Title

Longitudinal magnetic resonance imaging in progressive supranuclear palsy: A new combined score for clinical trials

Journal

Movement Disorders, 32(6)

ISSN

0885-3185

Authors

Höglinger, Günter U
Schöpe, Jakob
Stamelou, Maria
et al.

Publication Date

2017-06-01

DOI

10.1002/mds.26973

Peer reviewed

HHS Public Access

Author manuscript

Mov Disord. Author manuscript; available in PMC 2018 February 12.

Published in final edited form as:

Mov Disord. 2017 June ; 32(6): 842–852. doi:10.1002/mds.26973.

Longitudinal MRI in progressive supranuclear palsy: a new combined score for clinical trials

Günter U. Höglinger, MD[1,2,3], **Jakob Schöpe, MSc**[4], **Maria Stamelou, MD**[3,5], **Jan Kassubek, MD**[6], **Teodoro del Ser, MD**[7], **Adam L Boxer, MD, PhD**[8], **Stefan Wagenpfeil, PhD**[4], and **Hans-Jürgen Huppertz, MD**[9] **for the AL-108-231 Investigators, the Tauros MRI Investigators, and the Movement Disorder Society-endorsed PSP Study Group**

[1]Department of Neurology, Technische Universität München, Munich, Germany [2]German Center for Neurodegenerative Diseases (DZNE), Munich, Germany [3]Department of Neurology, University Hospital Gießen and Marburg, Marburg, Germany [4]Institute for Medical Biometry, Epidemiology and Medical Informatics, Saarland, University, Campus Homburg, Germany [5]Second Department of Neurology, Attikon University Hospital, University of Athens, Greece [6]Department of Neurology, University of Ulm, Ulm, Germany [7]Medical Department, Noscira SA, Madrid, Spain. Present affiliation: Alzheimer Project Research Unit, Fundación CIEN, Madrid, Spain [8]Memory and Aging Center, Department of Neurology, University of California, San Francisco, CA, USA [9]Swiss Epilepsy Centre, Klinik Lengg, Zurich, Switzerland

Abstract

Background—Two recent randomized, placebo-controlled phase II/III trials (clinicaltrials.gov: NCT01110720, NCT01049399) of davunetide and tideglusib in progressive supranuclear palsy (PSP) generated prospective, 1-year longitudinal datasets of high-resolution T1-weighted 3D MRI.

Objectives—Develop a quantitative MRI disease progression measurement for clinical trials.

Methods—We performed a fully automated quantitative MRI analysis employing atlas-based volumetry and provide sample size calculations based on data collected in N=99 PSP patients, assigned to placebo in these trials. Based on individual volumes of N=44 brain compartments and structures at baseline and 52 weeks follow-up, means and standard deviations of annualized percentage volume changes were used to estimate standardized effect sizes and the required sample sizes per group for future two-armed, placebo-controlled therapeutic trials.

Results—The highest standardized effect sizes were found for midbrain, frontal lobes, and third ventricle. Using the annualized percentage volume change of these structures to detect a 50% change in the 1-year progression (80% power, significance level 5%) required lower numbers of patients per group [third ventricle, N=32; midbrain, N=37; frontal lobe, N= 43] than the best clinical scale (PSP rating scale total score, N=58). A combination of volume changes in these

Correspondence to: Prof. Dr. Günter U. Höglinger, Dept. of Translational Neurodegeneration, German Center for Neurodegenerative Diseases (DZNE), Feodor-Lynen Str. 17, D-81677 Munich, Germany. Phone: +49-89-4400-46464, Fax: +49-89-4400-46565, guenter.hoeglinger@dzne.de.

Financial Disclosure/Conflict of Interest concerning the research related to the manuscript: none

Founding source for study: none

three structures reduced the number of required patients to only N=20 and correlated best with the progression in the clinical scales.

Conclusions—We propose the 1-year change in the volumes of third ventricle, midbrain, and frontal lobe as combined imaging read-out for clinical trials in PSP, requiring the least number of patients for detecting efficacy to reduce brain atrophy.

Keywords

Progressive supranuclear palsy; magnetic resonance imaging; volumetry; power calculation; clinical trials

Introduction

Progressive supranuclear palsy (PSP) is a neurodegenerative disease caused by intracellular aggregation of the tau protein. There is currently no treatment approved.[1] However, several compounds are awaiting clinical evaluation. Rational trial design is therefore important to allow clear conclusions while saving resources.

The randomized, placebo-controlled phase II/III trials of davunetide and tideglusib (clinicaltrials.gov: NCT01110720, NCT01049399) studied patients according to similar protocols for 12 months.[2–4] Their placebo-data allowed to identify rating scales suited to measure disease progression and to estimate sample sizes for future trials.[5] These trials also generated prospective high-resolution T1-weighted 3D MRI data. A recent study analyzed the atrophy rate of four selected volumes (whole brain, ventricles, superior cerebellar peduncle, and midbrain) using label propagation with Statistical Parametric Mapping (SPM) 5 software in all (i.e. placebo and verum) patients from the davunetide trial and concluded that midbrain volumetry as outcome measure in a 1-year parallel-group trial would require comparable numbers of patients[6] as the best performing clinical scale, the PSP-rating scale (PSPRS).[5,7] Another study[8] proposed that annual midbrain volume change measured with SPM12 would require 44% less patients than the PSPRS.

Unbiased by *a priori* hypotheses, we aimed to analyze which volumetric parameters perform best as outcome measures, allowing trials with fewer patients despite comparable power to detect efficacy to slow brain atrophy. We used fully-automated atlas-based volumetry (ABV) to quantify structural changes in individual patients with an intrascanner variability of <1%.[9,10] In the tideglusib study, this method revealed lower cerebral atrophy rates in verum- vs. placebo-treated patients.[4] Furthermore, in conjunction with support vector machine classification, this method reliably identified PSP patients in a large mixture of parkinsonian syndromes.[11] Here, we performed ABV using SPM12 on 44 brain structures of N=99 placebo-patients from the davunetide and tideglusib trials at baseline and 1-year follow-up, calculated annualized volume changes, and estimated standardized effect sizes and sample sizes for two-armed, placebo-controlled therapeutic trials. We also studied if the cross-sectional regional atrophy at baseline would predict longitudinal atrophy rates. Finally, we correlated longitudinal regional volumes changes with change in clinical scales, to examine their relevance as surrogate markers for disease progression.

Materials and Methods

Study population

Raw data were obtained from PSP patients of the placebo arms from two randomized, controlled trials with similar inclusion-exclusion criteria.[2,3] In the davunetide trial,[2] N=313 patients from 48 centres met the following criteria: at least a 12-month history of postural instability or falls during the first 3 years from onset; reduced downward saccade velocity or supranuclear gaze palsy; an akinetic-rigid syndrome with prominent axial rigidity; ability to take at least five steps with minimal assistance. In the tideglusib trial,[3] N=139 patients from 24 centres met the possible or probable NINDS-SPSP criteria;[12] ability to ambulate independently or with minimal assistance (PSP staging system[13] score < 5). The tideglusib MRI sub-study was conducted in 17 European sites.[4] Detailed inclusion and exclusion criteria, clinical assessments and follow-up visits schedules are reported elsewhere.[2–4]

For comparison, healthy controls without neurological or psychiatric disease were recruited in a similar time period in the Departments of Neurology at Marburg and Ulm, Germany.[11] A subset of N=50 of these controls was matched to the PSP-cohort for age and gender distribution.

Ethics approval was obtained at each site from the local ethics committee, and all participants gave written informed consent.

