OPEN ACCESS

Original research

Added value of FDG-PET for detection of progressive supranuclear palsy

Ralph Buchert ,[1] Hans-Jürgen Huppertz ,[2] Florian Wegner ,[3] Georg Berding ,[4] Matthias Brendel,[5,6,7] Ivayla Apostolova ,[1] Carsten Buhmann ,[8] Monika Poetter-Nerger ,[8] Alexander Dierks,[9] Sabrina Katzdobler ,[6,7,10] Martin Klietz ,[3] Johannes Levin ,[6,7,10] Nima Mahmoudi ,[11] Andreas Rinscheid ,[12] Andrea Quattrone ,[10,13] Sophia Rogozinski,[3] Jost-Julian Rumpf ,[14] Christine Schneider ,[15] Sophia Stoecklein,[16] Phoebe G Spetsieris ,[17] David Eidelberg ,[17] Osama Sabri,[18] Henryk Barthel,[18] Mike P Wattjes ,[11,19] Günter Höglinger ,[3,6,7,10] On behalf of the Alzheimer's Disease Neuroimaging Initiative

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jnnp-2024-333590).

For numbered affiliations see end of article.

Correspondence to
Dr Ralph Buchert; r.buchert@uke.de

RB, H-JH and FW are joint first authors.
OS, HB, MPW and GH are joint senior authors.

Received 8 February 2024
Accepted 17 July 2024
Published Online First 6 August 2024

To cite: Buchert R, Huppertz H-J, Wegner F, *et al*. *J Neurol Neurosurg Psychiatry* 2025;**96**:287–295.

ABSTRACT

Background Diagnostic criteria for progressive supranuclear palsy (PSP) include midbrain atrophy in MRI and hypometabolism in [^{18}F]fluorodeoxyglucose (FDG)-positron emission tomography (PET) as supportive features. Due to limited data regarding their relative and sequential value, there is no recommendation for an algorithm to combine both modalities to increase diagnostic accuracy. This study evaluated the added value of sequential imaging using state-of-the-art methods to analyse the images regarding PSP features.

Methods The retrospective study included 41 PSP patients, 21 with Richardson's syndrome (PSP-RS), 20 with variant PSP phenotypes (vPSP) and 46 sex- and age-matched healthy controls. A pretrained support vector machine (SVM) for the classification of atrophy profiles from automatic MRI volumetry was used to analyse T1w-MRI (output: MRI-SVM-PSP score). Covariance pattern analysis was applied to compute the expression of a predefined PSP-related pattern in FDG-PET (output: PET-PSPRP expression score).

Results The area under the receiver operating characteristic curve for the detection of PSP did not differ between MRI-SVM-PSP and PET-PSPRP expression score ($p \geq 0.63$): about 0.90, 0.95 and 0.85 for detection of all PSP, PSP-RS and vPSP. The MRI-SVM-PSP score achieved about 13% higher specificity and about 15% lower sensitivity than the PET-PSPRP expression score. Decision tree models selected the MRI-SVM-PSP score for the first branching and the PET-PSPRP expression score for a second split of the subgroup with normal MRI-SVM-PSP score, both in the whole sample and when restricted to PSP-RS or vPSP.

Conclusions FDG-PET provides added value for PSP-suspected patients with normal/inconclusive T1w-MRI, regardless of PSP phenotype and the methods to analyse the images for PSP-typical features.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Current diagnostic criteria list midbrain atrophy on structural MRI and hypometabolism on brain [^{18}F]fluorodeoxyglucose (FDG)-positron emission tomography (PET) as supportive progressive supranuclear palsy (PSP) features, but due to the lack of data, they do not make any recommendations on the sequential use of both imaging modalities in patients where the clinical status and one imaging modality remain inconclusive. This contributes to the lack of consensus on how neuroimaging should be used in the diagnosis of PSP.

WHAT THIS STUDY ADDS

⇒ This study indicates that brain FDG-PET is about 15% more sensitive for the detection of PSP than structural MRI and, therefore, might be clinically useful for the detection of PSP in patients with normal or inconclusive structural MRI, irrespective of the PSP predominance type.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ These findings can guide decision-making on sequential MRI and FDG-PET imaging in clinically unclear patients with suspected PSP.

INTRODUCTION

Progressive supranuclear palsy (PSP) is a primary 4-repeat tauopathy[1 2] presenting with a broad phenotypic spectrum including the 'classical' Richardson syndrome (PSP-RS) and various other variants (vPSP).[3 4] Despite the fact that accounting for the large phenotypic variability in the revised criteria of the PSP study group of the International Parkinson and Movement Disorder Society (MDS) improved the PSP diagnosis, particularly at early disease stages and for vPSP,[3 5] there is a need for supportive features to assist the clinical diagnosis in uncertain cases.[6 7] The MDS criteria list midbrain atrophy on structural MRI and hypometabolism on [^{18}F]

fluorodeoxyglucose (FDG) positron emission tomography (PET) of the brain as supportive PSP features.[3 8]

There are numerous studies on either structural MRI or FDG-PET alone for the (differential) diagnosis of PSP.[8–19] However, there is a striking lack of head-to-head comparisons of both modalities in the same patients. In particular, there are no studies on sequential administration of both modalities in patients where the clinical status and one imaging modality remain inconclusive. This contributes to the lack of consensus on how neuroimaging should be used in the diagnosis of parkinsonian syndromes.[16]

Structural MRI and FDG-PET can be assessed visually for descriptive PSP features/signs. This can be supported by quantitative and/or statistical analyses in predefined regions-of-interest or voxel-by-voxel. However, recent studies demonstrated that fully automatic multivariable analyses using a support vector machine (SVM) to classify volumetric MRI profiles and covariance analysis of FDG uptake patterns clearly outperform visual analyses (with and without support by univariate quantitative/statistical analyses) in the detection of PSP, most pronounced in vPSP.[10 18]

The current multicentre study performed a head-to-head comparison between SVM classification of MRI atrophy profiles and covariance pattern analysis of FDG-PET in the same clinical sample including about 50% vPSP. The a priori hypothesis was that covariance pattern analysis of FDG-PET is more sensitive for the detection of PSP than SVM classification of MRI atrophy profiles so that additional FDG-PET is most useful in case of normal or inconclusive MRI.

METHODS

Subjects

This retrospective multicentre study was designed as phase 2 study ('ability to discriminate patients from controls') according to the five-phase framework for biomarker validtion[20] and, therefore, included well-characterised patients with established PSP diagnosis and healthy controls (HCs).

The archives of the university hospitals of Augsburg, Hamburg, Hannover, Leipzig and Munich were searched for consecutive patients using the following inclusion criteria: (1) clinical PSP diagnosis according to the MDS criteria by a movement disorder specialist and (2) structural MRI and FDG-PET performed within 12 months. 63 patients fulfilling these criteria were identified. From these, 18 patients (28.6%) were excluded because the T1w sequence was not adequate for morphometric/volumetric analyses, 4 patients (6.3%) because of severe ischaemic small vessel disease (Fazekas grade 3).[21] The remaining 41 patients were included in the analyses (age at PET 69.6±7.8 years, range 53–81 years, 49% females). Mean disease duration at PET was 3.1±2.6 years (range 0.4–12.2 years). Median and IQR of the delay between the date of the clinical reference diagnosis and FDG-PET were 0 day and −6 days to 6 days, respectively (range −210 to 441 days).

