BMJ Open

Diagnostic Accuracy of Imaging Brain Vesicular Monoamine Transporter type-2 (VMAT2) in Clinically Uncertain Parkinsonian Syndrome (CUPS): a 3 year follow-up study.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025533
Article Type:	Research
Date Submitted by the Author:	21-Jul-2018
Complete List of Authors:	Xu, San San; Austin Health, Department of Molecular Imaging and Therapy; Austin Health, Department of Neurology Alexander, Paschal; Austin Health, Department of Molecular Imaging and Therapy Lie, Yenni; Austin Health, Department of Molecular Imaging and Therapy Dore, Vincent; Austin Health, Department of Molecular Imaging and Therapy Bozinovski, Svetlana; Austin Health, Department of Molecular Imaging and Therapy Mulligan, Rachel; Austin Health, Department of Molecular Imaging and Therapy Young, Kenneth; Austin Health, Department of Molecular Imaging and Therapy Villemagne, Victor; Austin Health, Department of Molecular Imaging and Therapy; Florey Institute of Neuroscience and Mental Health Rowe, Christopher C.; Austin Health, Department of Molecular Imaging and Therapy; Florey Institute of Neuroscience and Mental Health
Keywords:	Diagnostic accuracy, Molecular imaging, Parkinson-s disease < NEUROLOGY, PET, VMAT2

TITLE PAGE

Diagnostic Accuracy of Imaging Brain Vesicular Monoamine Transporter type-2 (VMAT2) in Clinically Uncertain Parkinsonian Syndrome (CUPS): a 3 year follow-up study.

Authors:

San San Xu[1,2], Paschal K Alexander[1], Yenni Lie[1], Vincent Dore[1, 3, 4], Svetlana Bozinovski[1], Rachel S. Mulligan[1], Kenneth Young[1], Victor L. Villemagne[1,3], Christopher C. Rowe[1,3].

1. Department of Molecular Imaging and Therapy, Austin Health, Melbourne, Victoria, Australia.
2. Department of Neurology, Austin Health, Melbourne Victoria, Australia
3. Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, Victoria, Australia.
4. Health & Biosecurity Flagship, The Australian eHealth Research Centre, The Commonwealth Scientific and Industrial Research Organisation, Melbourne, Vicotira, Australia

Corresponding author:

San San Xu,

Neurologist, Department of Neurology, Austin Heath,

145 Studley Road, Heidelberg, Victoria, 3084, Australia

Telephone: +61400851823, +61-3-9496 5529

Fax: +61-3-9496 4065

Email: sansan.xu@austin.org.au

Word count: 2905

Abstract word count: 248

Number of references: 38

Keywords: Diagnostic accuracy; molecular imaging; Parkinson's disease; PET; VMAT2

ABSTRACT

Objectives: In a previous study, we reported that ^{18}F-AV-133 Vesicular monoamine transporter type-2 (VMAT2) PET in patients with Clinically Uncertain Parkinsonian Syndromes (CUPS) changed diagnosis and management and increased diagnostic confidence. Here, we aim to further validate the diagnostic utility of ^{18}F-AV-133 PET by comparison to follow-up diagnosis in the CUPS cohort.

Methods: We obtained the current clinical diagnosis from the patient and treating specialist and compared this to the diagnosis suggested three years earlier by the ^{18}F-AV-133 PET. A second ^{18}F-AV-133 PET was available in those with a discordant or inconclusive final diagnosis.

Results: 81 of the 85 patients (95%) previously recruited to the CUPS study had follow-up of which 79 had a clinical diagnosis and two remained CUPS. The diagnosis was in agreement with the initial ^{18}F-AV-133 PET scan result in 74 cases (94%). Five patients (6%) had a discordant diagnosis; one patient with a diagnosis of rubral tremor had a severely abnormal scan that had further worsened when rescanned; four cases with normal scans had a clinical diagnosis of Parkinson's disease but their repeat scans remained normal. Two patients with suspected genetic disorders remained classified as CUPS and both had normal scans. In the 24 CUPS cohort patients where ^{18}F-AV-133 PET initially changed diagnosis, this change was supported by follow-up clinical diagnosis in all but the one rubral tremor case.

Conclusion: ^{18}F-AV-133 PET is a useful tool in improving diagnostic accuracy in CUPS providing results and diagnostic changes that remain robust after 3 years follow-up.

STRENGTHS AND LIMITATIONS

- This study supports the diagnostic accuracy of the novel imaging technique ^{18}F-AV-133 PET in patients with Clinically Uncertain Parkinsonian Syndromes after 3 years follow-up.
- The findings further validate the optimal binding reduction threshold of 50% for abnormal scans.
- The ^{18}F-AV-133 PET offers advantages compared to DaTSCAN including improved spatial resolution and better image quality, more precise quantitation and reduced tracer uptake and scan acquisition times, without the need for thyroid blockade or other patient preparation.
- A limitation of this study is the reliance on long term clinical follow-up as the standard of truth rather than histopathological diagnosis.