Clinical assessments

The Schwab & England Activities of Daily Living (SEADL) scale,[14] PSPRS,[7] and Clinical Global Impression of Disease Severity (CGIDS)[15] were obtained at the baseline and 52-week visit in both trials.

Image acquisition

The core MRI protocol for all participants comprised 3D T1-weighted 3D magnetization prepared rapid gradient echo (MPRAGE) sequence with 1x1x1 mm resolution according to the Alzheimer's Disease Neuroimaging Initiative (UCLA, CA, USA; www.loni.ucla.edu/ADNI) recommendations for volumetric analysis, and axial T2 sequences to detect vascular lesions, acquired on 1.5 or 3T scanners. Patients were scanned within 4 weeks prior to baseline and within ±1 (davunetide) or ±4 (tideglusib) weeks of the 52-week visit. For each patient, baseline and follow-up MRI were acquired on the same scanner using the same sequence parameters. Controls were scanned only once (baseline). We performed quality control of each scan and excluded those with obvious factors biasing volumetry (e.g. movement artefacts, signs of head traumatization, or different sequences at baseline and follow-up).

Image processing/atlas-based volumetry

MPRAGE sequences were pseudonymised, converted to ANALYZE 7.5 format, and analyzed by a fully-automated, observer-independent method of atlas- and mask-based volumetry on MATLAB (R2014b, Mathworks, USA) using SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK, www.fil.ion.ucl.ac.uk/spm), as described (Fig. 1).[9–11] In

short, each T1-weighted volume dataset was normalized to Montreal Neurological Institute (MNI) template space using diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL)[16] and segmented into different brain compartments, i.e. gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) using the 'unified segmentation' algorithm of SPM12 with default parameters. Volumetric measures of brain structures were calculated by voxel-by-voxel multiplication and subsequent integration of normalized and modulated component images (GM, WM or CSF) with predefined masks in the same space. Masks were derived from different probabilistic brain atlases, because not all regions of interest are comprised in a single atlas: the Harvard-Oxford atlas of subcortical structures distributed with the FSL package[17–20] for the hippocampus, amygdala, caudate, putamen, nucleus accumbens, pallidum and thalamus; the Hammers_mith atlas n30r83[21] for third and lateral ventricles; the LONI Probabilistic Brain Atlas (LPBA40)[22] for all other structures. GM, WM and CSF volumes and intracranial volume (ICV) were determined by the "Tissue Volumes" utility of SPM12.[23] The volumes of N=37 brain structures and compartments and the areas of midsagittal planes across N=7 structures were determined. In the midsagittal planes, several previous studies suggested the midbrain area and midbrain tegmentum area as reliable markers of atrophy in PSP.[24,25] For comparison, the midsagittal areas of other structures (corpus callosum, pons and pars basilaris of pons, medulla oblongata, cerebellar vermis) have also been determined. All cross-sectional results of ABV were ICV-corrected and normalized to the mean ICV of the whole study population.[26] Each dataset was processed independently with the same protocol. Processing of follow-up scans did not require co-registration to baseline scans. Differences in absolute volumes and areas between baseline and follow-up were transformed to annualized percentage changes for each patient.

Statistical analysis

Data are given as mean ± standard deviation, unless indicated otherwise. Normality of data was verified with the Kolmogorov-Smirnov-test. Patients and controls were compared by two-sample t-tests for age, and by chi-squared test for gender distribution. Score changes in patients between baseline and follow-up were compared by paired t-test. P-values are shown both uncorrected and corrected for multiple testing based on Holm's method.[27] We calculated Pearson's r for linear correlation analysis.

Calculation of standardized effect sizes and sample sizes

Sample size calculation was conducted using the pwr-package in R version 3.2.1 [28] with cases for which both baseline and follow-up measurements were available. Standardized effect sizes were estimated as:

$$d = \frac{\mu \Delta \mathrm{v}}{\sigma \Delta \mathrm{v}} \quad (1)$$

where μΔv and σΔv are the mean and standard deviation of the observed annualized volume differences in a particular intracranial structure or compartment without treatment. Sample size calculations were based on a two-sided significance level (α) of 5%, and a power (1-β) of 80%. Assuming a normal distribution and equal variances in two equally sized groups

(control and intervention group), minimum required sample sizes per group for an independent two-sample t-test were obtained using the following equation, where d is the standardized effect size according to (1), and z^x denotes the x^{th} quantile of the standard normal distribution:

$$N_{group} = \frac{1+q}{q} * \frac{\left(z_{1-\frac{\alpha}{2}} + z_{1-\beta}\right)^2}{d^2} + \frac{z^2_{1-\alpha/2}}{2\,(1+q)}$$

Assuming equally sized groups defined by the allocation ratio (with q = 1), one would yield:

$$N_{group} = \frac{15.6978}{d^2} + 0.9604$$

Including e as the expected treatment effect (e.g. e = 0.5 for 50% reduction of atrophy rate) the formula would be:

$$N_{group} = \frac{15.6978}{(d * e)^2} + 0.9604$$

For a standardized effect size of 1, e.g., this would yield:

$$N_{group} = \frac{15.6978}{(1 * 0.5)^2} + 0.9604 = 63.8 \cong 64$$

For the Mann-Whitney U test, the result is derived by dividing N_{group} by 0.864.[29]

Results

Study population

N=106 patients (97 from davunetide/9 from tideglusib; 56 female/50 male) received placebo and had MRI data at baseline and follow-up of sufficient quality to be considered for analysis. We excluded N=7 of these because of factors biasing volumetry (movement artefacts, signs of head traumatization, different sequences at baseline and follow-up; Supplementary Fig. 1). N=50 healthy controls from our previously reported series[11] were matched for age at baseline and gender distribution.

The clinical features of the included PSP patients and controls are reported in Table 1, showing demographic data, rating scales scores at baseline and follow-up visits, 1-year differences, standard effect sizes, and sample sizes required for a two-arm 1-year follow-up therapeutic trial to detect 50%-change in the progression of these scales without adjusting for an expected drop-out rate. These data differ slightly from our previous report[5] since the former included all N=187 placebo-patients both trials, while the current work focuses patients with available MRI data only. The davunetide study did not use NINDS-SPSP criteria and thus did not distinguish between possible and probable PSP. From the tideglusib study, 6 patients had probable, 3 possible PSP. Based on their inclusion criteria, both studies

mainly recruited Richardson's syndrome, but characterization of PSP-phenotypes was not part of their protocol.

Volumetric analyses

Table 2 shows the ABV quantification of the N=44 brain structures of controls and patients at baseline and follow-up, the mean differences between controls and patients at baseline, the mean differences between patients at baseline and follow-up, and the standardized effect sizes. The last three columns display the required numbers of patients per group in a therapeutic study with an expected treatment effect of 20%, 30% or 50%, respectively, i.e. when the annual volume change expected by the natural course of the disease is reduced by these amounts (the two-sample t-test can be applied for normal distributions of volume results, otherwise the Mann-Whitney U test would be appropriate).

The compartment with the highest effect size, i.e. the third ventricle, required a 45% smaller number of patients than the clinical rating scale with the highest effect size (PSPRS total score) to detect a treatment effect of 50% (N=32 vs. N=58, cf. last column in Table 1 and 2, respectively).

To further minimize the numbers of patients required by combined analysis of individual structures, we selected the three non-overlapping structures with the highest individual effect sizes, i.e. the third ventricle, frontal lobe, and midbrain. The annualized percentage volume changes of these structures were summed up for each individual patient, with negative signs for the third ventricle, since the enlargement of the CSF space runs counter the atrophy of the brain parenchyma. This approach further increased the effect size and reduced the required sample size for detecting a 50% treatment effect to N=20, i.e. 65% less than the PSPRS total score (Table 2).

