (n = 303) of individuals who were randomized to davunetide or placebo and had at least one post-baseline assessment of both primary and secondary outcomes was used for analyses of baseline variables that predicted dropout.

2.2. Inclusion criteria

To be included in the AL-108-231 study, participants had to meet modified criteria for probable or possible PSP based on the Neuroprotection and Natural History in Parkinson Plus Syndromes (NNIPPS) study [5]. Individuals had to be between 41 and 85 years of age at disease onset with at least a 12-month history of early and prominent postural instability or falls, supranuclear ophthalmoplegia or decreased downward saccade velocity, and prominent axial rigidity. Participants were required to be able to either ambulate independently or take at least five steps with minimal assistance. Individuals could participate only if they had PSP symptoms for less than 5 years, or if for more than 5 years with a PSPRS of 40 or greater at screening. Detailed inclusion and exclusion criteria are described in the primary study manuscript [6].

2.3. Clinical data

The primary endpoints were the change in PSPRS and Schwab and England activities of daily living scale (SEADL) [9] over one year. The PSPRS consists of six categories including daily activities, behavior, bulbar, oculomotor, limb motor, and gait/midline. Scores range from 0 to 100, with higher scores indicating more severe disease. SEADL is a measure of overall disability based on interviews with the patient and the informant, and is scored on an 11-point ordinal scale (10% intervals starting with 0 indicating vegetative functions, up to 100% indicating complete independence).

Secondary outcome measures included the Clinical Global Impression of Change (CGIC) [10] and brain ventricular volume as measured on MRI scans as described below [6]. In addition, exploratory outcomes were obtained including: the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS; test domains include memory, visuospatial, language, and attention) [11], three additional assessments of executive function (color trails, phonemic fluency, and letter-number sequencing), the Geriatric Depression Scale (GDS) [12], the Clinical Global Impression of Disease Severity (CGIds) [13], and additional volumetric MRI scan measurements of the whole brain, midbrain, and superior cerebellar peduncle (SCP). A total of 217 patients completed the neuropsychological testing.

2.4. MRI data (n = 214)

To be included in the clinical trial all participants had to complete a baseline volumetric T1-weighted MRI scan on a 1.5 T or 3 T scanner. Whole brain and ventricular volumes were generated using the boundary shift integral technique, and midbrain and SCP volumes were generated using label propagation in statistical parametric mapping 5 (SPM5) at the Mayo Clinic Aging and Dementia Imaging Research Laboratory as previously described [6]. Brain volumes were adjusted for total intracranial volume (TIV) to control for head size differences where appropriate. Pons volume was not obtained. All scanners were calibrated using a standard phantom and MRI analyses were conducted blinded to treatment assignment. Five subjects were excluded from the original dataset (n = 219), because they were deemed to be influential outliers (well below the 25th or far above the 75th percentile of annual change), likely due to artifacts introduced during the initial MRI analysis.

2.5. Statistical analysis

We combined data from the placebo and davunetide groups since extensive analyses of the davunetide trial dataset revealed no differences between groups at baseline. The absolute value of Cohen's d between the treatment arms for baseline demographics, primary and secondary outcomes ranged [0.01, 0.14]; for MRI measures [0.01, 0.07], and in change in measures over time for primary and secondary outcomes ranged [0.01, 0.24]; and for change in MRI measures [0.01, 0.15]. Baseline values and 52-week change from baseline values in clinical ratings, neuropsychological measures, and MRI volumes were presented using estimates of central tendency (mean, proportion) and variance. Effect of baseline characteristics on drop out was examined using logistic regression models. Concordance between the observed 52-week change in PSPRS and the corresponding change in other measures was measured using Pearson R^2 and Spearman correlation coefficients where appropriate. All estimates include 95% confidence intervals. The relationship of the baseline evaluations to the 52-week change in PSPRS was explored with univariate and multivariate linear regression models. These models were performed with and without adjustments for potential confounders: baseline PSPRS, age, gender, disease duration, treatment group assignment (davunetide or placebo), tau haplotype, CoQ10 use, and MMSE.

We further examined the effect of baseline evaluations on trajectory of PSPRS across all patients using linear mixed effects models. These models accommodate repeated measures of PSPRS and allow the baseline evaluations to have impact on the overall trajectory both in slope (speed of PSPRS change) and intercept. The

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

random coefficients model that we employed includes two random variables allowing shift in the overall trajectory and modification of the speed of PSPRS change due to patient-specific characteristics.

To investigate the ability of the baseline evaluations to define subpopulations with either a faster rate or smaller standard deviation of change in PSPRS, a minimum p-value method and sample size/power analyses were performed. The search for a cut-off was performed on an equally spaced grid of values ranging from minimum to maximum of a baseline measure. The goal of this search was to find a cut-off that separated the population into two subpopulations with significantly different speeds of progression as measured by the two-sided, two-sample Student's *t*-test of the 52-week change in PSPRS. If multiple cut-offs resulted in significant separation ($p < 0.05$), final cut-offs were chosen to be the ones that defined clinically relevant subpopulations with maximum effect size (PSPRS/SD[PSPRS]). Longitudinal behavior of PSPRS in the inferred subpopulations is presented based on average changes modeled at 13, 26, 39, and 52 week follow-up visits, accompanied by standard error values. Sample size analyses were performed to estimate minimum sample sizes required to detect 10%, 25%, 37.5%, and 50% change in PSPRS progression attributable to the treatment effect with 90% power using two-sample Student's *t*-test at the two-sided 0.05 significance level, not accounting for dropout or accounting for the observed 23% dropout rate from the AL-108-231 study. The proportion of inter-patient variability in change of PSPRS explained by multivariate regression models was broken down into portions corresponding to each variable using analyses of relative importance [14]. False discovery rate (FDR) adjustment was applied to correct for multiple testing [15].