MANUSCRIPT

INTRODUCTION

Parkinson's disease (PD) is a common neurodegenerative condition, second only to Alzheimer's disease and the most prevalent of the Parkinsonian syndromes. Diagnostic certainty of Parkinson's disease and the other Parkinsonian syndromes (multiple system atrophy (MSA)[1], progressive supranuclear palsy (PSP)[2] and cortico-basal syndrome (CBS)[3] can only be confirmed by histopathological demonstration of the characteristic pathology and resultant nigrostriatal degeneration. In clinical practice, diagnosis of Parkinson's disease relies on the presence of bradykinesia and at least one of rest tremor, rigidity or postural instability[4]. Atypical or mild clinical features may delay diagnosis and introduction of appropriate therapies. In a tertiary movement disorders centre, the diagnostic accuracy of clinical assessment in post-mortem clinicopathologically confirmed Parkinson's disease did improve from 79% to 90% over a 10 year period[5 6]. However, the rate of misdiagnosis is likely higher in early disease and in the primary care setting. In community patients with Parkinsonian features or on anti-parkinsonian medications, only 53% - 83% of patients fulfilled the criteria for probable Parkinson's disease[7 8]. Other conditions that may mimic Parkinson's disease include essential tremor, dystonia, drug induced parkinsonism (DIP), vascular parkinsonism and functional movement disorder. Misdiagnosis of these disorders can have significant prognostic and management implications.

Positron emission tomography (PET) and single photon emission computerised tomography (SPECT) imaging can accurately evaluate the nigrostriatal system and aid in early diagnosis of Parkinson's disease[9]. Vesicular monoamine transporter type 2 (VMAT2) plays an integral role in pre-synaptic dopamine uptake and storage and is a reliable marker of nigrostriatal terminal integrity [10 11]. ^{18}F-AV-133 is a novel ^{18}F–labelled dihydrotetrabenazine analogue that selectively binds to VMAT2 with high affinity and allows for in vivo evaluation of VMAT2

density. Several clinical studies have demonstrated the feasibility of ^{18}F-AV-133 PET technique to assist in the diagnosis of Parkinson's disease and dementia with Lewy bodies[10-14]. ^{18}F-AV-133 PET has multiple potential advantages compared to dopamine transporter SPECT, including improved image quality and quantification, reduced tracer administration to scan interval time and reduced scan duration. Further, there is no requirement for pre-scan thyroid blockade in contrast to the iodine-123 labelled SPECT dopamine transporter tracers.

In a previous study, we investigated the management impact of ^{18}F-AV-133 PET imaging in patients with Clinically Uncertain Parkinsonian Syndromes (CUPS)[14]. The results of the ^{18}F-AV-133 PET altered diagnosis in 23% of participants (11 of 47) and changed management in more than half of the cases (53%; 25 of 47). Furthermore, diagnostic confidence in clinicians increased in 74% of the participants after the scan, regardless of whether the result was normal or abnormal. Total enrolment in the CUPS study subsequently reached 85 and the present study aims to further confirm the diagnostic accuracy of ^{18}F-AV-133 PET by comparing the results of the initial scan with the clinical diagnosis at 3 years follow-up in the total cohort.

METHODS

Study subjects

All patients previously recruited in our CUPS study were eligible for the current study[14]. Patients with CUPS were recruited from the private and public clinics of movement disorders specialists from across the city of Melbourne, Australia. The criteria for uncertainty of diagnosis was at the discretion of the referring clinician and included the presence of atypical features of parkinsonism including poor levodopa responsiveness, lack of disease progression, dystonia and young age of onset. Participants were excluded if they had a

history of malignancy within the last 5 years or if they were unable to provide informed written consent.

Patient and public involvement

Patients and the public were not involved in the design or analysis of this study. Once published, the results of the study will be summarised in a letter and disseminated to the participants and doctors involved in patient recruitment.

Study design

This was a single-centre, prospective experimental study with a mean follow-up interval of 3 years ± 6 months (range: 18 – 68 months). The study was conducted in accordance with local and international standards and approved by the Austin Health Human Research Ethics Committee. All participants provided written consent prior to their inclusion in the study. At the time of follow-up, phone calls were made to the treating movement disorders specialist and study participant to establish the current diagnosis for the patient. If the treating neurologist had changed during the follow-up period, the diagnosis was made by the most recent clinician involved in the care of the participant. The clinicians had access to the initial ^{18}F-AV-133 PET scan results. The diagnostic categories were classified into parkinsonism with nigrostriatal degeneration (including idiopathic Parkinson's disease, multiple system atrophy, progressive supranuclear palsy and cortico-basal syndrome), other (including essential tremor, dystonia, drug induced parkinsonism, functional (psychogenic), monosymptomatic resting tremor) or an unclear diagnosis i.e. remained CUPS. Follow-up diagnosis was considered in agreement with the initial PET scan diagnosis if it remained in the same binary diagnostic category i.e. parkinsonism with nigrostriatal degeneration or other. For example, if a participant had a change of diagnosis from Parkinson's disease to multiple system atrophy during the follow-up period, this was still considered a concordant

result. All participants had a follow-up AV-133 PET but only those for patients who had a clinical diagnosis which was discordant to the initial PET result are reported here.