For completeness, we also present the demographic and clinical data (Supplementary Table 1) and ABV analyses (Supplementary Table 2) of the total PSP cohort (N=106 patients) without exclusion of technically compromised MRIs, which only marginally decreased the standardized effect sizes and increased the required sample sizes, demonstrating the strong reliability of this approach.

Correlation analysis

First, we analyzed if the degree of regional atrophy at baseline would predict the rate of longitudinal atrophy in PSP patients during the follow-up period. Therefore, we calculated on a *group* level the correlation of a) the mean cross-sectional difference at baseline between the N=99 patients vs. all N=50 controls, with b) the mean longitudinal annualized change in all patients. The correlation was moderate when including all structures as data points (N=44, $r=0.69$, $p<0.01$), strong when including CSF structures only (N=3, $r=0.99$, $p<0.05$), absent when including parenchymal structures only (N=41, $r=0.08$, not significant), and absent when including frontal lobe, midbrain, and third ventricle only (N=3, $r=0.97$, not significant). These data demonstrate that on a group level, the baseline atrophy of parenchymal structures in PSP does not predict the atrophy rate in the following 1-year period.

Also when using the *individual* patients' values as data points (N=99), there was no significant correlation between a) the cross-sectional %-differences to controls at baseline with b) the longitudinal annualized %-change for the frontal lobe, midbrain, and third ventricle and for their combination (frontal lobe + midbrain − third ventricle). These data demonstrate that the degree of atrophy in individual PSP patients at baseline does not predict the ensuing atrophy in these regions of interest. However, the %-difference vs. controls at baseline in the CSF volumes in individual patients significantly correlated with their longitudinal %-changes ($r=0.31$, $p<0.01$).

Finally, we analyzed if the rate of regional brain atrophy during the follow-up period would predict the clinical disease progression. Therefore, we correlated the annualized change of the three structures with the highest individual effect sizes (and their combination) with the annualized change in the three most relevant clinical scales (Table 3). SEADL and PSPRS total score significantly correlated with the volume change in third ventricle, frontal lobe and midbrain individually. The highest correlation of SEADL, CGIDS and PSPRS total score was observed with the combined volume change of these three structures.

Discussion

The study evaluated the 1-year change in brain structures by ABV on prospectively acquired 3D MRI datasets of N=99 placebo-treated PSP patients from the davunetide and tideglusib trials. We found that the annualized % volume change of one single structure (third ventricle) required 45% less patients than the clinical scale with the best effect size (PSPRS total score) to detect a 50% treatment effect (N=32 vs. N=58), when including only patients with excellent MRIs. Combined analysis of the three non-overlapping compartments with the highest individual effect sizes (third ventricle, frontal lobe, midbrain) even allowed to reduce the required sample size to N=20 (65% less than the PSPRS total score), demonstrating low variability in the atrophy pattern in our cohort. Furthermore, the combined volume changes of these three structures significantly correlated with the score changes in relevant clinical scales (SEADL, CGIDS, PSPRS total score), suggesting this imaging parameter to have clinical relevance as biomarker of disease progression. Thus, we propose the combined analysis of volume change within 1-year in the third ventricle, frontal lobe and midbrain as outcome measure for clinical trials aimed at determining disease modification in PSP.

One prior monocentric prospective study in N=17 PSP patients analyzed MRI volumetric changes in N=7 brain structures (brain, pons, midbrain, superior cerebellar peduncle, cerebellum, lateral ventricles, third ventricle) at baseline and 1-year follow-up[30] and determined the midbrain volume as the parameter requiring the smallest sample size (N=147 for a presumed treatment effect of 30% and power of 90%). Another monocentric prospective study in N=16 PSP patients analyzed MRI volumetric changes in N=5 brain structures (brain, ventricles, superior frontal lobe, thalamus, midbrain) after a 1-year interval[31] and also identified the midbrain volume as the parameter requiring the smallest sample size (N=84 for a treatment effect of 30% and power of 90%). A recent study[6] analyzed the change rate in N=4 brain volumes (whole brain, ventricles, superior cerebellar peduncle, midbrain) in N=189 placebo- and verum-patients from the davunetide trial, and

again identified midbrain volume as best performing outcome measure for a 1-year trial (N=113 at a treatment effect of 25% and power of 90%). The numbers in the studies including both *possible* and *probable* PSP patients[6,30] compare well with our results for midbrain volume (N=100 for a treatment effect of 30% and power of 80%) and thus, suggest reliability of this parameter across different populations and methods of volumetric analyses. The smaller numbers reported by Whitwell et al.[31] may result from the fact that they only included patients with a clinical diagnosis of *probable* PSP (NINDS-SPSP criteria), whereas the other studies[6,30] and our study also included patients with *possible* PSP or modified diagnostic criteria, which may affect the variability in disease progression markers. Finally, a most recent study [8] used SPM12-based longitudinal morphometric analysis and found that the annual change in midbrain volume as outcome measure requires only 56% of patients as the PSPRS. This rate compares well with the 63% observed in our study, which also used SPM12. The different absolute case numbers resulting from the power calculations between the former[8] and the current work were most probably the result from differences in the study populations.

The regions analyzed in these prior studies[6,8,30,31] have been selected on the basis of preceding imaging and pathological studies, which had demonstrated atrophy in PSP patients vs. healthy controls. Measurements of these regions, and ratios thereof (e.g. midbrain-pons ratio, MR-Parkinson-index) proofed therefore very helpful as markers for the cross-sectional differential diagnosis of Parkinson syndromes,[24,25] however, only limited data addressed the utility of these ratios as progression measurement.[8]

In contrast to these studies, we hypothesized here that the cross-sectional degree of atrophy at baseline in patients vs. controls might not be reliable predictors for the ensuing longitudinal atrophy rate in PSP patients, since regional atrophy rates might change during the course of the disease. Particularly, regions showing already marked atrophy at baseline might have lower atrophy rates during follow-up. Consistently, we did not find significant correlations of baseline atrophy rates with the longitudinal changes in the N=41 parenchymal structures analyzed. Moreover, regions with similar degrees of baseline atrophy (about 15%–16%) showed surprisingly high divergence in the 1-year volume change in our cohort (e.g. cerebellum white matter +0.1%, pallidum −1.5%, midbrain −2.3%). Furthermore, longitudinal atrophy rates do not translate linearly into sample sizes, since the latter incorporate both the atrophy rate and the variability of the measurement, which again depends on both inter-individual and methodological variability.

Therefore, we undertook for the first time an unbiased approach to identify the MRI measurement providing the highest effect size by starting our analysis with N=44 brain structures or compartments and identified unexpectedly the third ventricle as best individual structure. Also, the approach to combine the annualized volume change in three non-redundant brain structures had not been attempted in any prior study in PSP to our knowledge. This approach seems promising, since it allows further reducing the required sample sizes for clinical trials.

The observation that outcome measures based on neuroimaging can translate into very small sample sizes for clinical studies has already been described for Alzheimer’s or Huntington’s

disease or multiple sclerosis.[32–35] Advantages of imaging above clinical parameters, which have been reported previously[36] and was also recognized in our study, are probably due to the higher variability of clinical ratings, which reduces effect sizes. Among the neuroimaging parameters ventricular volumes often achieve highest effect sizes, perhaps due to more distinct boundaries than in other structures.[32,33] Consequently, also in our study the highest effect size was determined for a CSF compartment, i.e. the third ventricle. However, combining several measures has been proposed to further increase effect sizes[37] and also in our study turned out to be the optimal approach.