2.6. Standard protocol approvals, registrations, and patient consents

Ethics approval was obtained at each site from the local ethics committee and all participants gave written informed consent at the recruitment visit as per local regulations.

3. Results

3.1. Baseline characteristics of AL-108-231 trial completers

The baseline characteristics of the davunetide trial participants are shown in Table 1. Baseline characteristics of the participants with complete 52-week PSPRS data ($n = 241$) did not substantially differ from the overall population ($n = 303$) used for the primary, intent to treat analyses [6]. The baseline characteristics of the participants who dropped out ($n = 62$) were similar to those of completers. However, participants with worse cognitive function at baseline (RBANS scores) had slightly higher risk of dropping out (RR = 1.05 [95% CI: 1.02, 1.08]). There were no other characteristics that differed between individuals who dropped out or completed the trial.

3.2. Correlation of changes in outcome measures with change in PSPRS

Of the clinical scales, the largest effect sizes (Cohen's d) for change over 52 weeks were observed with the PSPRS and CGIC in the trial completers (Table 2). As expected, the two primary outcome measures, the PSPRS and SEADL, showed concordant declines over one year. Other clinical (CGIds, CGIC, GDS) and neuropsychological (RBANS Total Raw and Total Scaled, Letter Number Sequence) outcomes changed over one year and these declines also correlated with changes in PSPRS. Midbrain volume (absolute change and percent change), ventricular volume (percent change), and whole brain volume (percent change) had the largest effect sizes for 52 week change and were correlated with changes in PSPRS.

3.3. Individual baseline predictors of clinical progression

In the trial completers, the baseline values for a variety of clinical outcome measures, including disease duration, GDS, RBANS, and color trails 1 and 2 were strongly related to 52-week change in PSPRS (Table 3) with adjustment for baseline PSPRS, and other potential confounding variables including age, gender, disease duration, treatment group assignment (davunetide or placebo), tau haplotype, CoQ10 use, and MMSE. Using linear mixed effect models, we further examined the effect of baseline measurements on trajectory of PSPRS change of all participants including completers and dropouts with adjustments for patient-specific characteristics, including potential confounders (Table 4). The following baseline evaluations had significant effect on both the intercept and slope (speed of change) of PSPRS trajectory: color trails 2 ($z = 3.2$), color trails 1 ($z = 3.09$), phonemic fluency ($z = -2.9$), RBANS [$z = -2.32$(scaled) and $z = -2.28$(raw)] and SEADL ($z = -2.06$).

3.4. Multivariate models predicting clinical progression

The ability of multiple baseline characteristics to explain inter-patient variability in change of PSPRS was examined using multivariate regression models. We found that a model that combined baseline demographic, clinical, and neuropsychological measures (PSPRS, color trails 2, GDS, total raw RBANS score, phonemic fluency), and potential confounders (age, disease duration, treatment group, co-enzyme Q10 use) explained 16.4% of variance in 52-week PSPRS change (Model 1; Supplementary Table 1). In this model ($N = 226$), disease duration [$\beta = -5.77$ (-10.1, -1.4)] and GDS [$\beta = 0.24$ (0.06, 0.41)], were significant contributors. Adding volumetric MRI measurements to this model (Model 2; $N = 213$) did not improve the ability to explain variance in 52-week PSPRS change.

3.5. Utility of baseline values in determining sample size of hypothetical clinical trials

We determined the estimated number of patients per arm in a two-arm parallel study required to demonstrate an effect of a hypothetical treatment for PSP in slowing the rate of change in PSPRS scores over 52 weeks, assuming all participants completed the trial (Supplementary Table 2). The planned AL-108-231 sample size ($n = 150$ per arm) was based on the number required to detect a treatment effect of a 37.5% difference in rate of decline in PSPRS over one year based on the published rate of PSPRS change (11 ± 11 points) at a single center [8] with 90% power at $\alpha = 0.05$. However, using the observed rate of PSPRS change in the patients who completed the study (11 ± 9 points), only 106 patients per arm would be required to detect this treatment effect.

Using a minimum p-value approach, we identified cut points in baseline values that could define sub-populations with significantly different rates of change in PSPRS over one year based on the observed data. For example, baseline phonemic fluency values defined subpopulations with faster rates of PSPRS change than the overall population, leading to a reduction in sample size required to detect a treatment effect (Supplementary Table 2). Other cut points defined subpopulations with slower rates of change than the overall population, such as individuals with disease duration at baseline of greater than 5 years.

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

Table 1
Baseline characteristics of completers and dropouts.