PET scan protocol and image analysis

As previously described, a 20-minute emission PET scan was obtained two hours after intravenous injection of approximately 250 MBq of ^{18}F-AV133[11 14]. For attenuation-correction purposes, a rotation transmission sinogram in 3D mode with a single ^{137}Cs point source was acquired prior to radiotracer injection. The final images were reconstructed using a 3D row-action maximum-likelihood algorithm. The regional tracer binding of the caudate nucleus, anterior and posterior putamen were calculated using the ratio of regional activity to primary visual cortex, the latter being a region devoid of monoaminergic terminals, and therefore suitable as a reference standard. Each individual image was spatially normalised to a normal AV-133 template using statistical parametric mapping software (SPM8; Wellcome Department of Cognitive Neurology). A standard region of interest (ROI) template was previously constructed manually over 13 slices for the caudate and 8 slices for the putamen (each slice 2mm thick). The putamen ROI was bisected to give anterior and posterior putamen binding. Abnormal images were determined quantitatively and visually. Quantitatively, abnormal images were defined as those with a greater than 50% reduction in binding in the most affected posterior putamen, which corresponds to 4 standard deviations below the mean of the healthy control reference group[11 14] that consisted of 16 healthy controls; 9 males and 7 females, with a mean age of 72 +/- 5.1 (range 64 – 78 years). Scans were called visually abnormal when there was significant asymmetry in the posterior putamen or marked reduction in uptake in the putamen relative to the caudate nucleus. In two cases, visual analysis was abnormal when quantitative results were not but otherwise all classifications were concordant. The binding percentage for each region was calculated by subtracting the regional control group mean binding ratio from the patient result, then dividing this by the control group mean and then multiplying by one hundred. This threshold

is based on multiple histopathological studies which suggest that motor symptoms of Parkinson's disease only emerge after at least 50% loss of dopaminergic neurons in the substantia nigra [15 16]. The 3 year follow-up data was also used to test the validity of the 50% binding reduction threshold by comparison to ROC curve using the Youden criterion.

Outcome variables

The primary endpoint of the study was the proportion of patients who had a clinical diagnosis at 3 years follow-up, which was in agreement with the results of their initial ^{18}F-AV-133 PET scan. Secondary endpoints were the proportion of patients who had the same diagnosis in the follow-up period as that reached after the initial scan and the stability of diagnostic changes made after the first scan.

Statistical analysis

The results of the study are expressed as a mean ± standard deviation with accuracy figures derived from two by two contingency tables. The diagnostic accuracy of ^{18}F-AV-133 PET was further investigated using a receiver operator characteristics (ROC) analysis and area under the curve (AUC)[17], with binding reduction threshold determined using the Youden criterion[18]. Data processing and statistical analysis was conducted using Microsoft ® Excel ® 2016 software, Minitab 18 (Minitab Inc., Pennsylvania, USA) and R Version 3.4.3[19].

Data sharing

No additional data is available for sharing.

RESULTS

Study population

81 of the 85 participants (95%) enrolled in our CUPS study (47 previously described in Alexander et al, 2017[11 14]) were reviewed after 3 years. Four patients were lost to follow-up. Patient characteristics are outlined in Table 1. There were no significant differences when comparing participants involved in the study with those who were lost to follow-up.

Table 1: Patient characteristics at baseline in the follow-up and lost to follow-up groups

	Follow-up	Lost to follow-up
Demographics		
n	81	4
Age (mean ± SD)	57 ± 13.1	57 ± 17.0
Female	44 (54%)	2 (50%)
UPDRS Motor score mean ± SD	10.3 ± 4.2	8.4 ± 5.9
Hoehn and Yahr Score stage average	1.6 ± 0.76	1.1 ± 0.25
MMSE mean ± SD	28.9 ± 1.6	29.0 ± 1.4
Scan results		
Abnormal ^{18}F-AV-133 Scan	42 (52%)	2 (50%)
Baseline Pre-scan Diagnosis		
Neurodegenerative conditions	**45**	**2**
Parkinson's disease	31	1
Multiple system atrophy	1	
Progressive supranuclear palsy	1	
Corticobasal syndrome	1	
Alzheimer's disease	1	
Undefined	10	1
Non-degenerative conditions	**36**	**2**
Functional	13	2
Dystonia	10	0
Drug induced parkinsonism	5	0
Essential tremor	3	0
Monosymptomatic resting tremor	3	0
Vascular parkinsonism	1	0
Rubral tremor	1	0

UPDRS = Unified Parkinson's Disease Rating Scale

SD = Standard Deviation

Diagnosis summary

Of the 81 participants followed up, 79 (98%) had a specific clinical diagnosis and two cases had an inconclusive diagnosis i.e. remained CUPS. When a specific clinical diagnosis was available, the follow-up diagnosis was concordant with that suggested by the initial ^{18}F-AV-133 PET scan in 74 cases (94%), with a positive predictive value of 98%, a negative predictive value of 89% and a sensitivity and specificity of 91% and 97% respectively (derived from Table 2). The diagnostic agreement rate was slightly lower in the parkinsonism associated with nigrostriatal degeneration category compared to the "other" category (91% vs 97%). The agreement rate of clinical diagnosis at follow-up was 67% when compared to the most likely pre-scan clinical diagnosis and was 97% when compared to the initial post-scan diagnosis. The diagnostic accuracy of ^{18}F-AV-133 PET in predicting Parkinsonism with nigrostriatal degeneration was further evaluated with a ROC curve (Figure 1). The calculated AUC was 0.94 (95% confidence interval 0.88 - 0.99) with an optimal binding reduction threshold of 50% as per Youden criterion. The clinical diagnosis of all participants in the pre-scan, post-scan and follow-up period are listed individually in supplementary Table 1.

Table 2: Correlation between AV-133 PET scan result and diagnosis after 3 years

AV-133 PET scan result	**Diagnosis at 3 year follow-up**		
	Parkinsonism with nigrostriatal degeneration	Other diagnosis	Inconclusive diagnosis
Abnormal	41	1	0
Normal	4	33	2
Total	45	34	2

Discordant clinical diagnosis with imaging results

Five patients (6%) had a follow-up clinical diagnosis that did not concur with the results of their ^{18}F-AV-133 PET scan (Table 3). One participant with an abnormal scan was diagnosed

pre scan as rubral tremor, post scan as PD and at follow-up diagnosis had returned to rubral tremor despite worsening of the scan (Figure 2). This patient has an asymmetrical, resting, action and postural upper limb tremor that is levodopa responsive but has remained relatively stable for 30 years with no bradykinesia or rigidity. Brain MRI did not reveal a structural lesion that accounted for reduced AV-133 uptake in the absence of nigrostriatal degeneration.