Limitations of our study concern the study population, intra- scanner variability, longitudinal fitting models, inter-scanner variability, and the generalizability of the results. All patients included in our study fulfilled the MDS-criteria for probable PSP-Richardson syndrome (‘vertical supranuclear gaze palsy’ or ‘slow vertical saccades’, with ‘repeated unprovoked falls within 3 years’ or ‘tendency to fall on the pull-test within 3 years’)[38] and had an annualized PSP-RS progression consistent with prior cohorts,[5] however, they were not consecutively referred, but volunteers participating in clinical trials, and might therefore differ from other PSP patients.

Controls were only scanned once, which did not allow accurately correcting for the influence of aging on ABV. The calculation of longitudinal volume changes in individual patients was mainly affected by intra-scanner variability, since baseline and follow-up MRI was acquired on the same scanner using the same sequences for each patient. Intra-scanner variability of atlas-based volumetry was less than 1% for most structures in a prior study comparing two MRIs in single subjects.[39] In the present study, atrophy rates (and subsequently effect sizes and sample sizes) are based on the average of N=99 patients. Consequently, intra-scanner variability influenced the results by a factor of 99 less, i.e., in the range of 0.01% of the measured volumes. Compared to atrophy rates of the relevant structures [2.6–7.6%], this influence can be regarded as negligible. The ideal modelling of longitudinal volumetric changes and the choice of a suitable fitting model (e.g. linear or non-linear, for example quadratic) are debated issues. There is evidence that non-parametric approaches may be advantageous in certain instances.[40–42] With only two measurements (baseline, follow-up) in the present study, we decided to calculate a simple volume difference for each patient (corresponding to a linear fit) and to determine the mean volume change in the cohort by averaging the patients’ results. We deliberately refrained from applying more sophisticated nonlinear models for fitting volume changes since they would only be sufficiently supported by more than two measurements in time. The MRIs of different patients in this study were acquired in various centres and on different scanners. These scanners may differ from each other by imaging quality, signal homogeneity and contrast. Although regular phantom scans have ascertained quality standards in both the tideglusib and davunetide studies, inter-scanner variability can be regarded as a disadvantage. However, in the davunetide dataset, no significant effect of the MRI scanner on longitudinal atrophy rates was identified.[6] Furthermore, recruitment at multiple sites and imaging at different scanners reflects the realistic scenario for future clinical trials and should be seen as advantage, since such multi-center analyses have better generalizability, the feasibility of which is demonstrated in the current work.

Although imaging-based outcome measures are not accepted as primary endpoints in phase III trials aimed at demonstration of clinical efficacy to improve the feeling, functioning and survival of patients, they might still be of high relevance in small scale proof-of-concept studies, as interim read-out in longer efficacy studies, or as secondary outcome measures in efficacy studies with a primary clinical read-out.[33]

In conclusion, we propose the 1-year change in the volumes of third ventricle, frontal lobe and midbrain as combined imaging read-out for clinical trials in PSP, requiring the least number of patients for detecting biological evidence for efficacy to slow down disease progression.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The Davunetide study was sponsored by Allon Therapeutics, Inc. (Vancouver, BC, CAN), a company that was purchased in August, 2013, by Paladin Laboratories. The Tauros study was sponsored by Noscira SA, Madrid, Spain. MRI datasets obtained from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (http://adni.loni.usc.edu) and from the International Consortium for Brain Mapping (ICBM) database (http://www.loni.usc.edu/ICBM) have been used to improve and test the methods of volumetric MRI analysis applied in this study. Dr. Höglinger was supported by the Deutsche Forschungsgemeinschaft (DFG, HO2402/6-2). Dr. Huppertz is supported by the Swiss Epilepsy Foundation. Dr. Boxer was supported by NIH grants R01AG038791, U54NS092089, Corticobasal Degeneration Solutions and the Tau Consortium. The authors are grateful to Volkmar Glauche, MSc, group leader at the Freiburg Brain Imaging Center, University Medical Center Freiburg, Germany, for his help with implementing the SPM12 algorithms.

References

1. Stamelou M, Höglinger G. A Review of Treatment Options for Progressive Supranuclear Palsy. CNS Drugs. 2016; 30(7):629–36. [PubMed: 27222018]
2. Boxer AL, Lang AE, Grossman M, et al. Davunetide in patients with progressive supranuclear palsy: a randomised, double-blind, placebo-controlled phase 2/3 trial. Lancet Neurol. 2014; 13(7): 676–685. [PubMed: 24873720]
3. Tolosa E, Litvan I, Hoglinger GU, et al. A phase 2 trial of the GSK-3 inhibitor tideglusib in progressive supranuclear palsy. Mov Disord. 2014; 29(4):470–478. [PubMed: 24532007]
4. Höglinger GU, Huppertz HJ, Wagenpfeil S, et al. Tideglusib reduces progression of brain atrophy in progressive supranuclear palsy in a randomized trial. Mov Disord. 2014; 29(4):479–487. [PubMed: 24488721]
5. Stamelou M, Schöpe J, Wagenpfeil S, del Ser T, Antonelou RCh, Oertel WH, Boxer AL, Höglinger GU. Power Calculations and Placebo Effect Estimation for Clinical Trials in Progressive Supranuclear Palsy. Movement Disorders. 2016; 31(5):742–7. [PubMed: 26948290]
6. Tsai RM, Lobach I, Bang J, Whitwell JL, Senjem ML, Jack CR Jr, Rosen H, Miller B, Boxer AL. AL-108-231 Investigators. Clinical correlates of longitudinal brain atrophy in progressive supranuclear palsy. Parkinsonism Relat Disord. 2016; 28:29–35. [PubMed: 27132501]
7. Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. Brain. 2007; 130(Pt 6):1552–1565. [PubMed: 17405767]
8. Dutt S, Binney RJ, Heuer HW, Luong P, Attygalle S, Bhatt P, Marx GA, Elofson J, Tartaglia MC, Litvan I, McGinnis SM, Dickerson BC, Kornak J, Waltzman D, Voltarelli L, Schuff N, Rabinovici GD, Kramer JH, Jack CR Jr, Miller BL, Rosen HJ, Boxer AL. AL-108-231 investigators. Progression of brain atrophy in PSP and CBS over 6 months and 1 year. Neurology. 2016; 87:2016–2025. [PubMed: 27742814]