Value		Completers	Dropouts	Completers vs dropouts
		Mean (95% confidence interval) or count (%)		Relative risk* (95% confidence interval)
Demographics		*N = 241*	*N = 62*	
Age	years	67.4 (66.5, 68.1)	68.6 (66.9, 70.2)	0.99 (0.94, 1.02)
Sex	female	125 (51.9%)	27 (43.5%)	0.85 (0.47, 1.5)
Weight	kg	78.4 (76.2, 80.2)	74.1 (69.9, 78.4)	1.03 (1.01, 1.06)
Race	White	214 (88.8%)	52 (83.9%)	1.1 (1.0, 1.2)
MMSE		26.5 (26.1, 26.9)	25.5 (24.5, 26.4)	1.1 (1.0, 1.2)
Treatment	Davunetide	118 (49%)	30 (48.4%)	1.1 (0.60, 1.9)
Tau haplotype	H1/H1	181 (95.8%)	47 (75.8%)	0.67 (0.34, 1.3)
n (% genotyped)	H1/H2	8 (4.2%)	5 (8.1%)	
	H2/H2	0 (0%)	0 (0%)	
	Missing	52 (21.6%)	10 (16.1%)	
Disease duration	%>5 years	19 (8.4%)	8 (12.9%)	0.63 (0.25, 1.7)
Concomitant medication used during study		*N = 241*	*N = 62*	
CoQ10 use		49 (20.3%)	11 (17.7%)	1.1 (0.52, 2.4)
Levodopa use		111 (46.1%)	26 (41.9%)	1.2 (0.64, 2.1)
Primary outcomes		*N = 241*	*N = 62*	
PSPRS		39.1 (37.7, 40.5)	41.2 (38.7, 43.7)	0.99 (0.96, 1.02)
SEADL		0.54 (0.51, 0.57)	0.46 (0.41, 0.51)	3.48 (0.52, 24.0)
Secondary/exploratory outcomes		*N = 241*	*N = 62*	
GDS		12.7 (11.8, 13.6)	12.6 (11.1, 12.6)	1.01 (0.97, 1.06)
CGIds		3.9 (3.8, 4.0)	4.2 (3.9, 4.4)	0.78 (0.53, 1.2)
RBANS	Total raw	144.9 (140.7, 149.1)	128 (119.7, 136.3)	1.02# (1.00, 1.03)
	Total scaled	74.4 (72.8, 76.0)	67.6 (64.7, 70.4)	1.05# (1.02, 1.08)
Phonemic Fluency	Words/min	11.5 (10.6, 12.4)	9.6 (8.0, 11.1)	1.03 (0.98, 1.08)
Letter number seq.	Score	7.1 (6.7, 7.5)	6.1 (5.4, 6.8)	1.06 (0.93, 1.2)
Color Trails 1	Seconds	160.5 (151.6, 169.4)	189.2 (173, 205.4)	0.99 (0.99, 1.0)
Color Trails 2	Seconds	235 (226, 244)	262.2 (246, 278.4)	0.99 (0.99, 1.0)
MR Imaging		*N = 223*	*N = 58*	
Ventricular volume/TIV	$\times 10^{-4}$	333.1 (315.6, 357.1)	330.8 (295.8, 365.7)	1.0 (0.99, 1.0)
Whole brain volume/TIV	$\times 10^{-4}$	9008 (8949, 9067)	8827 (8716, 8938)	1.0 (0.99, 1.0)
Midbrain volume/TIV	$\times 10^{-4}$	46.9 (46.2, 47.7)	51 (48, 53)	1.0 (0.99, 1.0)
SCP volume/TIV	$\times 10^{-4}$	2.6 (2.5, 2.7)	2.5 (2.3, 2.8)	1.0 (0.99, 1.0)

MMSE = Mini Mental Status Exam; PSPRS = Progressive Supranuclear Palsy Rating Scale; SEADL = Schwab and England Activities of Daily Living Scale; GDS = Geriatric Depression Scale; CGIds = Clinical Global Impression of Disease Severity; RBANS = Repeatable Battery for the Assessment of Neuropsychological Disease Status; SCP = superior cerebellar peduncle; TIV = total intracranial volume.
*Adjusted for PSPRS, age, gender, disease duration, treatment group assignment (davunetide or placebo), tau haplotype, CoQ10 use, and MMSE.
#$p < 0.05$ (FDR corrected).

4. Discussion

We found that PSPRS, CGIC, CGIds and RBANS were the best clinical measurements, and midbrain and ventricular volume were the best MRI measurements for capturing longitudinal change in PSP in a large, international clinical trial. These measurements had comparable effect sizes for measuring change; however, fewer high quality MRI data were available for analysis after one year which might limit the utility of this measurement. Baseline cognitive status had a small effect on predicting patient attrition. In the trial completers, disease duration, baseline measures of depression (GDS), RBANS, and color trails scores were significantly associated with annual changes in PSPRS scores. Patients with longer disease duration at baseline had a slower rate of progression. This might possibly be because they initially had variant forms of PSP such as pure akinesia with gait freezing (PAGF) or PSP parkinsonism (PSP-P) preceding their evolution into Richardson's syndrome at study entry. PAGF and PSP-P are known to have slower rates of progression than Richardson's syndrome [22]. In the overall population, SEADL, RBANS, and executive function tests were the strongest predictors of PSPRS change. Multivariate models that included baseline clinical and MR imaging variables were no better at explaining variance in annual PSPRS change than individual variables. Together, these results demonstrate that a number of clinical and imaging biomarkers are sensitive to change in a typical, multicenter PSP clinical trial population. Volumetric MRI measurements are likely to be informative, but did not provide substantially greater power to detect change than clinical rating scales.