Diagnostic certainty was probable/possible/suggestive of PSP in 35 (85.4%)/1 (2.4%)/5 (12.2%) patients. PSP-RS was the most frequent predominance type: n=21 (51.2%). Among the 20 patients with vPSP, 11 presented with a 'cortical' syndrome[22]: 10 with predominant corticobasal syndrome (24.4% of all PSP patients) and 1 with predominant frontal presentation (2.4%). The remaining nine vPSP patients presented with a 'subcortical' syndrome: eight with predominant parkinsonism (19.5% of all PSP patients) and one with progressive gait freezing (2.4%). The patients scored 30±14 (range 9–86) on the PSP rating scale[23] and 21.4±4.4 (n=32, range 10–29) on the Montreal Cognitive Assessment scale. Depressive symptoms scored 5.7±2.6 (n=28, range 1–12) on the 15-point Geriatric Depression Scale.

46 HC subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI-1) were included (age at PET 72.2±4.7 years, range 59–79 years, 39% females). The HC subjects did not differ significantly from the PSP patients with respect to age (p=0.07) or sex (p=0.37). The ADNI was launched in 2003 as a public–private partnership, led by principal investigator Michael W Weiner. For up-to-date information, see www.adni-info.org.

The same subjects had been included in previous studies on imaging-based detection of PSP with T1w-MRI[18] or FDG-PET.[10] These previous studies compared different approaches to identify PSP features separately in MRI and FDG-PET. In T1w-MRI, fully automatic classification of atrophy profiles by an SVM clearly outperformed descriptive signs (hummingbird, morning glory, Mickey-Mouse), visual reading supported by manual and automatic planimetric measures as well as visual reading supported by automatic volumetry.[18] In FDG-PET, automatic covariance pattern analysis clearly outperformed visual interpretation of FDG-PET supported by voxel-based statistical testing.[10] The current study first compared optimal (SVM) evaluation of T1w-MRI and optimal (covariance expression) evaluation of FDG-PET for the detection of PSP. Then, the putative added value of a sequential administration of these modalities was evaluated.

MRI

MRI was performed at 3/1.5 Tesla in 28/13 PSP patients (68.3/31.7%) with 15 different MR systems of 3 manufacturers (Siemens, Philips, Toshiba).[18]

For the ADNI HC subjects, the first of the two back-to-back 3D T1w-MRI from the baseline session was used.[24]

SVM classification of T1W-MRI

For the automatic classification of T1w-MRI as PSP or HC, an SVM was used that had been trained previously using an independent dataset consisting of 106 PSP patients and 73 HCs.[11] The input to the SVM comprised the results of automated atlas-based MRI volumetry for 44 different compartments, brain structures, and planes, all corrected for intracranial volume and age and scaled to a range of 0–1 (online supplemental figure 1A). The output of the SVM consisted of an MRI-SVM-PSP score ranging between 0 (most likely HC) and 1 (most likely PSP). The pretrained SVM of 2016 was used without any changes.

FDG-PET imaging

In the PSP patients, brain FDG-PET was performed with six different PET systems (five PET/CT and one PET/MRI) from two manufacturers (Siemens, Philips) after intravenous injection of 202±37 MBq (range 128–317 MBq) according to the local standard operating procedure at each site. PET and MRI had been acquired simultaneously with a PET/MRI system in five patients. In the remaining 36 patients, median and IQR of the delay (absolute value) between PET and MRI were 7 days and 2–99 days, respectively. MRI had been performed before PET in 21 of the 36 patients (58.3%). The PET images were harmonised with respect to spatial resolution (8 mm full-width-at-half-maximum) as described previously.[10]

For the ADNI HC subjects, the baseline FDG-PET was used, fully preprocessed by the ADNI imaging core lab including filtering to 8 mm full width at half maximum.[25]

PET images with harmonised resolution were spatially normalised to the Montreal Neurological Institute space using

the Statistical Parametric Mapping software package (SPM12) with default parameter settings.[26] The spatially normalised images were smoothed to spatial resolution of 12 mm full width at half maximum. No intensity scaling was performed (not required for covariance pattern analysis).

Covariance pattern analysis of FDG-PET

Spatial covariance analysis of FDG-PET was performed with scaled subprofile model principal component analysis[27–29] implemented in the Scan Analysis and Visualization Processor software package (ScAnVP7.0w)[28–30] freely available from the Feinstein Institute for Medical Research (https://feinsteinneuroscience.org/imaging-software/download-software). This software was used for automatic computation of the expression of the predefined 'North American' PSP-related pattern (PSPRP)[31] in each individual FDG-PET. The 'North American' PSPRP was derived in an independent sample of 10 patients with PSP diagnosis according to the NINDS criteria[32] and 10 age- and sex-matched healthy volunteers (online supplemental figure 1B).[31] No attempt was made to adapt the 'North American' PSPRP to the current dataset. The PET-PSPRP expression scores were corrected for age as described previously.[10]

Statistical analyses

MRI-SVM-PSP and PET-PSPRP expression score were transformed to z-scores relative to the mean and SD in the HC subjects. The relationship between both z-scores was tested by linear regression (including constant term) of the PET-PSPRP expression z-score with the MRI-SVM-PSP score as independent variable.

Receiver operating characteristic (ROC) analysis was used to characterise the utility for PSP detection separately for the MRI-SVM-PSP z-score and the PET-PSPRP expression z-score. DeLong's test was used to compare the areas under the ROC curves.[33] ROC analyses were first performed including all PSP patients. Additional analyses were performed including only PSP-RS, only vPSP (all subtypes), only cortical vPSP or only subcortical vPSP.

Cut-off values for dichotomisation of the z-scores were obtained from the ROC curves according to Youden's criterion. Contingency tables were used to characterise balanced accuracy, sensitivity and specificity, separately for both z-scores. McNemar's method was applied to the 2×2 contingency table of the predictions (PSP or HC) by the dichotomised MRI-SVM-PSP z-score versus the dichotomised PET-PSPRP expression z-score to test their difference in the proportion of PSP predictions for significance.

To examine whether the combination of the MRI-SVM-PSP z-score with the PET-PSPRP expression z-score could enhance PSP detection, their sum (MRI-SVM-PSP z-score+PET-PSPRP expression z-score) and product (MRI-SVM-PSP z-score×PET-PSPRP expression z-score) as well as a binary logistic regression model of the group label (PSP or HC) with the two z-scores as predictors were tested. The same ROC analyses were applied as for each z-score alone.