9. Huppertz HJ, Kroll-Seger J, Kloppel S, Ganz RE, Kassubek J. Intra- and interscanner variability of automated voxel-based volumetry based on a 3D probabilistic atlas of human cerebral structures. Neuroimage. 2010; 49:2216–2224. [PubMed: 19878722]
10. Opfer R, Suppa P, Kepp T, Spies L, Schippling S, Huppertz HJ. Alzheimer's Disease Neuroimaging Initiative. Atlas based brain volumetry: How to distinguish regional volume changes due to biological or physiological effects from inherent noise of the methodology. Magn Reson Imaging. 2016; 34(4):455–61. [PubMed: 26723849]
11. Huppertz HJ, Möller L, Südmeyer M, Hilker R, Hattingen E, Egger K, Amtage F, Respondek G, Stamelou M, Schnitzler A, Pinkhardt EH, Oertel WH, Knake S, Kassubek J, Höglinger GU. Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. Mov Disord. 2016; doi: 10.1002/mds.26715
12. Litvan I, Agid Y, Calne D, 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; 47:1–9. [PubMed: 8710059]
13. Golbe LI. The Medical Advisory Board of the Society for Progressive Supranuclear Palsy. A clinical rating scale and staging system for progressive supranuclear palsy. Neurology. 1997; 48(Suppl):A326.
14. Schwab, RS., England, AC. Projection technique for evaluating surgery in Parkinson's disease. In: Gillingham, FJ., Donaldson, IML., editors. Third symposium on surgery in Parkinson's disease. Edinburgh; Livingstone: 1969. p. 152-157.
15. Guy, W. ECDEU Assessment Manual for Psychopharmacology. Rockville: Revised DHEW Pub. (ADM) MD: National Institute for Mental Health; 1976. Clinical Global Impressions; p. 218-222.
16. Ashburner J. A fast diffeomorphic image registration algorithm. Neuroimage. 2007; 38:95–113. [PubMed: 17761438]
17. Frazier JA, Chiu S, Breeze JL, et al. Structural brain magnetic resonance imaging of limbic and thalamic volumes in pediatric bipolar disorder. Am J Psychiatry. 2005; 162:1256–1265. [PubMed: 15994707]
18. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006; 31:968–980. [PubMed: 16530430]
19. Makris N, Goldstein JM, Kennedy D, et al. Decreased volume of left and total anterior insular lobule in schizophrenia. Schizophr Res. 2006; 83:155–171. [PubMed: 16448806]
20. Goldstein JM, Seidman LJ, Makris N, et al. Hypothalamic abnormalities in schizophrenia: sex effects and genetic vulnerability. Biol Psychiatry. 2007; 61:935–945. [PubMed: 17046727]
21. Hammers A, Allom R, Koepp MJ, Free SL, Myers R, Lemieux L, Mitchell TN, Brooks DJ, Duncan JS. Three-dimensional maximum probability atlas of the human brain, with particular reference to the temporal lobe. Hum Brain Mapp. 2003; 19:224–47. [PubMed: 12874777]
22. Shattuck DW, Mirza M, Adisetiyo V, et al. Construction of a 3D probabilistic atlas of human cortical structures. Neuroimage. 2008; 39:1064–1080. [PubMed: 18037310]
23. Malone IB, Leung KK, Clegg S, et al. Accurate automatic estimation of total intracranial volume: a nuisance variable with less nuisance. Neuroimage. 2015; 104:366–372. [PubMed: 25255942]
24. Oba H, Yagishita A, Terada H, et al. New and reliable MRI diagnosis for progressive supranuclear palsy. Neurology. 2005; 64:2050–5. [PubMed: 15985570]
25. Möller L, Huppertz HJ, Südmeyer M, et al. Manual MRI Morphometry in Parkinsonian Syndromes. Movement Disorders. in press.
26. Whitwell JL, Crum WR, Watt HC, Fox NC. Normalization of cerebral volumes by use of intracranial volume: implications for longitudinal quantitative MR imaging. AJNR Am J Neuroradiol. 2001; 22:1483–9. [PubMed: 11559495]
27. Holm S. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics. 1979; 6:65–70.
28. R Core Team. A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna, Austria: https://www.r-project.org

29. Lehmann, E. Nonparametrics: Statistical methods based on ranks. New York, United States of America: Springer; 2006.
30. 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; 254:478–81. [PubMed: 17401522]
31. Whitwell JL, Xu J, Mandrekar JN, et al. Rates of brain atrophy and clinical decline over 6 and 12-month intervals in PSP: determining sample size for treatment trials. Parkinsonism Relat Disord. 2012; 18(3):252–256. [PubMed: 22079523]
32. Cash DM, Frost C, Iheme LO, Ünay D, Kandemir M, Fripp J, Salvado O, Bourgeat P, Reuter M, Fischl B, Lorenzi M, Frisoni GB, Pennec X, Pierson RK, Gunter JL, Senjem ML, Jack CR Jr, Guizard N, Fonov VS, Collins DL, Modat M, Cardoso MJ, Leung KK, Wang H1, Das SR, Yushkevich PA, Malone IB, Fox NC, Schott JM, Ourselin S. Assessing atrophy measurement techniques in dementia: Results from the MIRIAD atrophy challenge. Neuroimage. 2015; 123:149–64. [PubMed: 26275383]
33. Hobbs NZ, Farmer RE, Rees EM, Cole JH, Haider S, Malone IB, Sprengelmeyer R, Johnson H, Mueller HP, Sussmuth SD, Roos RA, Durr A, Frost C, Scahill RI, Landwehrmeyer B, Tabrizi SJ. Short-interval observational data to inform clinical trial design in Huntington's disease. J Neurol Neurosurg Psychiatry. 2015; 86(12):1291–8. [PubMed: 25669748]
34. Majid DS, Aron AR, Thompson W, Sheldon S, Hamza S, Stoffers D, Holland D, Goldstein J, Corey-Bloom J, Dale AM. Basal ganglia atrophy in prodromal Huntington's disease is detectable over one year using automated segmentation. Mov Disord. 2011; 26(14):2544–51. [PubMed: 21932302]
35. van den Elskamp IJ, Boden B, Dattola V, Knol DL, Filippi M, Kappos L, Fazekas F, Wagner K, Pohl C, Sandbrink R, Polman CH, Uitdehaag BM, Barkhof F. Cerebral atrophy as outcome measure in short-term phase 2 clinical trials in multiple sclerosis. Neuroradiology. 2010; 52(10): 875–81. [PubMed: 20049424]
36. Tabrizi SJ, Reilmann R, Roos RA, Durr A, Leavitt B, Owen G, Jones R, Johnson H, Craufurd D, Hicks SL, Kennard C, Landwehrmeyer B, Stout JC, Borowsky B, Scahill RI, Frost C, Langbehn DR. TRACK-HD investigators. Potential endpoints for clinical trials in premanifest and early Huntington's disease in the TRACK-HD study: analysis of 24 month observational data. Lancet Neurol. 2012; 11(1):42–53. [PubMed: 22137354]
37. Jones R, Stout JC, Labuschagne I, Say M, Justo D, Coleman A, Dumas EM, Hart E, Owen G, Durr A, Leavitt BR, Roos R, O'Regan A, Langbehn D, Tabrizi SJ, Frost C. The potential of composite cognitive scores for tracking progression in Huntington's disease. J Huntingtons Dis. 2014; 3(2): 197–207. [PubMed: 25062862]
38. Höglinger GU, Respondek G, Stamelou M, Kurz C, Josephs KA, Lang AE, et al. Clinical Diagnosis of Progressive Supranuclear Palsy - The Movement Disorder Society Criteria. Mov Disord. In press.
39. Opfer R, Suppa P, Kepp T, Spies L, Schippling S, Huppertz HJ. Alzheimer's Disease Neuroimaging Initiative. Atlas based brain volumetry: How to distinguish regional volume changes due to biological or physiological effects from inherent noise of the methodology. Magn Reson Imaging. 2016; 34:455–61. [PubMed: 26723849]
40. Fjell AM, Walhovd KB, Westlye LT, et al. When does brain aging accelerate? Dangers of quadratic fits in cross-sectional studies. Neuroimage. 2010; 50:1376–83. [PubMed: 20109562]
41. Fjell AM, Westlye LT, Grydeland H, et al. Critical ages in the life course of the adult brain: nonlinear subcortical aging. Neurobiol Aging. 2013; 34:2239–47. [PubMed: 23643484]
42. Schippling S, Ostwaldt AC, Suppa P, et al. Global and regional annual brain volume loss rates in physiological aging. J Neurol. 2017; doi: 10.1007/s00415-016-8374-y

Appendix

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, Günter Höglinger, Axel Lipp, Heinz Reichmann, Dirk Woitalla