4.1. Measuring longitudinal change in PSP clinical trials

After the NNIPPS study [5], the AL-108-231 study was the largest multicenter clinical trial that has been completed in PSP. Although there was little experience with use of the PSPRS in a multicenter clinical trial setting prior to this study, the current analyses show that this measure and the CGIC (a secondary outcome measure for the trial) were the best clinical scales for measuring disease progression, based on Cohen's d, over one year. Neuropsychological testing and MRI changes also showed modest correlations with changes in PSPRS. Based on effect size of change over one year and strength of correlation with change in PSPRS, midbrain volume seemed to be the most promising of the MRI volumetric measures for tracking disease progression. Similar to a smaller, single center study [16], all four region of interest volumes examined from MRIs collected from 48 different clinical centers were to some extent correlated with clinical change, although baseline volumes were not related to annual PSPRS change in trial completers. Importantly, the standard deviation of change in PSPRS score over one year in the AL-108-231 trial was less than previously reported, which resulted in greater power to detect a treatment effect than originally planned (Supplementary Table 2).

We identified a number of other clinical measures that could predict clinical decline in PSP, even after controlling for baseline disease severity and other potential confounding factors. Executive function is often the only reported neuropsychological deficit in the typical PSP (Richardson's) syndrome [17–19], and individuals with worse color trails scores, a measure of executive function, declined

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

Table 2
Correlation of 52-week changes in outcome measures with change in PSPRS.

Outcome measures		Mean 52 week change (95% confidence interval)	Range [min, max]	Cohen's d	Correlation with 52 week change in PSPRS (95% confidence interval)
Clinical (n = 241)					
PSPRS[a]		11.1* (9.9, 12.3)	[−9, 52]	1.1	–
SEADL[a]		−0.17* (−0.19, −0.15)	[−0.8, 0.3]	−1.0	−0.40* (−0.50, −0.29)
CGIC[b]		5.0* (4.8, 5.2)	[3, 7]	1.1#	0.43* (0.32, 0.53)
CGIds[c]		0.89* (0.78, 1.0)	[−1, 4]	1.0#	0.44* (0.33, 0.54)
GDS[c]		0.63 (−0.04, 1.3)	[−26, 13]	0.12	0.23* (0.10, 0.35)
Neuropsychological (n = 217)					
RBANS[c]	Total raw	−22.0* (−24.8, −19.3)	[−120, 21]	−0.88	−0.33* (−0.45, −0.21)
RBANS[c]	Total scaled	−6.4* (−5.3, −7.5)	[−36, 13]	−0.65	−0.26* (−0.38, −0.13)
Phonemic Fluency[c]	Words/min	−2.1* (−2.7, −1.5)	[−29, 9]	−0.46	−0.11 (−0.24, 0.02)
Letter number seq.[c]	Score	−1.1* (−1.4, -0.72)	[−12, 7]	−0.41	−0.22* (−0.34, −0.09)
Color Trails 1[c]	Seconds	33.6* (26.4, 40.9)	[−120, 202]	0.63	0.09 (−0.04, 0.27)
Color Trails 2[c]	Seconds	33.8* (26.6, 40.9)	[−138, 220]	0.56	0.02 (−0.11, 0.16)
MRI absolute volume change/TIV (n = 214)					
Ventricular volume[b]	$x10^{-4}$	28.8* (25.2, 32.4)	[−22, 156]	1.1	0.17 (0.03, 0.30)
Whole brain volume[c]	$x10^{-4}$	−83.8* (−101.7, −64.9)	[−341, 157]	−0.81	−0.16 (−0.29, −0.02)
Midbrain volume[c]	$x10^{-4}$	−1.8* (−1.9, −1.6)	[−6.1, 2.2]	−1.2	−0.25* (−0.37, −0.12)
SCP volume[c]	$x10^{-4}$	−0.19* (−0.23, −0.16)	[−1.2, 0.7]	−0.66	−0.16 (−0.29, −0.03)
MRI percent volume change (n = 214)					
Ventricular volume		9.3* (8.6, 10.0)	[−6.1, 39.1]	1.3	0.20* (0.07, 0.33)
Whole brain volume		−0.86* (−0.99, −0.73)	[−3.6, 2.0]	−0.82	−0.25* (−0.41, −0.15)
Midbrain volume		−3.5* (−3.9, -3.1)	[−14, 4.7]	−1.2	−0.31* (−0.43, −0.18)
SCP volume		−7.3* (−8.6, -6.0)	[−36, 22.7]	−0.75	−0.15 (−0.28, −0.01)

n values are from participants with complete 52 week data.
*$p < 0.05$ (FDR corrected).
#Cohen's h. Correlation coefficients are Pearson R^2 for all the parameters.
Abbreviations: PSPRS = Progressive Supranuclear Palsy Rating Scale; SEADL = Schwab and England Activities of Daily Living Scale; GDS = Geriatric Depression Scale; CGIds = Clinical Global Impression of Disease Severity; RBANS = Repeatable Battery for the Assessment of Neuropsychological Disease Status; SCP = superior cerebellar peduncle; TIV = total intracranial volume.
[a] Primary endpoints.
[b] Secondary endpoints.
[c] Exploratory endpoints.