Three participants with a follow-up clinical diagnosis of Parkinson's disease had a normal ^{18}F-AV-133 PET scan. In two of these patients the managing clinician now considers them as having PD phenotype due to Symptoms Without Evidence of Dopaminergic Deficit (SWEDD). Follow-up imaging in these two patients 2 years later remained stable (Figure 2 shows one of these cases). One patient had an initial pre-scan and post-scan diagnosis of dystonia. During the follow-up period, the diagnosis was revised to PD in the context of emerging bradykinesia and a good response to levodopa. Follow-up imaging at three years remained normal.

One patient had a follow-up diagnosis of progressive atypical parkinsonian syndrome. The initial and repeat ^{18}F-AV-133 PET scan two years later showed stable and symmetrical binding in the lower range but less than the 50% reduced cut-off (posterior putamen binding of -39% and -34% respectively).

Table 3: Scan results and diagnosis of patients with clinical diagnosis discordant to scan results or unknown

Case	Age	UPDRS	H& Y	Post scan diagnosis	3 year follow-up diagnosis	Posterior putamen binding	Follow-up posterior putamen binding	Putamen to caudate ratio	Left-right asymmetry ratio	Time between scans (months)
1	79	6	1	NDG PD	Rubral tremor	-85%	-94%	0.31	0.46	24
2*	51	9	1	NDG PD	NDG PD	-18%	-12%	1.21	0.98	27
3*	61	7	2	NDG PD	NDG PD	-20%	-16%	1.20	0.91	29
4	53	4	1	Dystonia	NDG PD	2%	-14%	1.12	0.98	36
5	53	5	1	NDG PD	NDG AP	-39%	-34%	1.21	0.97	26
6	20	6	1	UNK	UNK (Neurogenetic)	-16%	N/A	1.14	0.97	N/A
7	28	8	1.5	UNK	UNK (Dystonia parkinsonism syndrome)	9%	-24%	1.18	1.02	19

Putamen binding figures represent most affected side. Putamen to caudate ratio and asymmetry ratio represent posterior putamen results from baseline scans.

*Cases considered consistent with SWEDD (Scans without evidence of dopaminergic deficit) by treating specialist at follow-up.

UPDRS = Unified Parkinson's Disease Rating Scale

H & Y = Hoehn and Yahr

NDG = Neurodegenerative

PD = Parkinson's disease

AP = Atypical parkinsonism

UNK = Unknown

N/A = Not available

Cases with uncertain diagnosis

Two cases continue to be CUPS. One participant, a young male, had a pre-scan diagnosis of an unknown neurodegenerative condition. He had a normal ^{18}F-AV-133 PET scan and the post scan diagnosis remained unclear. The patient's symptoms have progressed but the current diagnosis remains "undefined neurogenetic condition". The patient did not return for repeat ^{18}F-AV-133 PET imaging. The second participant also had a normal initial ^{18}F-AV-133 PET scan. The pre-scan diagnosis was Parkinson's disease and the immediate post scan diagnosis was unclear. At follow-up, the participant clinically presents with a dystonia parkinsonism syndrome. A second ^{18}F-AV-133 PET scan 1.5 years after the initial study remained in the normal range but showed a decline in posterior putamen tracer binding from 9% to-24%. This participant has a sibling with early onset Parkinson's disease who had an abnormal ^{18}F-AV-133 PET.

Follow-up After Pre-scan to Post-scan Diagnostic Change.

Of the 81 CUPS with 3 year follow-up, 24 had a change in binary diagnostic classification due to the initial ^{18}F-AV-133 PET (see supplementary Table 1 for details). Of these, the current clinical diagnosis remained the same as the post scan diagnosis in 23 (96%). The patient with an original diagnosis of rubral tremor was re-classified as Parkinson's disease following an initial abnormal ^{18}F-AV-133 PET but the diagnosis had reverted back to rubral tremor at follow-up.

DISCUSSION

Our study provides further evidence that ^{18}F-AV-133 PET is a feasible adjunctive tool in the diagnosis of degenerative parkinsonism. The 3 year follow-up data validated the 50% binding reduction threshold and clinical diagnosis remained remarkably concordant with the

results of the ^{18}F-AV-133 PET scan (94% agreement rate) with an impressive sensitivity and specificity of 91% and 97% respectively. This is highly comparable to the sensitivity (87-98%) and specificity (80-100%) reported in the widely used ^{123}I-FP-CIT SPECT scans (DaTSCAN) [20-23]. When a follow-up diagnosis had been made, the diagnosis was in agreement with the pre-scan diagnosis in only 67% of cases, but was in agreement with the post-scan diagnosis made with the aid of the ^{18}F-AV-133 PET scan in 97%, highlighting the diagnostic challenges in early Parkinson's disease.