United Kingdom: Dennis Chan, Adam Zermansky, David Burn, Andrew Lees

United States: Adam Boxer, Bruce L. Miller, Iryna V. Lobach, (Memory and Aging Center, Department of Neurology, University of California, San Francisco, CA, USA), Erik Roberson, Lawrence Honig, Edward Zamrini, Rajesh Pahwa, Yvette Bor-delon, 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 Jancovic, 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, Iryna V. Lobach, Illana Gozes, Steve Whitaker, Joe Hirman, Michael Gold, Bruce H. Morimoto

Tau Restoration on PSP (TAUROS) MRI Investigators

J.C. Gómez, MD, B. Tijero, MD; R. Villoria, MD (H. de Cruces. Barakaldo, Site Investigator); J. García de Yebenes, MD, J.L. Lopez Sendón, MD (H. Ramón y Cajal. Madrid, Site Investigators); E. Tolosa, MD, M.T. Buongiorno, MD, N. Bargalló, MD (H. Clinic. Barcelona, Site Investigators); J.A. Burguera, MD, I. Martinez, MD (H. La Fe. Valencia, Site Investigators); J. Ruiz-Martínez, MD, J. Villanua, MD (H. Donostia. San Sebastián, Site Investigator); F. Vivancos, MD, I. Ybot, MD (H. La Paz. Madrid, Site Investigators); M. Aguilar, MD, J.L. Dolz, MD (H. Mutua Terrassa. Terrassa, Site Investigator); M. Boada, MD, A. Lafuente, MD, M.A. Tejero, MD (Fundación ACE: Barcelona, Site Investigators); J.J. López-Lozano, MD, M. Mata, MD (H. Puerta de Hierro. Madrid, Site Investigators); A. Kupsch, MD, A. Lipp, MD (Virchow-Klinikum. Berlin, Site Investigators); M. Höllerhage, MD, W.H. Oertel, MD, G. Respondek, MD, M. Stamelou, MD, S. Knake, MD (Universitätsklinikum. Marburg, Site Investigators); D. Berg, MD, W. Maetzler, MD, K.K. Srulijes, MD, A. Gröger, MD (Universitätsklinikum. Tübingen, Site Investigators); A. Ludolph, MD, J. Kassubek, MD (Universitätsklinikum. Ulm, Site Investigators); M. Steiger, MD, K. Tyler, MD (Walton Center. Liverpool, Site Investigators); D.J. Burn, MD, L. Morris, MD (Clinical Ageing Research Unit. Newcastle upon Tyne, Site Investigator); A. Lees, MD, H. Ling, MD, L. Strycharczuk, MD (Reta Lila Weston Institute. London, Site Investigators).

The MDS endorsed PSP study group

Adam L Boxer, Alexander Pantelyat, Angelo Antonini, Anthony E Lang, Carlo Colosimo, Caroline Kurz, Christer Nilsson, Gesine Respondek, Günter U Höglinger, Huw R Morris,

Irene Litvan, James B Rowe, Jan Kassubek, Jean-Christophe Corvol, Jennifer L Whitwell, John van Swieten, Kailash P Bhatia, Keith A Josephs, Lawrence I Golbe, Maria Stamelou, Richard Dodel, Stefan Lorenzl, Thilo van Eimeren, Wolfgang H Oertel.

Author Roles

1. Research Project: A. Conception, B. Organization, C. Execution; 2. Statistical Analysis: A. Design, B. Execution, C. Review and Critique; 3. Manuscript Preparation: A. Writing the First Draft, B. Review and Critique.

Günter U. Höglinger: 1A, 1B, 1C, 2A, 2B, 3A; ***Jakob Schöpe***: 2A, 2B, 3B; ***Maria Stamelou***: 1A, 2C, 3B; ***Jan Kassubek***: 1A, 2C, 3B; ***Teodoro del Ser***: 1A, 1B, 1C, 2C, 3B; ***Adam Boxer***: 1A, 1B, 1C, 2C, 3B; ***Stefan Wagenpfeil***: 2A, 2B, 3B; ***Hans-Jürgen Huppertz***: 1A, 1B, 1C, 2A, 2B, 3A.

Financial Disclosures of all authors for the preceding 12 months

Günter U. Höglinger has served on the advisory boards for AbbVie, Alzprotect, Asceneuron, Bristol-Myers Squibb, Novartis, Roche, Sellas Life Sciences Group, UCB; has received honoraria for scientific presentations from Abbvie, Roche, Teva, UCB, has received research support from CurePSP, the German Academic Exchange Service (DAAD), German Parkinson's Disease Foundation (DPG), German PSP Association (PSP Gesellschaft), German Research Foundation (DFG) and the German Ministry of Education and Research (BMBF), International Parkinson's Fonds (IPF), the Sellas Life Sciences Group; has received institutional support from the German Center for Neurodegenerative Diseases (DZNE). ***Jakob Schöpe*** has nothing to declare. ***Maria Stamelou*** has served on the editorial board of Movement Disorders Journal and Frontiers in Movement Disorders; has received speaker and travel honoraria from Actelion and Abbvie Pharmaceuticals. ***Jan Kassubek*** received consulting fees as an advisory board member and honoraria as a speaker from UCB Pharma, Teva Pharmaceuticals, Zambon, Medtronic, Desitin, AbbVie, Boehringer Ingelheim, GlaxoSmithKline, Merz Pharmaceuticals and Hoffmann-La Roche. ***Teodoro del Ser*** was Medical Director of Noscira S.A. since 2004 to 2013 and holds ordinary shares of Zeltia, the major shareholder of Noscira SA, Madrid, Spain. Previously served on scientific advisory boards of Bayer, Novartis and Pfizer Inc, and received speaker honoraria from Pfizer Inc, Novartis, Eisai Inc. and Janssen, and research support from Janssen. ***Adam Boxer*** is supported by National Institutes of Health grants U54NS092089, R01AG038791, U01AG045390, U01AG052943-041, U19AG010483, the University of California, the Bluefield Project to Cure FTD and the Tau Research Consortium; received research support from Avid, Biogen, Bristol Myers Squibb, C2N Diagnostics, Cortice Biosciences, Eli Lilly, Forum Pharmaceuticals, Genentech, Roche and TauRx; has served as a consultant for Abbvie, Asceneuron, Ipierian, Isis Pharmaceuticals, Janssen, Merck and Novartis; serves on a Data and Safety Monitoring Board for Neurogenetics Pharmaceuticals; has stock and/or options in Alector and Delos. ***Stefan Wagenpfeil*** has served on the advisory board for the NeoVita Trial; has served as a consultant for Apogepha GmbH; has served for expert testimony for "Endoscopy" and "European Heart Journal"; has received institutional support from University of Tromsø, Munich Business School and Münchner Studienzentrum. ***Hans-***

Jürgen Huppertz has served on the advisory board for Novartis. His research was supported by the Swiss National Foundation and is continuously supported by the Swiss Epilepsy Foundation.