Additionally, decision tree analyses were used, employing 'classification and regression trees' (CART). A major feature of CART is that each tree branching results in two subgroups, simplifying the interpretation compared with other decision tree variants that allow more than two subgroups at each branching. The growth method of CART aims at minimum intranode impurity, that is, maximum intranode homogeneity (optimally, all cases of a terminal node belong to the same category). The measure of impurity to be minimised by the CART was the Gini-Simpson diversity index, which is the standard for categorical dependent variables. The Gini-Simpson index of a node characterises the rate of false categorisation when randomly selected cases are randomly categorised according to the distribution of the categories within the node.[34] The minimum improvement in the Gini-Simpson index required to split a node was set to the default value of 0.0001. The primary CART analysis included all PSP patients and was free to select the predictor for the first branching (MRI-SVM-PSP or PET-PSPRP expression z-score). Secondary CART analyses were restricted to patients with PSP-RS or vPSP (all subtypes) in order to test for a potential impact of the clinical phenotype on the utility of sequential MRI and FDG-PET. The secondary CART analyses were also free regarding the selection of the predictor for the first branching. A tertiary CART analysis again included all PSP patients but was forced to select the PET-PSPRP expression z-score for the first branching in order to test for an added value of MRI if FDG-PET has already been performed. For PSP detection independent

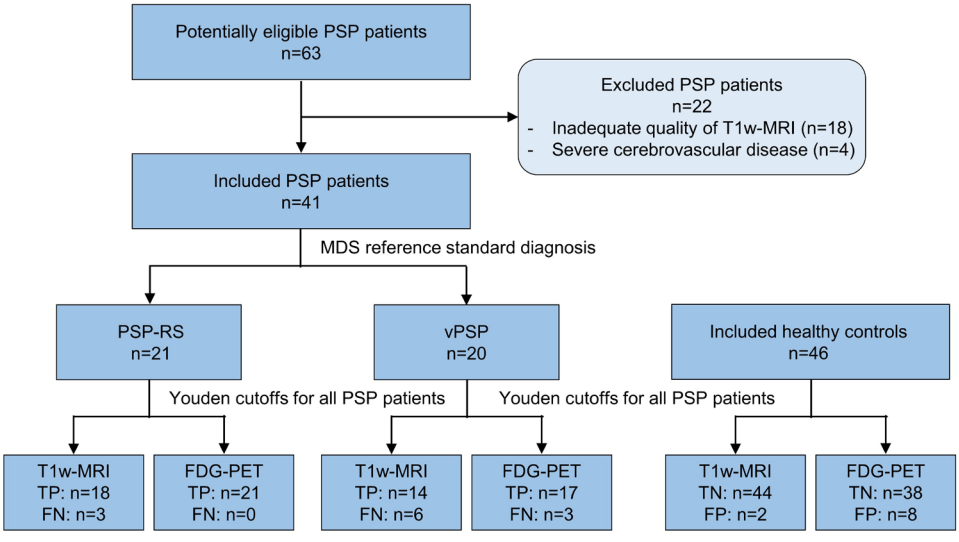

Figure 1 Flow of study participants. FDG-PET, [^{18}F]fluorodeoxyglucose-positron emission tomography; FN, false negative; FP, false positive; PSP, progressive supranuclear palsy; RS, Richardson's syndrome; TN, true negative; TP, true positive; vPSP, variant PSP.

Table 1 MRI-SVM-PSP z-score and PET-PSPRP expression z-score and their performance in the discrimination of the PSP patients (all, PSP-RS, vPSP, cortical vPSP and subcortical vPSP) from the HC subjects

	All PSP (n=41)	PSP-RS (n=21)	vPSP (n=20)	Cortical vPSP (n=11)	Subcortical vPSP (n=9)
z-score, mean±SD					
MRI-SVM-PSP	2.85±1.63	3.38±1.42	2.29±1.69	2.19±1.87	2.42±1.50
PET-PSPRP expression	2.14±1.51	2.57±1.23	1.69±1.67	1.68±1.83	1.71±1.56
Area under ROC curve (95% CI)					
MRI-SVM-PSP	0.910 (0.848 to 0.973)	0.950 (0.891 to 1.000)	0.868 (0.764 to 0.972)	0.832 (0.665 to 0.999)	0.913 (0.819 to 1.000)
PET-PSPRP	0.894 (0.826 to 0.963)	0.952 (0.908 to 0.996)	0.834 (0.726 to 0.942)	0.812 (0.662 to 0.963)	0.860 (0.738 to 0.981)
DeLong P alue	0.71	0.95	0.63	0.85	0.50
Youden cut-off for dichotomisation of z-scores					
MRI-SVM-PSP	2.013	2.327	2.013		
PET-PSPRP	0.516	0.728	0.516		
McNemar's test of dichotomised z-scores (MRI-SVM-PSP vs PET-PSPRP expression)					
P value	0.012	0.012	0.049	0.092	0.057
Balanced accuracy in % (95% CI)					
MRI-SVM-PSP	86.9 (79.9 to 93.8)	91.8 (84.0 to 99.5)	82.8 (72.4 to 93.3)	84.2 (70.7 to 97.7)	81.2 (65.5 to 96.9)
PET-PSPRP	87.6 (80.9 to 94.4)	92.4 (87.2 to 97.6)	83.8 (74.3 to 93.3)	82.2 (69.6 to 94.8)	85.7 (74.1 to 97.4)
Sensitivity in % (95% CI)					
MRI-SVM-PSP	78.0 (63.9 to 88.8)	85.7 (67.0 to 96.2)	70.0 (48.3 to 86.8)	72.7 (43.5 to 92.4)	66.7 (34.5 to 90.5)
PET-PSPRP	92.7 (82.1 to 98.1)	100 (83.9 to 100)	85.0 (65.6 to 96.0)	81.8 (53.7 to 96.7)	88.9 (59.5 to 99.3)
Specificity in % (95% CI)					
MRI-SVM-PSP	95.7 (87.2 to 99.3)	97.8 (90.8 to 99.9)	95.7 (87.2 to 99.3)	95.7 (87.2 to 99.3)	95.7 (87.2 to 99.3)
PET-PSPRP	82.6 (70.0 to 91.6)	84.8 (72.6 to 93.2)	82.6 (70.0 to 91.6)	82.6 (70.0 to 91.6)	82.6 (70.0 to 91.6)

Among the 11 patients with cortical vPSP, 10 presented with predominant corticobasal syndrome and 1 with predominant frontal presentation. From the nine patients with subcortical vPSP, eight presented with predominant parkinsonism and one with progressive gait freezing. The Youden cut-offs optimised for the detection of vPSP (independent of the vPSP subtype) were also applied for the detection of cortical vPSP and subcortical vPSP separately. The small sample size of the latter subgroups did not allow reliable determination of specific cut-offs.
HC, healthy controls; PET, positron emission tomography; PSPRP, progressive supranuclear palsy-related pattern; RS, Richardson's syndrome; SVM, support vector machine; vPSP, variant PSP.

of phenotype (primary and tertiary CART), the minimum size of parent and child nodes was set to 30 and 15, respectively, to account for the rather small sample size. For detection of PSP-RS or vPSP (secondary CART), with consideration for a further reduction of the sample size, the minimum sizes were set at 20 and 10. The maximum depth was set to two for all CART.

IBM SPSS (V.27) was used for the statistical analyses.

RESULTS

The flow of study participants is shown in figure 1.