Table 3
Regression of the baseline values on 52-week change in PSPRS in completers.

Baseline characteristics		Univariate slope (95% confidence interval)	Adjusted for multiple baseline characteristics[a] slope (95% confidence interval)
Demographics			
Age		−0.21 (−0.39, −0.02)	−0.14 (−0.33, 0.04)
Sex		−0.25 (−2.6, 2.1)	0.88 (−1.5, 3.2)
Weight		−0.04 (−0.11, 0.04)	−0.07 (−0.17, 0.02)
Tau haplotype			
H1/H1		−1.2 (−4.1, 1.7)	−1.0 (−4.0, 2.0)
H1/H2		−2.0 (−9.0, 5.0)	−1.3 (−8.0, 5.4)
Treatment/Medication Use			
Davunetide		−0.44 (−2.8, 1.9)	−0.74 (−3.1, 2.6)
CoQ10 use		2.4 (−0.58, 5.3)	1.7 (−1.2, 4.5)
Levodopa use		−0.13 (−2.5, 2.3)	−0.81 (−3.2, 1.5)
Clinical			
Disease Duration		−8.2 (−12.3, −4.1)*	−8.2 (−12.5, −3.9)*
PSPRS		−0.05 (−0.16, 0.06)	0.00 (−0.12, 0.12)
SEADL		−3.1 (−8.5, 2.3)	−6.0 (−13.6, 1.6)
GDS		0.21 (0.04, 0.38)	0.26 (0.08, 0.44)*
CGIds		−0.41 (−1.7, 0.91)	−0.18 (−1.7, 1.4)
MMSE		−0.13 (−0.49, 0.23)	−0.11 (−3.1, 1.6)
RBANS	Total raw	−0.04 (−0.08, −0.008)	−0.07 (−0.12, −0.02)*
RBANS	Total scaled	−0.11 (−0.20, −0.02)	−0.12 (−0.23, 0.003)
Fluency	Words/min	−0.25 (−0.42, −0.08)*	−0.24 (−0.42, −0.06)
Letter number	Score	−0.32 (−0.73, 0.09)	−0.37 (−0.88, 0.15)
Color Trails 1	Seconds	0.03 (0.01, 0.04)*	0.03 (0.01, 0.05)*
Color Trails 2	Seconds	0.03 (0.01, 0.04)*	0.04 (0.01, 0.05)*
MR Imaging			
Ventricular volume/TIV		−6 (−95.5, 83.5)	9.1 (−86.2, 104.3)
Whole brain volume/TIV		−25 (−51.7, 1.6)	−32.1 (−61.6, −2.6)
Midbrain volume/TIV		−1762 (−4157, 632)	−1758 (−4274, 758)
SCP volume/TIV ($x10^3$)		−17.1 (−33.1, −1.1)	−19.6 (−36.0, −2.3)

*$p < 0.05$ (FDR corrected).
Abbreviations: PSPRS = PSP Rating Scale; SEADL = Schwab and England Activities of Daily Living Scale; GDS = Geriatric Depression Scale; CGIds = Clinical Global Impression of Disease Severity; RBANS = Repeatable Battery for the Assessment of Neuropsychological Disease Status; SCP = superior cerebellar peduncle; TIV = total intracranial volume.
[a] Linear regression adjusted for baseline PSPRS, age, gender, disease duration, treatment group assignment, tau haplotype, CoQ10 use, and MMSE.

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

Table 4
Effect of baseline measurements on the trajectory of PSPRS change using linear mixed effect models.

Baseline characteristics	Mixed effects model					
	Fixed portion				Random portion estimate (SE)	
	Term	Z-score	Coefficient	p	Intercept	Slope
SEADL	Time	9.4	0.28	0.001	59.1 (5.3)	0.02 (0.003)
	SEADL	−16	−34	<0.001		
	SEADL × Time	−2.06	−0.11	0.045		
Color Trails 1	Time	5.0	0.14	<0.001	87.4 (7.6)	0.02 (0.003)
	CCT1	9.5	0.08	<0.001		
	CCT1 × Time	3.09	0.0005	0.006		
Color Trails 2	Time	2.7	0.10	0.007	91.6 (8.0)	0.02 (0.003)
	CCT2	8.5	0.07	<0.001		
	CCT2 × Time	3.2	0.0005	0.005		
RBANS raw	Time	6.8	0.32	<0.001	83.5 (7.2)	0.02 (0.003)
	RBANS	−10.2	−0.16	<0.001		
	RBANS × Time	−2.28	−0.0007	0.04		
RBANS scaled	Time	5.7	0.36	<0.001	94.2 (8.1)	0.02 (0.003)
	RBANS	−7.8	−0.34	<0.001		
	RBANS × Time	−2.32	−0.002	0.03		
Phonemic Fluency	Time	12.9	0.27	<0.001	102.1 (8.7)	0.02 (0.003)
	PHON FLU	−6.1	−0.53	<0.001		
	PHON FLU × Time	−2.9	−0.005	0.01		