Four patients had a clinical diagnosis of Parkinson's disease or atypical parkinsonism despite a normal ^{18}F-AV-133 PET scan. This discrepancy has been widely reported with other measures of dopaminergic integrity and is referred to as Symptoms Without Dopaminergic Deficit (SWEDD). The number of SWEDD cases in the literature has been described to be between 3.5% to 20% in patients with clinical features of Parkinson's disease undergoing DAT scanning[24]. However, the entity of SWEDD remains controversial and may characterise a heterogenous group of conditions. Some studies have suggested that the tremor dominant presentation may represent adult onset dystonia[25 26]. However, pathologically confirmed cases of multiple system atrophy and corticobasal ganglionic degeneration with normal DaTSCAN have been described in the literature[27-29]. A patient with levodopa responsive parkinsonism and dyskinesia and a G2019S mutation in the LRRK2 gene with a normal ^{18}F-fluorodopa PET scan[30] has been reported, suggesting that imaging of the nigrostriatal pathway may be normal in some cases of early Parkinson's disease[31]. In the current study, an abnormal scan was defined as a $\geq$ 50% reduction in tracer binding in the posterior putamen compared to healthy controls. This threshold is based on post mortem studies suggesting that the clinical features of parkinsonism emerge after > 50% loss of dopaminergic neurons in the substantia nigra[15 16]. Therefore, it is feasible that this prescribed range will miss preclinical or early PD with very mild motor symptoms. This is supported by the observation of progressive clinical and AV-133 binding decline in our patient with a suspected genetic dystonia parkinsonism syndrome. Further studies evaluating ^{18}F-AV-133

PET in at risk patients such as those with REM sleep behaviour disorder[32] may shed light on the reasons for false negatives and could help identify an appropriate threshold for detection of preclinical individuals. Our 2 SWEDD cases had posterior putamen binding of -16.5% and -17% and demonstrated no decline with repeat scans after 2 years. Consistent with the literature of SWEDD, these subjects did not show clinical progression over the follow-up period[24 33]. The one patient who was reclassified from dystonia to Parkinson's disease at follow-up had minimal decline in the scans from a posterior putamen binding of 3% to -14% after 2 years.

There were three patients who had a change in diagnosis from Parkinson's disease to multiple system atrophy during the follow-up period. This is consistent with the literature, that suggests that pre-synaptic dopaminergic imaging cannot differentiate Parkinson's disease from atypical parkinsonian syndromes[9]. Evaluation of the postsynaptic dopaminergic systems with D2 receptor binding ligands or metabolic imaging with ^{18}F-FDG-PET has been suggested to further differentiate idiopathic Parkinson's disease from atypical parkinsonism[34 35].

There are advantages of ^{18}F-AV-133 PET compared to DaTSCAN including improved spatial resolution and reduced tracer uptake and scan acquisition time, without the need for thyroid blockade or other patient preparation. In addition to its role in diagnosis, the ^{18}F-AV-133 PET may prove to be a valuable tool for disease monitoring and in patient selection and evaluation of the therapeutic impact of interventions in clinical trials.

There are limitations to the current study. Firstly, the definitive diagnosis of Parkinson's disease relies on histopathological evidence[5] and this is not available in this cohort at this time. However, in view of the logistical challenges of post mortem studies, a clinical diagnosis such as that outlined by the UK Brain Bank criteria is commonly accepted as a substitute gold standard[4]. Additionally, a long clinical follow-up period, such as the one

employed in this study, has been reported to improve diagnostic accuracy[36]. Post-mortem studies of autopsy confirmed Parkinson's disease have revealed a correlation between ligand uptake in ^{123}I-FP-CIT SPECT and nigrostriatal neuronal loss[37 38]. Similar histopathological studies would be worthwhile to further validate the diagnostic accuracy of ^{18}F-AV-133 PET.

CONCLUSION

This study validates and extends the findings of our previous CUPS study, providing further evidence of the diagnostic value of ^{18}F-AV-133 PET, with a robust impact after 3 years of follow-up indicating that management change initiated by ^{18}F-AV-133 PET scan findings, was and remains appropriate.

ACKNOWLEDGMENTS

We wish to thank Drs Andrew Hughes, Andrew Evans, David Williams, John Merory, Katya Kotchet, Katrina Reardon, Richard Peppard, Malcolm Horne and Dominic Thyagarajan for their role in patient recruitment.

AUTHOR CONTRIBUTORSHIP STATEMENT

Dr San San Xu was involved in statistical analysis and wrote the manuscript. Dr Paschal K Alexander, Dr Yenni Lie and Dr Vincent Dore were involved in the research project execution and statistical analysis. Ms Svetlana Bozinvski was involved with research project organisation and execution. Ms Rachel S Mulligan and Mr Kenneth Young were involved in research project execution. Dr Victor L. Villemagne was involved in statistical analysis design, execution, review and critique. Professor Christopher C. Rowe was involved in

research project conception, organisation and execution, statistical analysis review and critique and manuscript review and critique.

FINANCIAL DISCLOSUES AND CONFLICT RELATED TO MANUSCRIPT

A grant for this study was provided to Professor Rowe to support this study by Avid Radiopharmaceuticals, who developed and own the rights to ^{18}F-AV-133.

FUNDING SOURCES

This was an investigator-initiated study supported by a grant from Avid Radiopharmaceuticals.

FINANCIAL DISCLOSURES OF ALL AUTHORS UNRELATED TO CURRENT RESEARCH

Professor Christopher Rowe has received research grants for imaging in dementia from Piramal Imaging, GE Healthcare, Cerveau, Astra Zeneca, Biogen and Navidea. He has been a consultant or paid speaker at sponsored conference sessions for Piramal Imaging, GE Healthcare, Astra Zeneca, Roche and Biogen. Dr Victor Villemagne has been a consultant or paid speaker at sponsored conference sessions for Piramal Imaging, GE Healthcare, Astra Zeneca and Novartis.