Mean z-scores in PSP patients are given in table 1. All differences compared with the HC subjects (z-score=0.0±1.0 by definition) were significant (all p<0.001).

The PET-PSPRP expression z-score was significantly correlated with the MRI-SVM z-score (standardised regression coefficient ß=0.660, p<0.001, figure 2A). However, the regression model explained only 43% of its variance, leaving room for improvement by the combination of both z-scores.

ROC results are given in table 1 and figure 2B. The area under the ROC curve was about 0.90 for detection of PSP independent of the phenotype, about 0.95 for detection of PSP-RS and about 0.85 for detection of vPSP, without significant differences between MRI-SVM-PSP and PET-PSPRP expression z-score (all p≥0.63, table 1). Balanced accuracy was very similar for both z-scores, about 87% for detection of PSP independent of the subtype, approximately 92% for detection of PSP-RS and about 83% for detection of vPSP (table 1). However, the MRI-SVM-PSP z-score provided about 13% higher specificity and about 15% lower sensitivity compared with the PET-PSPRP expression z-score, independent of the PSP phenotype (table 1).

Binary logistic regression of the group label (PSP vs HC) included both, the MRI-SVM-PSP z-score (coefficient B=0.814, p<0.001) and the PET-PSPRP expression z-score (B=1.122, p=0.004) in addition to a constant term (B=−2.022, p<0.001). The regression model resulted in a slightly higher area under the ROC curve of 0.934 (95% CI 0.883 to 0.984) for the detection of PSP in the whole sample compared with each z-score alone, but the differences did not reach significance (p≥0.15). Balanced accuracy, sensitivity and specificity of the regression model were 86.0%, 82.9% and 89.1%, respectively. The sum of the two z-scores achieved an area under the ROC curve of 0.931 (95% CI 0.878 to 0.983), their product achieved 0.835 (95% CI 0.736 to 0.934).

Primary and secondary CART selected the MRI-SVM-PSP z-score for the first branching and the PET-PSPRP expression z-score for a second split of the subgroup with normal MRI-SVM-PSP z-score after the first split, independent of the phenotype (figure 3A, figure 4). In the whole sample, the primary CART achieved balanced accuracy, sensitivity and specificity of 89.0%, 97.6% and 80.4%, respectively. When the analysis was restricted to PSP-RS/vPSP (secondary CART), balanced accuracy, sensitivity and specificity were 91.3/87.7%, 100/95.0% and 82.6/80.4%, respectively. The tertiary CART is shown in figure 3B.

DISCUSSION

The primary finding of this study was that the MRI-SVM-PSP score and the PET-PSPRP expression score provide about the same overall (balanced) accuracy for the detection of PSP, but with a relative shift towards higher specificity of the MRI score and higher sensitivity of the PET score, suggesting that FDG-PET

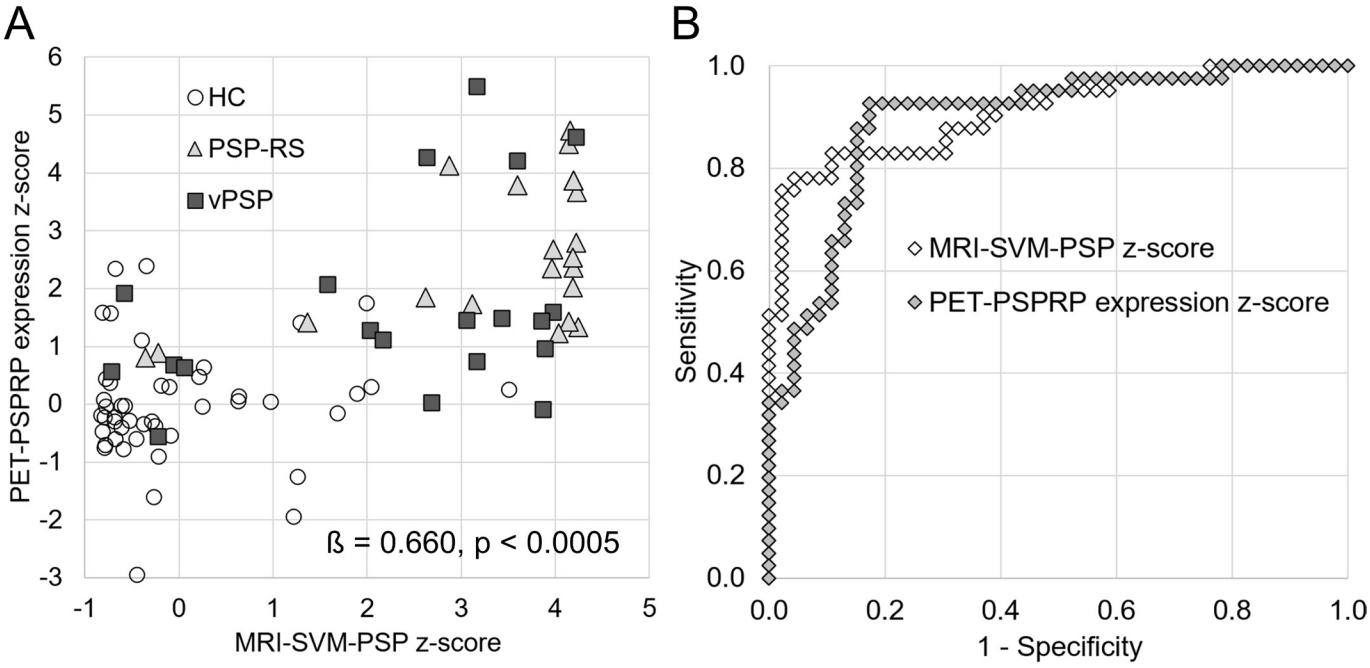

Figure 2 Scatter plot of the PET-PSPRP expression z-score versus the MRI-SVM-PSP z-score (A) and ROC curves for the detection of PSP independent of the phenotype (B). HC, healthy control; PET, positron emission tomography; PSPRP, progressive supranuclear palsy-related pattern; ROC, receiver operating characteristic; SVM, support vector machine; vPSP, variant PSP.

provides most added value in patients with normal (or inconclusive) MRI. This was confirmed by the primary decision tree analysis, which selected the MRI-SVM-PSP z-score for stratification at the first branching, followed by the PET-PSPRP expression z-score to identify PSP patients among those categorised as normal at the MRI-based split. The PET-PSPRP expression z-score identified 9 of 10 (90%) of the PSP patients with normal MRI (figure 3A). Only a single of the 41 PSP patients was misclassified as normal by the decision tree, resulting in 97.6% sensitivity.