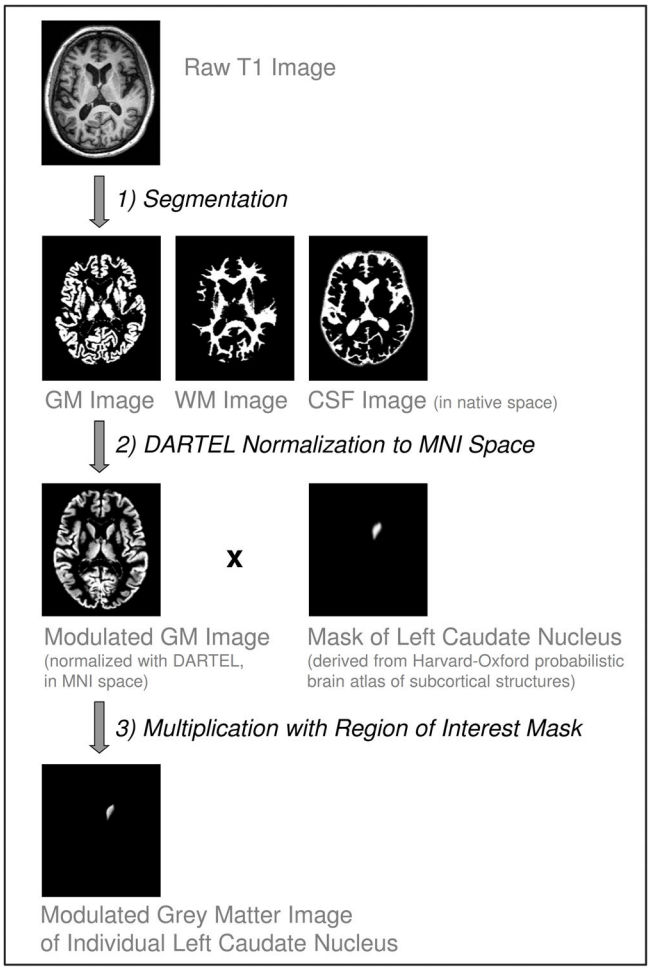

Figure 1.

Image processing and volume determination shown exemplarily for the left caudate nucleus: 1) Segmentation of gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) compartments by applying the 'unified segmentation' algorithm of SPM12 to a T1-weighted 3D image. 2) Normalization of the resulting GM image in native space using diffeomorphic anatomical registration through exponentiated Lie algebra (DARTEL) [16] with predefined templates in Montreal Neurological Institute (MNI) space. 3) Multiplication of the modulated GM image derived by DARTEL normalization with a mask of the left caudate nucleus. In this example the masking image is derived from the Harvard-Oxford probabilistic brain atlas of subcortical structures. [18–21] The voxel-wise multiplication results in a modulated GM image with the caudate nucleus isolated. Due to 'modulation' of the grey matter image, the effect of normalization (i.e. extension or shrinkage of the investigated structure) is compensated for so that the sum of the residual voxels in the final image represents the volume of the original structure in native space. For volume measurements of white matter structures or CSF compartments the same image processing steps are based on normalized and modulated white matter or CSF images, respectively.

Table 1

Demographic data of healthy controls and PSP patients, and one-year progression of patients in clinical rating scales.

	Controls	PSP				
		Baseline	After one year	One-year difference	Effect size	Sample size[#]
N=	50	99[§]	99[§]	n.a.	n.a.	n.a.
Gender (F/M)	20/30	53/46	53/46	n.a.	n.a.	n.a.
Age (Yrs.)	67.3 ± 5.2	67.9 ± 6.9	69.0 ± 6.9	1.1 ± 0.1	n.a.	n.a.
SEADL	n.a.	57.2 ± 22.9	40.6 ± 22.8	−17.1 ± 18.3 ***/***	−0.94	73 (84)
PSPRS						
Total score	n.a.	37.8 ± 10.9	47.1 ± 12.7	9.5 ± 9.1 ***/***	1.05	58 (67)
Bulbar score	n.a.	2.6 ± 1.5	3.5 ± 1.7	0.8 ± 1.3 ***/***	0.68	139 (161)
Gait score	n.a.	9.8 ± 3.9	12.7 ± 3.8	3.0 ± 3.2 ***/***	0.96	70 (81)
History score	n.a.	7.9 ± 3.3	10.0 ± 3.9	2.1 ± 3.2 ***/***	0.67	143 (165)
Limb score	n.a.	4.7 ± 2.1	5.7 ± 2.8	1.1 ± 2.1 ***/***	0.51	239 (276)
Mentation score	n.a.	3.6 ± 2.7	4.4 ± 3.1	0.7 ± 2.7 **/n.s.	0.31	826 (955)
Ocular score	n.a.	9.1 ± 2.9	10.9 ± 2.9	1.8 ± 2.5 ***/***	0.69	133 (154)
CGIDS	n.a.	3.8 ± 0.9	4.7 ± 1.0	0.9 ± 0.9 ***/***	0.97	69 (80)

Data are given as mean ± standard deviation.

Abbreviations: SEADL: Schwab and England Activities of Daily Living Scale; PSPRS: Progressive Supranuclear Palsy Rating Scale; CGIDS: Clinical Global Impression of Disease Severity. n.a. = not applicable.

Age and gender distributions between PSP-patients and controls were not statistically significant different (p>0.05).

[#]minimum number per group; based on a two-sample t-test with significance level of 5% and a power of 80% to detect a 50% change in progression; approximation of the sample size for the Mann-Whitney U test in parentheses. Sample sizes are given before adjusting for drop-out rate. To adjust for a drop-out rate of 26% (e.g. the combined drop-out rate in both trials), the following formula should be used: sample size/0.74 (e.g. 70/0.74 = 95).

[§]N=90 patients from the davunetide study and N=9 from the tideglusib study.

P-values are shown uncorrected/corrected for multiple testing based on Holm's method; n.s., non significant;

** P<0.01,

*** P<0.001; paired two-sample t-test baseline vs. 52 weeks.

Table 2

Brain atrophy, annualized volume changes, effect sizes and sample sizes in N=99 PSP patients with excellent imaging quality.

Structures/Compartments	Con-trols	Patients baseline	Patients follow-up	Mean differences to controls at baseline	Baseline z scores	Mean annualized volume changes*	SD of annualized volume changes*	Standardized effect sizes	Two-sample t-test	Mann-Whitney U test	Two-sample t-test	Mann-Whitney U test	Two-sample t-test	Mann-Whitney U test
Change due to treatment									20%		30%		50%	
Brain	1037.5	1003.9	986.1	−3.2%	−0.6	−2.0%	1.8%	−1.10	325	376	145	168	53	62
Gray Matter	587.1	600.4	586.9	2.3%	0.3	−2.4%	3.4%	−0.70	813	940	362	419	131	152
White Matter	450.3	403.5	399.1	−10.4%	−1.3	−1.2%	3.7%	−0.34	3421	3959	1521	1760	549	635
CSF	367.7	401.3	418.9	9.1%	0.6	4.0%	5.4%	0.73	732	848	326	378	118	137
Intracranial Volume	1405	1405	1405	0.0%	-	-	-	-	-	-	-	-	-	-
Cerebrum	886.9	861.1	845.9	−2.9%	−0.5	−2.0%	1.8%	−1.08	337	390	151	174	55	64
Cerebrum GM	486.7	500.2	488.7	2.8%	0.3	−2.4%	3.6%	−0.67	874	1011	389	450	141	163
Cerebrum WM	400.2	360.9	357.2	−9.8%	−1.2	−1.2%	3.8%	−0.32	3884	4495	1727	1999	623	721
Frontal lobe	291.9	274.0	267.0	−6.1%	−0.8	−2.7%	2.2%	−1.22	263	305	118	136	43	50
Frontal lobe GM	159.0	157.9	153.1	−0.7%	−0.1	−3.1%	3.7%	−0.83	570	659	254	294	92	107
Frontal lobe WM	133.0	116.1	113.9	−12.6%	−1.3	−2.1%	4.4%	−0.47	1792	2074	797	922	288	333
Temporal lobe	179.2	177.0	174.8	−1.2%	−0.2	−1.5%	2.9%	−0.51	1490	1724	663	767	240	277
Temporal lobe GM	122.0	123.8	121.7	1.5%	0.1	−1.9%	4.3%	−0.44	2064	2389	918	1063	331	384
Temporal lobe WM	57.2	53.2	53.1	−6.9%	−0.8	−0.5%	4.4%	−0.10	36103	41786	16047	18572	5778	6687
Parietal lobe	163.3	161.1	158.5	−1.4%	−0.2	−1.8%	2.2%	−0.82	586	678	261	302	95	110
Parietal lobe GM	88.2	93.3	91.1	5.7%	0.6	−2.4%	3.9%	−0.61	1039	1203	463	536	168	194
Parietal lobe WM	75.1	67.8	67.4	−9.7%	−0.9	−0.7%	4.9%	−0.15	17159	19860	7627	8828	2747	3179
Occipital lobe	116.2	118.2	116.7	1.8%	0.2	−1.5%	2.5%	−0.60	1105	1279	492	569	178	206
Occipital lobe GM	66.0	71.4	70.2	8.2%	0.9	−1.9%	4.5%	−0.41	2286	2646	1017	1177	367	425
Occipital lobe WM	50.2	46.9	46.6	−6.6%	−0.5	−0.6%	5.8%	−0.11	33304	38546	14803	17133	5330	6169
Insula	15.7	15.7	15.4	0.2%	0.0	−2.1%	3.6%	−0.59	1134	1312	505	584	183	211
Cerebellum	111.7	108.2	106.4	−3.2%	−0.4	−1.8%	3.1%	−0.58	1150	1331	512	593	185	214
Cerebellum GM	88.9	88.8	87.1	−0.2%	0.0	−2.1%	3.7%	−0.56	1272	1472	566	655	205	237
Cerebellum WM	22.8	19.4	19.3	−14.9%	−1.3	−0.4%	7.0%	−0.06	115355	133513	51270	59340	18458	21363