REFERENCES

1. Wenning GK, Colosimo C, Geser F, et al. Multiple system atrophy. *The Lancet Neurology* 2004;3(2):93-103. [published Online First: 2004/01/30]
2. Hauw J-J, Daniel S, Dickson D, et al. Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Neurology* 1994;44(11):2015-15.
3. Lee SE, Rabinovici GD, Mayo MC, et al. Clinicopathological correlations in corticobasal degeneration. *Annals of neurology* 2011;70(2):327-40.
4. Jankovic J. Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry* 2008;79(4):368-76.
5. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry* 1992;55(3):181-84.
6. Hughes AJ, Daniel SE, Lees AJ. Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. *Neurology* 2001;57(8):1497-99.
7. Meara J, Bhowmick BK, Hobson P. Accuracy of diagnosis in patients with presumed Parkinson's disease. *Age and ageing* 1999;28(2):99-102.
8. Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of Parkinson's disease in the community? *Journal of Neurology, Neurosurgery & Psychiatry* 2002;73(5):529-34.
9. Politis M. Neuroimaging in Parkinson disease: from research setting to clinical practice. *Nature Reviews Neurology* 2014;10(12):708-22.
10. Frey KA, Koeppe RA, Kilbourn MR, et al. Presynaptic monoaminergic vesicles in Parkinson's disease and normal aging. *Annals of neurology* 1996;40(6):873-84.
11. Okamura N, Villemagne VL, Drago J, et al. In vivo measurement of vesicular monoamine transporter type 2 density in Parkinson disease with 18F-AV-133. *Journal of Nuclear Medicine* 2010;51(2):223-28.
12. Villemagne VL, Okamura N, Pejoska S, et al. Differential diagnosis in Alzheimer's disease and dementia with Lewy bodies via VMAT2 and amyloid imaging. *Neurodegenerative Diseases* 2012;10(1-4):161-65.
13. Hsiao IT, Weng YH, Lin WY, et al. Comparison of 99mTc-TRODAT-1 SPECT and 18 F-AV-133 PET imaging in healthy controls and Parkinson's disease patients. *Nuclear medicine and biology* 2014;41(4):322-9. doi: 10.1016/j.nucmedbio.2013.12.017 [published Online First: 2014/02/08]

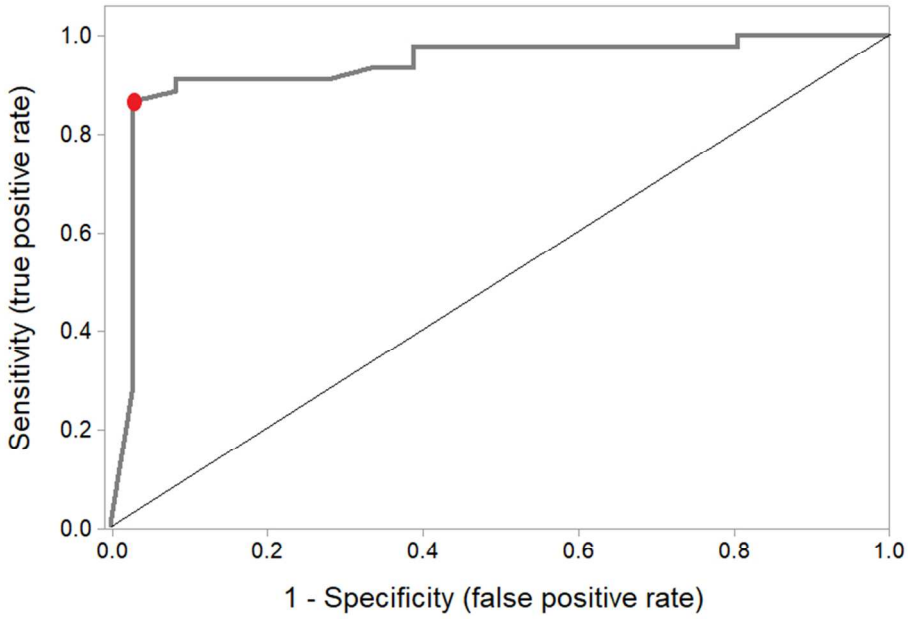
14. Alexander PK, Lie Y, Jones G, et al. Management impact of imaging brain vesicular monoamine transporter type-2 (VMAT2) in clinically uncertain parkinsonian syndrome (CUPS) with 18F-AV133 and PET. *Journal of Nuclear Medicine* 2017:jnumed. 116.189019.
15. Kordower JH, Olanow CW, Dodiya HB, et al. Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain : a journal of neurology* 2013;136(8):2419-31.
16. Damier P, Hirsch EC, Agid Y, et al. The substantia nigra of the human brain II. Patterns of loss of dopamine-containing neurons in Parkinson's disease. *Brain : a journal of neurology* 1999;122(8):1437-48. doi: 10.1093/brain/122.8.1437
17. Basic principles of ROC analysis. Seminars in nuclear medicine; 1978. Elsevier.
18. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1):32-35.
19. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014, 2016.
20. Benamer HT, Patterson J, Grosset DG, et al. Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group. *Movement Disorders* 2000;15(3):503-10.
21. Group PS. A multicenter assessment of dopamine transporter imaging with DOPASCAN/SPECT in parkinsonism. *Neurology* 2000;55(10):1540-47.
22. Tolosa E, Borght TV, Moreno E. Accuracy of DaTSCAN (123I-ioflupane) SPECT in diagnosis of patients with clinically uncertain parkinsonism: 2-Year follow-up of an open-label study. *Movement Disorders* 2007;22(16):2346-51.
23. Catafau AM, Tolosa E. Impact of dopamine transporter SPECT using 123I-Ioflupane on diagnosis and management of patients with clinically uncertain parkinsonian syndromes. *Movement Disorders* 2004;19(10):1175-82.
24. Erro R, Schneider SA, Quinn NP, et al. What do patients with scans without evidence of dopaminergic deficit (SWEDD) have? New evidence and continuing controversies. *J Neurol Neurosurg Psychiatry* 2015:jnnp-2014-310256.
25. Schneider SA, Edwards MJ, Mir P, et al. Patients with adult-onset dystonic tremor resembling parkinsonian tremor have scans without evidence of dopaminergic deficit (SWEDDs). *Movement Disorders* 2007;22(15):2210-15.
26. Schwingenschuh P, Ruge D, Edwards MJ, et al. Distinguishing SWEDDs patients with asymmetric resting tremor from Parkinson's disease: a clinical and electrophysiological study. *Movement disorders* 2010;25(5):560-69.