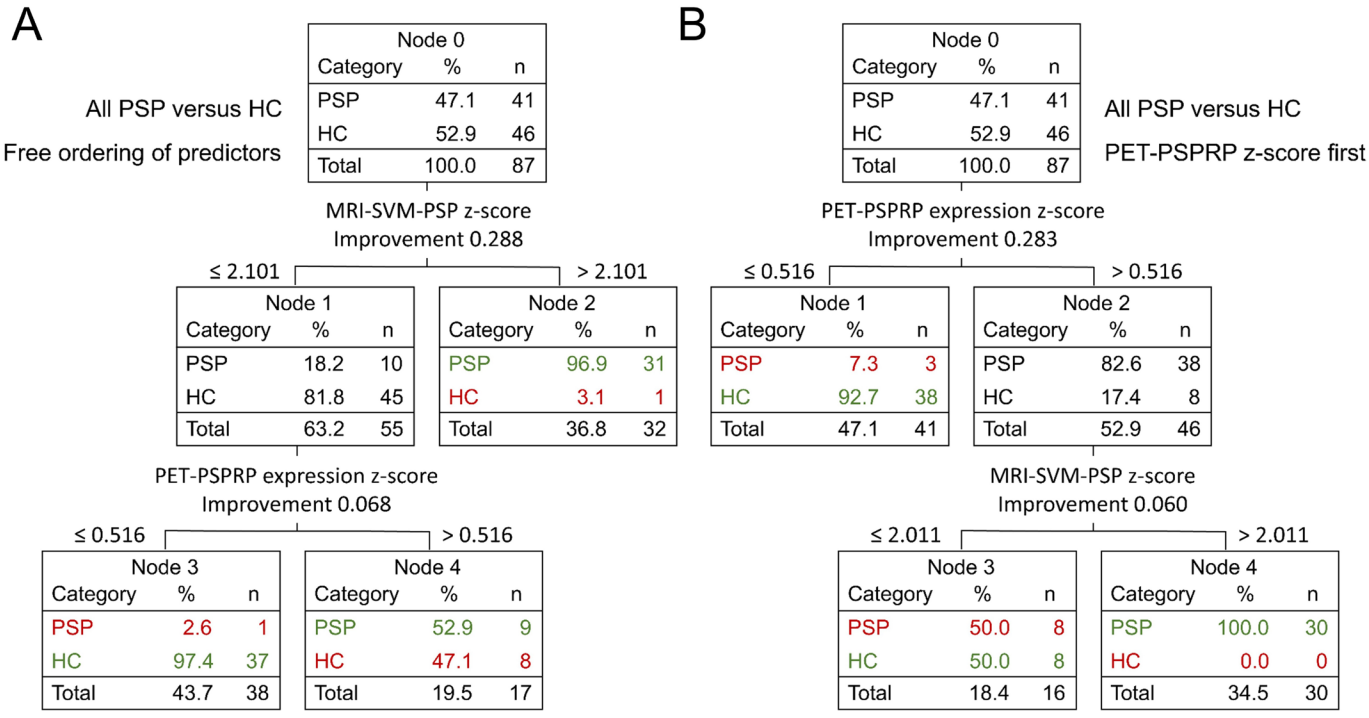

Figure 3 Decision tree analysis to combine the MRI-SVM-PSP z-score and the PET-PSPRP expression z-score for the detection of PSP independent of the phenotype. The primary decision tree (which was free to select a predictor for the first branching) selected the MRI-SVM-PSP z-score (A). (B) The tertiary decision tree that was forced to select the PET-PSPRP expression score for the first branching. Correctly classified cases are shown in green, falsely classified cases are shown in red. HC, healthy controls; PET, positron emission tomography; PSPRP, progressive supranuclear palsy-related pattern; SVM, support vector machine.

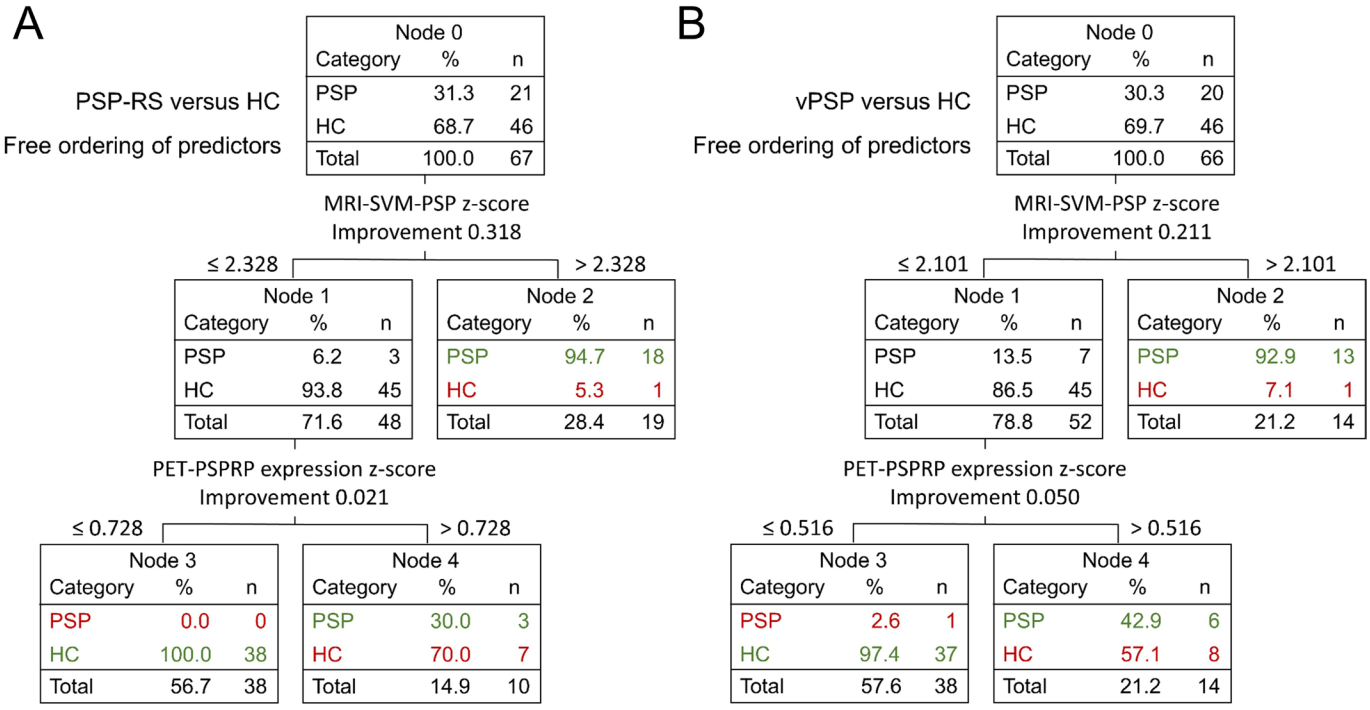

Figure 4 Secondary decision tree analyses to combine the MRI-SVM-PSP z-score and the PET-PSPRP expression z-score for the detection of PSP-RS (A) or vPSP (B). Correctly classified cases are shown in green, falsely classified cases are shown in red. HC, healthy controls; PET, positron emission tomography; PSPRP, progressive supranuclear palsy-related pattern; RS, Richardson's syndrome; SVM, support vector machine; vPSP, variant PSP.