Structures/Compartments	Con-trols	Patients baseline	Patients follow-up	Mean differences to controls at baseline	Baseline z scores	Mean annualized volume changes*	SD of annualized volume changes*	Standardized effect sizes	Two-sample t-test	Mann-Whitney U test	Two-sample t-test	Mann-Whitney U test	Two-sample t-test	Mann-Whitney U test
Midbrain	10.2	8.5	8.3	−16.4%	−2.7	−2.6%	1.9%	−1.33	223	258	100	115	37	43
Pons	15.5	14.0	13.7	−9.8%	−1.1	−2.5%	2.2%	−1.12	312	361	139	161	51	59
Medulla oblongata	4.6	4.3	4.3	−5.3%	−0.8	−2.0%	3.0%	−0.66	898	1039	400	463	145	168
Corpus callosum plane	582.5	525.2	516.6	−9.8%	−0.7	−1.8%	3.5%	−0.51	1484	1717	660	764	239	276
Cerebellar vermis plane	983.5	978.5	966.2	−0.5%	−0.1	−1.3%	3.0%	−0.45	1908	2209	849	983	307	355
Midbrain plane	281.6	245.3	235.8	−12.9%	−1.6	−3.8%	5.2%	−0.73	745	862	332	384	120	139
Midbrain tegmentum plane	168.6	133.8	128.9	−20.7%	−3.1	−3.6%	4.7%	−0.77	658	761	293	339	106	123
Pons plane	507.9	449.4	439.5	−11.5%	−1.3	−2.3%	3.1%	−0.73	745	862	332	384	120	139
Pons pars basilaris plane	347.5	310.4	304.0	−10.7%	−1.2	−2.2%	2.6%	−0.86	537	621	239	277	87	101
Medulla plane	270.5	251.9	249.6	−6.9%	−0.8	−1.0%	3.9%	−0.25	6143	7110	2731	3161	984	1139
Superior cerebellar peduncle	1.3	1.1	1.1	−13.6%	−1.9	−2.4%	2.7%	−0.89	493	571	220	255	80	93
Middle cerebellar peduncle	11.0	10.0	9.8	−9.1%	−1.1	−1.9%	2.9%	−0.64	950	1099	423	489	153	177
Third ventricle	1.3	1.7	1.8	24.8%	0.8	7.6%	5.3%	1.42	195	226	88	101	32	38
Lateral ventricles	24.4	32.7	36.0	33.9%	0.8	9.0%	7.0%	1.29	238	275	106	123	39	45
Hippocampus	6.1	6.2	6.0	0.7%	0.1	−2.2%	3.5%	−0.62	1039	1202	463	535	167	194
Amygdala	3.3	3.4	3.3	2.3%	0.2	−2.0%	5.3%	−0.38	2707	3133	1204	1393	434	503
Caudate	3.9	3.9	3.8	−0.6%	0.0	−3.0%	6.4%	−0.47	1787	2068	795	920	287	332
Putamen	5.4	5.6	5.4	3.5%	0.2	−3.9%	6.8%	−0.58	1178	1364	524	607	190	220
Accumbens	0.9	0.9	0.9	−0.3%	0.0	−3.4%	3.8%	−0.89	497	575	222	256	81	93
Pallidum	3.8	3.2	3.1	−15.8%	−2.6	−1.8%	3.7%	−0.48	1737	2010	773	894	279	323
Thalamus	11.2	10.8	10.5	−3.2%	−0.4	−2.8%	3.3%	−0.84	551	638	246	284	89	103
Combined: Midbrain + Frontal Lobe						−5.3%	3.3%	−1.60	156	180	70	81	26	30
Combined: Midbrain − Third ventricle						−10.2%	5.8%	−1.75	130	151	59	68	22	25
Combined: Frontal lobe − Third ventricle						−10.3%	6.5%	−1.59	156	180	70	81	26	30
Combined: Midbrain + Frontal Lobe − Third ventricle						−12.9%	7.1%	−1.83	119	137	54	62	20	23

Volume results are shown in mL, results for planes in mm^2. All cross-sectional results have been normalized to the mean intracranial volume (ICV) of controls. A 3-color-scale was used to rank the volume differences to controls, the volume changes between baseline and follow-up, and the standardized effect sizes, from shades of red (lower numbers or volume loss) over white to shades of blue (higher numbers or volume gain).

The last six columns display the required number of patients for the verum and placebo group each in a medication study when a treatment effect of 20%, 30% or 50%, respectively, is expected, i.e. when the atrophy due to the natural course of the disease is reduced by these amounts. For normal distributions of volume results the unpaired two-sample t-test can be applied, otherwise the Mann-Whitney U test would be appropriate. A 2-color-scale was used to rank the results in these six columns from higher (white) to lower numbers (shades of green).

* For combined structures the annualized percentage volume changes of the chosen structures have been summed up for each individual patient, with negative signs for the third ventricle since the enlargement of the CSF space runs counter the atrophy of the brain parenchyma. Subsequently, mean and standard deviations were calculated across the results of all investigated patients.

Table 3

Correlation between 1-year changes in clinical rating scales and annualized brain volume changes.

	SEADL	CGIDS	PSPRS-total
Third Ventricle	r = −0.35 ***/*	r = 0.25 */n.s.	r = 0.24 */n.s.
Frontal Lobe	r = 0.23 */n.s.	r = 0.09n.s./n.s.	r = −0.21 */n.s.
Midbrain	r = 0.26 */n.s.	r = 0.22 */n.s.	r = −0.31 **/*
Combined: Midbrain + Frontal Lobe − Third Ventricle	r = 0.42 ***/*	r = −0.28 **/*	r = −0.34 **/*

Pearson's r was calculated on the basis of N=95 cases with complete availability of all clinical and MRI data (in N=4 of the total N=99 cases individual clinical data points were missing); P-values are shown uncorrected/corrected for multiple testing based on Holm's method; n.s., non significant;

* P<0.05,

** P<0.01,

*** P<0.001.