SEADL = Schwab and England Activities of Daily Living Scale; RBANS = Repeatable Battery for the Assessment of Neuropsychological Disease Status.

at a faster rate. Surprisingly, the broader RBANS neuropsychological battery which includes tests of memory, attention, language and visuospatial function was also sensitive to change over one year, and individuals with more global cognitive impairments on RBANS at baseline had higher rates of annual PSPRS change. Depression and apathy are prominent in PSP [20,21], and we found that more severe baseline depression on the GDS was also associated with faster rates of PSPRS change. We identified cutoff values that could be used as selection criteria that could be used in a future study to define populations with more rapid or predictable disease progression.

4.2. Reducing attrition rate to improve power

Since patient drop out and a lack of evaluable data can diminish the power and quality of a trial, we first analyzed the baseline characteristics of study completers and dropouts, and found that there were no baseline characteristics that had a significant and meaningful effect on trial completion (Table 1). While dropouts had lower baseline RBANS scores than completers, the odds ratio for completing the trial based on RBANS was small (1.05). This suggests that in a PSP population similar to those recruited for the AL-108-231 trial, there are few changes to enrollment criteria that would have a major effect on study completion. Therefore it may be better to focus on other aspects of clinical trial procedures to identify ways to decrease attrition.

4.3. Limitations of current study

We combined data from the AL-108-231 study's treatment and placebo groups, which could have influenced the results if there was an undetected effect of the treatment (davunetide). This seems unlikely since the original trial analyses failed to identify an effect of treatment on any of the outcomes in either the primary intent-to-treat, completer, or a number of sensitivity analyses. Moreover, our analyses controlled for treatment group assignment.

4.4. Conclusion

We identified clinical and MRI measures that were able to capture change in a diverse population of PSP patients, and whose baseline values also related to clinical decline over the course of a year in a large multicenter trial. Together, these data provide support for inclusion of specific scales and imaging tools in future PSP clinical trials.

Author's roles

Jee Bang: Drafting/revising the manuscript for content, including medical writing for content, study concept/design, analysis/interpretation of data.

Iryna Lobach: Statistical analysis, drafting/revising the manuscript for content, including medical writing for content, study concept/design, analysis/interpretation of data.

Anthony Lang: Drafting/revising the manuscript for content.

Murray Grossman: Drafting/revising the manuscript for content.

David Knopman: Drafting/revising the manuscript for content.

Bruce Miller: Drafting/revising the manuscript for content.

Lon Schneider: Drafting/revising the manuscript for content.

Rachelle Doody: Drafting/revising the manuscript for content.

Andrew Lees: Drafting/revising the manuscript for content.

Michael Gold: Drafting/revising the manuscript for content.

Bruce Morimoto: Drafting/revising the manuscript for content.

Adam Boxer: Drafting/revising the manuscript for content, including medical writing for content, study concept/design, analysis/interpretation of data.

Financial disclosures of all authors

Jee Bang: none.

Iryna Lobach: none.

Anthony Lang: Dr. Lang has served as an advisor for Abbvie, Acorda, Avanir Pharmaceuticals, Bristol Myers Squibb, Cipla, Intekrin, and Merck; received honoraria from Medtronic, Teva, UCB, AbbVie; received grants from Brain Canada, Canadian Institutes of Health Research, Edmond J Safra Philanthropic Foundation, Michael J. Fox Foundation, the Ontario Brain Institute, National Parkinson Foundation, Parkinson Society Canada, Physicians Services Incorporated (PSI), W. Garfield Weston Foundation; received publishing royalties from Saunders, Wiley-Blackwell, Johns Hopkins Press, and Cambridge University Press; and has served as an expert witness in cases related to the welding industry.

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014

Murray Grossman: Dr. Grossman is an investigator on the BMS CN002003 study.

David Knopman: Dr. Knopman serves on a Data Safety Monitoring Board for Lundbeck Pharmaceuticals and for the DIAN study; is an investigator in clinical trials sponsored by Biogen, TauRX Pharmaceuticals, Lilly Pharmaceuticals and the Alzheimer's Disease Cooperative Study; and receives research support from the NIH.

Bruce Miller: none.

Lon Schneider: Dr. Schneider (within the past 3 years) has received grant or research support for industry-sponsored studies from Pfizer, Baxter, Eli Lilly, Forum, Lundbeck,, Merck, Novartis (with the Alzheimer Prevention Initiative and NIH), Roche/Genentech, Tau Rx; for NIH sponsored research, USC ADRC, ADCS (UCSD), ADNI (NCIRE), phytoSERMs, allopreganolone, clinical trials simulations, Banner Alzheimer Prevention Initiative, P50 AG05142, R01 AG033288, R01 AG037561, UF1 AG046148; from the State of California, the California Alzheimer Disease Center (CADC) and California Institute for Regenerative Medicine (CIRM). Dr. Schneider, within the past 3 years, has consulted or served on committees for AC Immune, Accera, Allon, Avraham, Axovant, Baxter, Boehringer Ingelheim, Cerespir, Cognition, Forum, Insys, Merck, Neurim, Novartis, Roche, Stemedica, Takeda, TauRx, Transition, vTv Therapeutics, Toyama/FujiFilm, and Zinfandel.