Both z-scores were provided as predictors to the primary decision tree on equal footing, that is, without any bias to prefer one or the other for the first branching. The fact that the decision tree preferred the MRI-SVM-PSP z-score for the first branching makes the model compatible with clinical practice. Structural MRI is part of the basic diagnostics in patients suspected to have PSP to rule out secondary causes[16] and, therefore, always should be performed prior to FDG-PET. Furthermore, procedure guideline for brain FDG-PET recommend that the findings of structural brain imaging, preferably MRI, are taken into account when interpreting brain FDG-PET in order to avoid misinterpretation of structural/vascular lesions (and their at-distance effects on the cerebral glucose metabolism) as indication of a primary neurodegenerative disease.[35] Thus, the primary decision tree model obtained in this study is ready for use in clinical practice.

The finding of about the same (balanced) accuracy with higher specificity of the MRI-SVM-PSP z-score and higher sensitivity of the PET-PSPRP expression z-score was independent of the PSP phenotype (table 1). Furthermore, the secondary decision trees for the detection of PSP-RS or vPSP also selected the MRI-SVM-PSP z-score for the first branching (figure 4). The improvement of within-node homogeneity (Gini-Simpson diversity index) by the MRI-SVM-PSP z-score at the first branching was somewhat smaller for the detection of vPSP compared with PSP-RS (0.211 vs 0.318), and the improvement by the PET-PSPRP expression z-score at the second branching was somewhat larger for the detection of vPSP (0.050 vs 0.021). This suggests that the added value of FDG-PET in patients with normal (or inconclusive) MRI is somewhat larger in vPSP than in PSP-RS.

Higher sensitivity of FDG-PET for PSP detection compared with structural MRI might be explained by neuronal dysfunction preceding brain volume loss resulting in greater and more consistent changes in FDG-PET compared with MRI-based volumetry at early PSP stages.[36] Concerning lower specificity of FDG-PET, retrospective inspection identified frontal widening of the longitudinal fissure as a major source of false positive FDG-PET.[10] Differences in added value of FDG-PET between vPSP and PSP-RS might be explained by phenotype-specific spatiotemporal patterns in FDG-PET and MRI-based volumetry.[37 38]

The head-to-head comparison of MRI-based and FDG-PET-based PSP detection in the current study was restricted to SVM classification of individual atrophy profiles from fully automatic MRI volumetry and fully automatic covariance pattern expression analysis of FDG-PET. However, better specificity of MRI and better sensitivity of FDG-PET are also true when comparing visual interpretation of the two modalities. In the same subjects as included in the current study, visual reading of MRI based on descriptive signs (hummingbird, morning glory, Mickey-Mouse) and supported by automatic volumetry (same compartments as used as input to the SVM) achieved balanced accuracy, sensitivity and specificity of 74.9%, 58.5% and 91.3% for the detection of PSP independent of the phenotype (majority vote of 3 readers).[18] Visual interpretation of FDG-PET supported by univariate voxel-based testing achieved balanced accuracy, sensitivity and specificity of 72.9%, 74.0% and 71.7%, respectively (majority vote of 5 readers).[10] Thus, the primary conclusion of this study, that is, that FDG-PET is most beneficial in patients with normal or inconclusive MRI, will also apply if the fully automatic multivariate methods are not available so that the image interpretation is based on visual inspection (without or with support by univariate quantitative/statistical features). Taken together, these findings indicate that the main conclusion applies regardless of the phenotypic range present in a specific setting and regardless of the approach utilised to assess PSP-typical patterns in T1w-MRI and FDG-PET. However, future guidelines should recommend the use of automatic multivariate techniques to support the interpretation of T1w-MRI and FDG-PET, as these techniques are clearly superior to visual inspection regarding the detection (and exclusion) of PSP features.

A literature search (online supplemental material) identified only two studies reporting results that allow comparison of T1w-MRI and FDG-PET regarding PSP detection.[39 40] Park *et*

al[39] found similar sensitivity of T1w-MRI and FDG-PET in 14 patients with probable PSP according to the NINDS-SPSP clinical criteria[32]: both, midbrain atrophy in T1w-MRI and midbrain hypometabolism in FDG-PET were detected in 12 patients (86%).[39] Zhao *et al*[40] found enlargement of the third ventricle in 15 (79%) and midbrain atrophy in 13 (68%) of 19 patients with MDS PSP diagnosis. Among the 13 patients who also underwent brain FDG-PET, 12 (92%) showed midbrain hypometabolism.[40] The authors concluded that FDG-PET provides better sensitivity for the detection of PSP, particularly at early clinical stages of the disease.[40] The analysis of T1w-MRI in this previous study was restricted to the visual detection of PSP-typical signs (hummingbird) and to planimetric measures. The analysis of FDG-PET was conducted using voxel-based statistical testing against an FDG-PET normal database. Multivariate techniques for the automatic classification of T1w-MRI and/or FDG-PET were not tested nor was any attempt made to combine MRI and PET or to identify a potential added value of FDG-PET after MRI or vice versa.

The current study has limitations. First, the patient sample was rather small. To some extent, this is explained by the fact that almost a third of the eligible PSP patients had to be excluded because the T1w-MRI was not adequate for morphometric/volumetric analyses, most often due to too large spacing between slices.[18] Future guidelines on the diagnosis of parkinsonian syndromes should recommend that T1w-MRI is performed with sufficient quality to allow reliable morphometric/volumetric analyses. Second, the rather small sample size did not allow splitting the dataset into training and test data. Thus, the cut-offs on the z-scores were determined in the same sample to which they were applied. This might have resulted in overly optimistic performance estimates due to overfitting. However, this applies to both z-scores and, therefore, most likely did not affect the findings with respect to their relative performance or with respect to their combination. Third, there was considerable variability of the time interval between PET and MRI in the PSP patients (due to the retrospective nature of the study). However, the PSP sample was rather balanced with respect to the sequence of PET and MRI so that a relevant bias to favour one modality over the other caused by their order of succession seems unlikely.

In conclusion, the current study indicates a clinically useful added value of brain FDG-PET beyond T1w-MRI for the diagnosis of PSP in patients with normal or inconclusive T1w-MRI, irrespective of the clinical presence of a PSP-RS or a vPSP predominance type, and irrespective of the methods used to test T1w-MRI and FDG-PET for PSP-typical features. The added value by FDG-PET might be somewhat larger in vPSP than in PSP-RS. If confirmed in independent patient samples, these findings might guide decision-making on sequential MRI and FDG-PET imaging in clinically unclear patients with suspected PSP.

Author affiliations
[1]Department of Diagnostic and Interventional Radiology and Nuclear Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[2]Swiss Epilepsy Center, Klinik Lengg, Zurich, Switzerland
[3]Department of Neurology, Hannover Medical School, Hannover, Germany
[4]Department of Nuclear Medicine, Hannover Medical School, Hannover, Germany
[5]Department of Nuclear Medicine, University Hospital of Munich, LMU Munich, Munich, Germany
[6]German Center for Neurodegenerative Diseases (DZNE), Munich, Germany
[7]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[8]Department of Neurology, University Medical Center Eppendorf, Hamburg, Germany
[9]Department of Nuclear Medicine, University Hospital Augsburg, Augsburg, Germany
[10]Department of Neurology, University Hospital of Munich, LMU Munich, Munich, Germany
[11]Department of Diagnostic and Interventional Neuroradiology, Hannover Medical School, Hannover, Germany
[12]Medical Physics and Radiation Protection, University Hospital Augsburg, Augsburg, Germany
[13]Institute of Neurology, Department of Medical and Surgical Sciences, University "Magna Graecia" of Catanzaro, Catanzaro, Italy
[14]Department of Neurology, University of Leipzig, Leipzig, Germany
[15]Department of Neurology and Clinical Neurophysiology, University Hospital Augsburg, Augsburg, Germany
[16]Department of Radiology, University Hospital of Munich, LMU Munich, Munich, Germany
[17]Feinstein Institutes for Medical Research Manhasset, Manhasset, New York, USA
[18]Department of Nuclear Medicine, University Hospital of Leipzig, Leipzig, Germany
[19]Department of Neuroradiology, Charité - Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany

Correction notice In August 2024, this paper was resupplied as open access.