27. Kaasinen V, Gardberg M, Röyttä M, et al. Normal dopamine transporter SPECT in neuropathologically confirmed corticobasal degeneration. *Journal of neurology* 2013;260(5):1410.
28. McKinley J, O'Connell M, Farrell M, et al. Normal dopamine transporter imaging does not exclude multiple system atrophy. *Parkinsonism Relat Disord* 2014;20:933-4.
29. O'Sullivan SS, Burn DJ, Holton JL, et al. Normal dopamine transporter single photon-emission CT scan in corticobasal degeneration. *Mov Disord* 2008;23(16):2424-6. doi: 10.1002/mds.22323 [published Online First: 2008/10/01]
30. Whone AL, Watts RL, Stoessl AJ, et al. Slower progression of Parkinson's disease with ropinirole versus levodopa: The REAL-PET study. *Annals of neurology* 2003;54(1):93-101.
31. Stoessl AJ, Halliday GM. DAT-SPECT diagnoses dopamine depletion, but not PD. *Movement Disorders* 2014;29(14):1705-06.
32. Iranzo A, Molinuevo JL, Santamaría J, et al. Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder: a descriptive study. *The Lancet Neurology* 2006;5(7):572-77.
33. Marek K, Seibyl J, Eberly S, et al. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. *Neurology* 2014;82(20):1791-97.
34. Hellwig S, Amtage F, Kreft A, et al. [18F] FDG-PET is superior to [123I] IBZM-SPECT for the differential diagnosis of parkinsonism. *Neurology* 2012;79(13):1314-22.
35. Südmeyer M, Antke C, Zizek T, et al. Diagnostic accuracy of combined FP-CIT, IBZM, and MIBG scintigraphy in the differential diagnosis of degenerative parkinsonism: a multidimensional statistical approach. *Journal of Nuclear Medicine* 2011;52(5):733-40.
36. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. *The Lancet Neurology* 2006;5(6):525-35. doi: 10.1016/S1474-4422(06)70471-9
37. Colloby SJ, McParland S, O'Brien JT, et al. Neuropathological correlates of dopaminergic imaging in Alzheimer's disease and Lewy body dementias. *Brain : a journal of neurology* 2012;135(9):2798-808.
38. Kraemmer J, Kovacs GG, Perju-Dumbrava L, et al. Correlation of striatal dopamine transporter imaging with post mortem substantia nigra cell counts. *Movement Disorders* 2014;29(14):1767-73. doi: 10.1002/mds.25975

FIGURE LEGEND

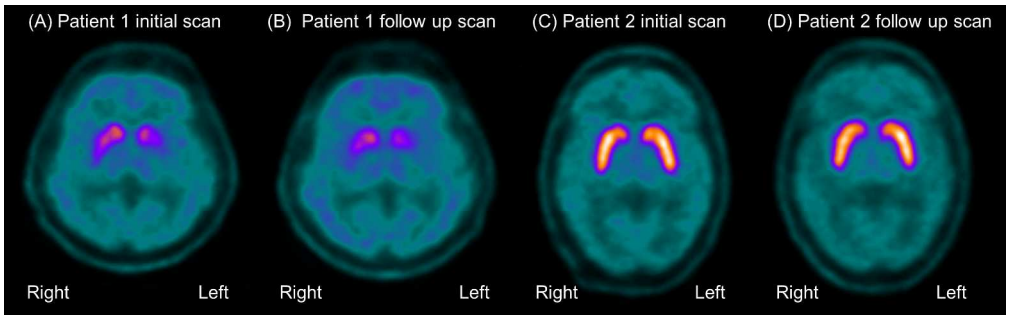
Figure 1. ROC curve of ^{18}F-AV-133 PET for predicting nigrostriatal degeneration in CUPS patients. The red line denotes the optimal binding reduction threshold of -50% as determined by the Youden criterion.

Figure 2. ^{18}F-AV-133 VMAT2 PET in two discordant cases. Patient with “rubral tremor” showing bilateral, asymmetrical reduction in tracer uptake at baseline (A) and follow-up two years later (B) with decline from -85% to –94% in the left posterior putamen. (C) Normal ^{18}F-AV-133 PET scan in a patient diagnosed with Parkinson’s disease, unchanged in two year follow-up scan (D).



ROC curve of 18F-AV-133 PET for predicting nigrostriatal degeneration in CUPS patients. The red dot denotes the optimal binding reduction threshold of 50% as determined by the Youden criterion.