Rachelle Doody: Dr. Doody's institution receives research support from Accera, Biogen, Genentech, Merck and Takeda. She has been a consultant to AC Immune, AgeneBio, Avanir, Axovant, AZ Therapies, Biogen, Cerespir, Corium, Forum, Hoffman LaRoche, Shanghai Green Valley, Suven, and Takeda, and holds stock options in AZ Therapies and QR Pharma.

Andrew Lees: none.

Michael Gold: Dr. Gold is employed full time at Pharmaceutical Product Development, Inc.

Bruce Morimoto: Dr. Morimoto is a former employee of Allon Therapeutics Inc.

Adam Boxer: 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 Abbvie, Asceneuron, Ipierian, Ionis, Janssen and Merck. He has stock/options in Alector and Delos.

Financial disclosure/conflict of interest

None.

Study funding

Supported by R01AG038791, U54NS092089, T32 AG23481, and the Tau Consortium. These sources had no involvement in the study design. The funding for AL-108-231 was provided by Allon Therapeutics Inc.

Acknowledgements

We thank the patients who participated in the AL-108-231 study and their families, as well as the investigators. **The AL-108-231 Study Group:** *Australia:* David Williams; *Canada:* Anne Louise Lafontaine, Connie Marras, Mandar Jog, Michael Panisset, Anthony Lang, Lesley Parker, Alistair J. Stewart; *France:* Jean-Christophe Corvol, Jean-Philippe Azulay, Philippe Couratier; *Germany:* Brit Mollenhauer, Stefan Lorenzl, Albert Ludolph, Reiner Benecke, Gunter Hoglinger, Axel Lipp, Heinz Reichmann, Dirk Woitalla; *United Kingdom*: Dennis Chan, Adam Zermansky, David Burn, Andrew Lees; *Israel:* Illana Gozes, *United States:* Adam L. Boxer, Bruce L. Miller, Iryna V. Lobach, Erik Roberson, Lawrence Honig, Edward Zamrini, Rajesh Pahwa, Yvette Bordelon, Erika Driver-Dunkley, Stephanie Lessig, Mark Lew, Kyle Womack, Brad Boeve, Joseph Ferrara, Argyle Hillis, Daniel Kaufer, Rajeev Kumar, Tao Xie, Steven Gunzler, Theresa Zesiewicz, Praveen Dayalu, Lawrence Golbe, Murray Grossman, Joseph Jankovic, Scott McGinnis, Anthony Santiago, Paul Tuite, Stuart Isaacson, Julie Leegwater-Kim, Irene Litvan, Murray Grossman, David S. Knopman, Bruce L. Miller, Lon S. Schneider, Rachelle S. Doody, Lawrence I. Golbe, Erik D. Roberson, Mary Koestler, Clifford R. Jack, Jr., Viviana Van Deerlin, Christopher Randolph, Steve Whitaker, Joe Hirman, Michael Gold, Bruce H. Morimoto.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.parkreldis.2016.04.014.