Acknowledgements We thank Karl-Titus Hoffmann, Institute for Neuroradiology, University Hospital of Leipzig, Leipzig, Germany and Xiaoqi Ding, Department of Diagnostic and Interventional Neuroradiology, Hannover Medical School, Hannover, Germany, for support in data collection. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica; Biogen; Bristol-Myers Squibb Company; CereSpir; Cogstate; Eisai; Elan Pharmaceuticals; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech; Fujirebio; GE Healthcare; IXICO; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development.; Lumosity; Lundbeck; Merck & Co; Meso Scale Diagnostics; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer; Piramal Imaging; Servier; Takeda Pharmaceutical Company and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organisation is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. GH is a member of the Alzheimer's Disease Neuroimaging Initiative.

Contributors RB: design, execution, analysis, drafting of the manuscript. H-JH, FW, GB, MB, IA, CB, MK, JL, NM, SS, OS, HB, MPW and GH: design, execution, analysis, editing of final version of the manuscript. MP-N, AD, SK, AR, AQ, SR, J-JR, CS, PGS and DE: execution, analysis, editing of final version of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Disclaimer Data used for preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Competing interests H-JH has used atlas-based volumetric MRI analysis in industry-sponsored research projects. CB received a grant from the Hilde-Ulrichs-Stiftung, served as a consultant for Bial, Hormosan Pharma, Merz Pharmaceuticals and Zambon and received honoraria for scientific presentations from Abbvie, Bial, Stada Pharma, TAD Pharma, UCB Pharma and Zambon. MP-N received lecture fees from Abbott, Abbvie, Boston Scientific and served as consultant for Medtronic, Boston Scientific, Abbott, Zambon and Abbvie. SK was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy–ID 390857198), the Ehrmann Foundation and the Lüneburg Heritage. SK receives research funding from CurePSP and reports travel support from Life Molecular Imaging outside the submitted work. MK received honoraria for scientific presentations from Abbvie and Ever Pharma. JL reports speaker fees from Bayer Vital, Biogen, EISAI, TEVA and Roche, consulting fees from Axon Neuroscience and Biogen, author fees from Thieme medical publishers and W. Kohlhammer medical publishers and is inventor in a patent 'Oral Phenylbutyrate for Treatment of Human 4-Repeat Tauopathies' (EP 23 156 122.6) filed by LMU Munich. In addition, he reports compensation for serving as chief medical officer for MODAG, is beneficiary of the phantom share program of MODAG and is inventor in a patent 'Pharmaceutical Composition and Methods of Use' (EP 22 159 408.8) filed by MODAG, all activities outside the submitted work. J-JR received speaker honoraria from GE Healthcare. OS received research support from Life Molecular Imaging. HB received reader honoraria from Life Molecular Imaging and speaker honoraria from Novartis/AAA. MPW received speaker or consultancy honoraria from

Alexion, Bayer Healthcare, Biogen, Biologix, Bristol Myers Squibb, Celgene, Genilac, Imcyse, IXICO, Icometrix, Medison, Merck-Serono, Novartis, Roche, Sanofi-Genzyme. Publication royalties from Springer and Elsevier. GH was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy–ID 390857198) and within the Hannover Cluster RESIST (EXC 2155–project number 39087428), the EU/EFPIA/Innovative Medicines Initiative (2) Joint Undertaking (IMPRIND grant no 116060), the European Joint Programme on Rare Diseases (Improve-PSP), Deutsche Forschungsgemeinschaft (DFG, HO2402/6-2 Heisenberg Program, HO2402/18-1 MSAomics), the VolkswagenStiftung (Niedersächsisches Vorab), the Petermax-Müller Foundation (Etiology and Therapy of Synucleinopathies and Tauopathies); participated in indurtry-sponsored research projects from Abbvie, Biogen, Biohaven, Novartis, Roche, Sanofi, UCB; served as a consultant for Abbvie, Alzprotect, Aprineua, Asceneuron, Bial, Biogen, Biohaven, Kyowa Kirin, Lundbeck, Novartis, Retrotope, Roche, Sanofi, UCB; received honoraria for scientific presentations from Abbvie, Bayer Vital, Bial, Biogen, Bristol Myers Squibb, Kyowa Kirin, Roche, Teva, UCB, Zambon; received publication royalties from Academic Press, Kohlhammer and Thieme. All other authors declare that they have no potential conflicts of interest.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Ethics Committee of the Hannover Medical School, reference number 9400_BO_K_2020. Waiver of informed consent for the retrospective analyses in this study was obtained from the relevant ethics review boards.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Ralph Buchert http://orcid.org/0000-0002-0945-0724
Hans-Jürgen Huppertz http://orcid.org/0000-0003-3856-9094
Florian Wegner http://orcid.org/0000-0002-9931-2666
Georg Berding http://orcid.org/0000-0001-5592-8373
Ivayla Apostolova http://orcid.org/0000-0003-0290-7186
Carsten Buhmann http://orcid.org/0000-0001-8540-3789
Monika Poetter-Nerger http://orcid.org/0000-0001-7680-2147
Sabrina Katzdobler http://orcid.org/0000-0002-3512-5984
Martin Klietz http://orcid.org/0000-0002-3054-9905
Johannes Levin http://orcid.org/0000-0001-5092-4306
Nima Mahmoudi http://orcid.org/0000-0002-2053-9623
Andreas Rinscheid http://orcid.org/0009-0008-8625-0217
Andrea Quattrone http://orcid.org/0000-0003-2071-2083
Jost-Julian Rumpf http://orcid.org/0000-0002-7357-713X
Christine Schneider http://orcid.org/0009-0003-2815-0262
Phoebe G Spetsieris http://orcid.org/0000-0001-7368-2354
David Eidelberg http://orcid.org/0000-0002-0854-864X
Mike P Wattjes http://orcid.org/0000-0001-9298-2897
Günter Höglinger http://orcid.org/0000-0001-7587-6187