References

[1] D.R. Williams, A.J. Lees, Progressive supranuclear palsy: clinicopathological concepts and diagnostic challenges, Lancet Neurol. 8 (2009) 270–279.
[2] I. Litvan, Y. Agid, D. Calne, G. Campbell, B. Dubois, R.C. Duvoisin, C.G. Goetz, L.I. Golbe, J. Grafman, J.H. Growdon, M. Hallett, J. Jankovic, N.P. Quinn, E. Tolosa, D.S. Zee, Clinical research criteria for the diagnosis of progressive supranuclear palsy (Steele-Richardson-Olszewski syndrome): report of the NINDS-SPSP international workshop, Neurology 47 (1996) 1–9.
[3] G. Respondek, M. Stamelou, C. Kurz, L.W. Ferguson, A. Rajput, W.Z. Chiu, J.C. van Swieten, C. Troakes, S. Al Sarraj, E. Gelpi, C. Gaig, E. Tolosa, W.H. Oertel, A. Giese, S. Roeber, T. Arzberger, S. Wagenpfeil, G.U. Höglinger, Movement Disorder Society endorsed PSP Study Group, The phenotypic spectrum of progressive supranuclear palsy: a retrospective multicenter study of 100 definite cases, Mov. Disord. 29 (2014) 1758–1766.
[4] R.M. Tsai, A.L. Boxer, Clinical trials: past, current, and future for atypical parkinsonian syndromes, Semin. Neurol. 34 (2014) 225–234.
[5] G. Bensimon, A. Ludolph, Y. Agid, M. Vidailhet, C. Payan, P.N. Leigh, NNIPPS Study Group, Riluzole treatment, survival and diagnostic criteria in Parkinson plus disorders: the NNIPPS study, Brain 132 (2009) 156–171.
[6] A.L. Boxer, A.E. Lang, M. Grossman, D.S. Knopman, B.L. Miller, L.S. Schneider, R.S. Doody, A. Lees, L.I. Golbe, D.R. Williams, J.C. Corvol, A. Ludolph, D. Burn, S. Lorenzl, I. Litvan, E.D. Roberson, G.U. Höglinger, M. Koestler, C.R. Jack Jr., V. Van Deerlin, C. Randolph, I.V. Lobach, H.W. Heuer, I. Gozes, L. Parker, S. Whitaker, J. Hirman, A.J. Stewart, M. Gold, B.H. Morimoto, AL-108-231 Investigators, Davunetide in patients with progressive supranuclear palsy: a randomised, double-blind, placebo-controlled phase 2/3 trial, Lancet Neurol. 13 (2014) 676–685.
[7] E. Tolosa, I. Litvan, G.U. Höglinger, D. Burn, A. Lees, M.V. Andrés, B. Gómez-Carillo, T. León, T. Del Ser, TAUROS Investigators, A phase 2 trial of the GSK-3 inhibitor tideglusib in progressive supranuclear palsy, Mov. Disord. 29 (4) (2014 Apr) 470–478.
[8] L.I. Golbe, P.A. Ohman-Strickland, A clinical rating scale for progressive supranuclear palsy, Brain 130 (2007) 1552–1565.
[9] R. Schwab, A. England, Projecton technique for evaluating surgery in Parkinson's disease, in: F. Gillingham, M. Donaldson (Eds.), Third Symposium on Parkinson's Disease Research, ES Livingston, Edinburgh, Scotland, 1969, pp. 152–157.
[10] L.S. Schneider, J.T. Olin, R.S. Doody, C.M. Clark, J.C. Morris, B. Reisberg, F.A. Schmitt, M. Grundman, R.G. Thomas, S.H. Ferris, 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. 11 (2) (1997) S22–S32.
[11] C. Randolph, M.C. Tierney, E. Mohr, T.N. Chase, The repeatable battery for the assessment of neuropsychological status (RBANS): preliminary clinical validity, J. Clin. Exp. Neuropsychol. 20 (1998) 310–319.
[12] J.A. Yesavage, T.L. Brink, T.L. Rolse, O. Lum, V. Huang, M. Adey, V.O. Leirer, Development and validity of a geriatric depression scale: a preliminary report, J. Psychiatr. Res. 17 (1983) 37–49.
[13] C.A. Payan, F. Viallet, B.G. Landwehrmeyer, A.M. Bonnet, M. Borg, F. Durf, L. Lacomblez, M. Verny, J. Fermanian, Y. Agid, A.C. Ludolph, P.N. Leigh, G. Bensimon, NNIPPS Study Group, Disease severity and progression in progressive supranuclear palsy and multiple system atrophy: validation of the NNIPPS-Parkinson plus scale, PLoS One 6 (8) (2011) e22293.
[14] U. Gromping, Relative importance for linear regression in R: the package relaimpo, J. Stat. Softw. 17 (1) (2006 Oct) 1–27.
[15] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, J. R. Stat. Soc. Ser. B Stat. Methodol. 57 (1995) 289–300.
[16] J.L. Whitwell, J. Xu, J.N. Mandrekar, J.L. Gunter, C.R. Jack Jr., K.A. Josephs, Rates of brain atrophy and clinical decline over 6 and 12-month intervals in PSP: determining sample size for treatment trials, Park. Relat. Disord. 18 (3) (2012 Mar) 252–256.

ARTICLE IN PRESS

[17] D. Millar, P. Griffiths, A.J. Zermansky, D.J. Burn, Characterizing behavioral and cognitive dysexecutive changes in progressive supranuclear palsy, Mov. Disord. 21 (2006) 199–207.
[18] R.G. Brown, L. Lacomblez, B.G. Landwehrmeyer, T. Bak, I. Uttner, B. Dubois, Y. Agid, A. Ludolph, G. Bensimon, C. Payan, N.P. Leigh, NNIPPS Study Group, Cognitive impairment in patients with multiple system atrophy and progressive supranuclear palsy, Brain 133 (2010) 2382–2393.
[19] A. Gerstenecker, B. Mast, K. Duff, T.J. Ferman, I. Litvan, Executive dysfunction is the primary cognitive impairment in progressive supranuclear palsy, Arch. Clin. Neuropsychol. 28 (2013) 104–113.
[20] I. Litvan, M.S. Mega, J.L. Cummings, L. Fairbanks, Neuropsychiatric aspects of progressive supranuclear palsy, Neurology 47 (1996) 1184–1189.
[21] A. Schrag, S. Sheikh, N.P. Quinn, A.J. Lees, C. Selai, C. Mathias, I. Litvan, A.E. Lang, J.H. Bower, D.J. Burn, P. Low, M. Jahanshahi, A comparison of depression, anxiety, and health status in patients with progressive supranuclear palsy and multiple system atrophy, Mov. Disord. 25 (2010) 1077–1081.
[22] D.R. Williams, A.M. Pittman, T. Revesz, A.J. Lees, R. de Silva, Genetic variation at the tau locus and clinical syndromes associated with progressive supranuclear palsy, Mov. Disord. 22 (2007) 895–897.

Please cite this article in press as: J. Bang, et al., Predicting disease progression in progressive supranuclear palsy in multicenter clinical trials, Parkinsonism and Related Disorders (2016), http://dx.doi.org/10.1016/j.parkreldis.2016.04.014