REFERENCES

1 Stamelou M, Respondek G, Giagkou N, *et al*. Evolving concepts in progressive supranuclear palsy and other 4-repeat tauopathies. *Nat Rev Neurol* 2021;17:601–20.
2 Roemer SF, Grinberg LT, Crary JF, *et al*. Rainwater charitable foundation criteria for the neuropathologic diagnosis of progressive supranuclear palsy. *Acta Neuropathol* 2022;144:603–14.
3 Höglinger GU, Respondek G, Stamelou M, *et al*. Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria. *Mov Disord* 2017;32:853–64.
4 Respondek G, Stamelou M, Kurz C, *et al*. The phenotypic spectrum of progressive supranuclear palsy: a retrospective multicenter study of 100 definite cases. *Mov Disord* 2014;29:1758–66.
5 Ali F, Martin PR, Botha H, *et al*. Sensitivity and specificity of diagnostic criteria for progressive supranuclear palsy. *Mov Disord* 2019;34:1144–53.
6 van Eimeren T, Antonini A, Berg D, *et al*. Neuroimaging biomarkers for clinical trials in atypical parkinsonian disorders: proposal for a neuroimaging biomarker utility system. *Alz & Dem Diag Ass & Dis Mo* 2019;11:301–9.
7 Saeed U, Lang AE, Masellis M. Neuroimaging advances in parkinson's disease and atypical parkinsonian syndromes. *Front Neurol* 2020;11:572976.
8 Whitwell JL, Höglinger GU, Antonini A, *et al*. Radiological biomarkers for diagnosis in PSP: where are we and where do we need to be? *Mov Disord* 2017;32:955–71.
9 Brooks DJ. Imaging of genetic and degenerative disorders primarily causing parkinsonism. *Handb Clin Neurol* 2016;135:493–505.
10 Buchert R, Wegner F, Huppertz H, *et al*. automatic covariance pattern analysis outperforms visual reading of ^{18}F-fluorodeoxyglucose-positron emission tomography (FDG-PET)in variant progressive supranuclear palsy. *Mov Disord* 2023;38:1901–13.
11 Huppertz H-J, Möller L, Südmeyer M, *et al*. Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification. *Mov Disord* 2016;31:1506–17.
12 Lee W. conventional magnetic resonance imaging in the diagnosis of parkinsonian disorders: ameta-analysis. *Movement Disord Clin Pract* 2021;8:217–23.
13 Mahlknecht P, Hotter A, Hussl A, *et al*. Significance of MRI in diagnosis and differential diagnosis of parkinson's disease. *Neurodegener Dis* 2010;7:300–18.
14 Meyer PT, Frings L, Rücker G, *et al*. ^{18}F-FDG PET in parkinsonism: differential diagnosis and evaluation of cognitive impairment. *J Nucl Med* 2017;58:1888–98.
15 Möller L, Kassubek J, Südmeyer M, *et al*. Manual MRI morphometry in parkinsonian syndromes. *Mov Disord* 2017;32:778–82.
16 Peralta C, Strafella AP, van Eimeren T, *et al*. Pragmatic approach on neuroimaging techniques for the differential diagnosis of parkinsonisms. *Mov Disord Clin Pract* 2022;9:6–19.
17 Sakurai K, Tokumaru AM, Shimoji K, *et al*. Beyond the midbrain atrophy: wide spectrum of structural MRI finding in cases of pathologically proven progressive supranuclear palsy. *Neuroradiology* 2017;59:431–43.
18 Wattjes MP, Huppertz H, Mahmoudi N, *et al*. brainMRIin progressive supranuclear palsy with richardson's syndrome and variant phenotypes. *Mov Disord* 2023;38:1891–900.
19 Caminiti SP, Alongi P, Majno L, *et al*. Evaluation of an optimized [^{18}f]fluoro-deoxy-glucose positron emission tomography voxel-wise method to early support differential diagnosis in atypical parkinsonian disorders. *Eur J Neurol* 2017;24:687–e26.
20 Garibotto V, Herholz K, Boccardi M, *et al*. Clinical validity of brain fluorodeoxyglucose positron emission tomography as a biomarker for alzheimer's disease in the context of a structured 5-phase development framework. *Neurobiol Aging* 2017;52:183–95.
21 Fazekas F, Chawluk JB, Alavi A, *et al*. MR signal abnormalities at 1.5 T in alzheimer's dementia and normal aging. *AJR Am J Roentgenol* 1987;149:351–6.
22 Jabbari E, Holland N, Chelban V, *et al*. Diagnosis across the spectrum of progressive supranuclear palsy and corticobasal syndrome. *JAMA Neurol* 2020;77:377–87.
23 Golbe LI, Ohman-Strickland PA. A clinical rating scale for progressive supranuclear palsy. *Brain (Bacau)* 2007;130:1552–65.
24 Jack CR Jr, Bernstein MA, Fox NC, *et al*. The alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008;27:685–91.
25 Joshi A, Koeppe RA, Fessler JA. Reducing between scanner differences in multi-center PET studies. *Neuroimage* 2009;46:154–9.
26 Statistical Parametric Mapping: The Analysis of Functional Brain Images. Academic Press, 2006.
27 Moeller JR, Strother SC, Sidtis JJ, *et al*. Scaled subprofile model: a statistical approach to the analysis of functional patterns in positron emission tomographic data. *J Cereb Blood Flow Metab* 1987;7:649–58.
28 Spetsieris P, Ma Y, Peng S, *et al*. Identification of disease-related spatial covariance patterns using neuroimaging data. *J Vis Exp* 2013;2013:50319:76:.
29 Eidelberg D. Metabolic brain networks in neurodegenerative disorders: a functional imaging approach. *Trends Neurosci* 2009;32:548–57.
30 Peng S, Ma Y, Spetsieris PG, *et al*. Characterization of disease-related covariance topographies with SSMPCA toolbox: effects of spatial normalization and PET scanners. *Hum Brain Mapp* 2014;35:1801–14.
31 Eckert T, Tang C, Ma Y, *et al*. Abnormal metabolic networks in atypical parkinsonism. *Mov Disord* 2008;23:727–33.
32 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 (ECronicon)* 1996;47:1–9.
33 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.
34 Simpson EH. Measurement of diversity. *Nat New Biol* 1949;163:688.
35 Guedj E, Varrone A, Boellaard R, *et al*. EANM procedure guidelines for brain pet imaging using [^{18}f]fdg, version 3. *Eur J Nucl Med Mol Imaging* 2022;49:632–51.
36 Dukart J, Kherif F, Mueller K, *et al*. Generative FDG-PET and MRI model of aging and disease progression in alzheimer's disease. *PLoS Comput Biol* 2013;9:e1002987.

37 Quattrone A, Sarica A, Buonocore J, *et al*. Differentiating between common PSP phenotypes using structural MRI: a machine learning study. *J Neurol* 2023;270:5502–15.
38 Srulijes K, Reimold M, Liscic RM, *et al*. Fluorodeoxyglucose positron emission tomography in richardson's syndrome and progressive supranuclear palsy-parkinsonism. *Mov Disord* 2012;27:151–5.
39 Park HK, Kim JS, Im KC, *et al*. Functional brain imaging in pure akinesia with gait freezing: [^{18}F] FDG PET and [^{18}F] FP-CIT PET analyses. *Mov Disord* 2009;24:237–45.
40 Zhao P, Zhang B, Gao S, *et al*. Clinical, MRI and ^{18}F-FDG-PET/CT analysis of progressive supranuclear palsy. *J Clin Neurosci* 2020;80:318–23.