233x160mm (96 x 96 DPI)



18F-AV-133 VMAT2 PET in two discordant cases. Patient with "rubral tremor" showing bilateral, asymmetrical reduction in tracer uptake at baseline (A) and follow-up two years later (B) with decline from -85% to –94% in the left posterior putamen. (C) Normal 18F-AV-133 PET scan in a patient diagnosed with Parkinson's disease, unchanged in two year follow-up scan (D).

260x80mm (300 x 300 DPI)

Supplementary table 1. Individual list of cases with baseline putamen binding on the most affected side, and the pre-scan, post-scan and follow-up diagnoses.

Age	Putamen binding (%)*	Pre-scan diagnosis	Post-scan diagnosis	Follow-up diagnosis
52	-93%	FUNCTIONAL	NDG PD	NDG PD
71	-91%	NDG PD	NDG PD	NDG PD
69	-90%	NDG PD	NDG PD	NDG PD
45	-89%	NDG UNK	NDG PD	NDG PD
70	-88%	NDG UNK	NDG PD	NDG PD
53	-88%	NDG CBS	NDG PD	NDG PD
66	-87%	NDG PD	NDG PD	NDG PD
54	-87%	NDG PD	NDG PD	NDG PD
53	-86%	NDG PD	NDG PD	NDG PD
63	-85%	NDG PSP	NDG PD	NDG PD
74	-85%	MRT	NDG PD	NDG PD
68	-85%	NDG PD	NDG PD	NDG PD
49	-85%	NDG PD	NDG PD	NDG PD
79	-85%	RUBRAL TREMOR	NDG PD	RUBRAL TREMOR
51	-84%	NDG UNK	NDG PD	NDG PD
64	-84%	FUNCTIONAL	NDG PSP	NDG PSP
65	-83%	NDG PD	NDG PD	NDG MSA
45	-83%	NDG PD	NDG PD	NDG PD
69	-82%	NDG PD	NDG PD	NDG MSA
68	-80%	NDG PD	NDG PD	NDG PD
79	-80%	NDG PD	NDG PD	NDG PD
44	-79%	NDG PD	NDG PD	NDG PD
48	-79%	NDG PD	NDG PD	NDG PD
50	-79%	DRUG INDUCED	NDG PD	NDG PD
65	-79%	VASCULAR	NDG PD	NDG PD
57	-78%	MRT	NDG PD	NDG PD
30	-78%	NDG PD	NDG PD	NDG PD
63	-78%	DYSTONIA	NDG PD	NDG PD
59	-76%	NDG UNK	NDG PD	NDG PD
62	-76%	FUNCTIONAL	NDG PD	NDG PD
46	-72%	NDG PD	NDG PD	NDG PD
58	-72%	DYSTONIA	NDG PD	NDG PD
66	-71%	NDG PD	NDG PD	NDG PD
72	-71%	FUNCTIONAL	NDG PD	NDG PD
48	-69%	NDG PD	NDG PD	NDG PD
67	-69%	DYSTONIA	NDG PD or MSA	NDG MSA
46	-62%	NDG PD	NDG PD	NDG PD
51	-60%	NDG UNK	NDG PD	NDG PD
38	-55%	DYSTONIA	NDG PD	NDG PD

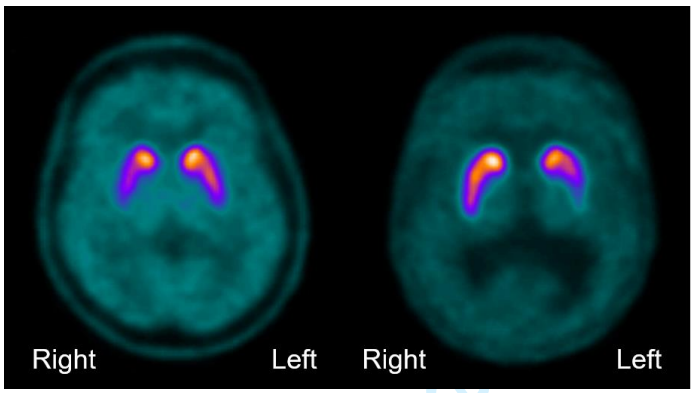
70	-54%	NDG UNK	NDG PD	NDG PD
68	-39%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
66	-39%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
53	-39%	MRT	NDG PD	NDG UNK
73**	-35%	NDG UNK	NDG PD	NDG PD
57	-27%	NDG PD	DYSTONIA	DYSTONIA
61	-25%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
63	-24%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
67	-24%	DYSTONIA	DYSTONIA	DYSTONIA
57	-23%	NDG PD	ET	ET
69	-23%	ET	ET	ET
55	-22%	ET	ET	ET
69	-20%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
51	-20%	NDG UNK	OTHER	FUNCTIONAL
61	-20%	NDG PD	NDG PD (SWEDD)	NDG PD (SWEDD)
69	-19%	DYSTONIA	DYSTONIA	DYSTONIA
47	-19%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
65**	-18%	NDG MSA	NDG MSA	NDG MSA
51	-18%	NDG PD	NDG PD (SWEDD)	NDG PD (SWEDD)
71	-16%	NDG PD	ET	ET
20	-16%	NDG (UNK)	UNK (NEUROGENETIC)	UNK (NEUROGENETIC)
68	-16%	DYSTONIA	LOST TO FOLLOW-UP	ET
61	-16%	NDG AD	NDG AD	NDG AD
68	-13%	ET	ET	ET
63	-11%	NDG PD	ET	FUNCTIONAL
43	-11%	NDG UNK	MRT	MRT
25	-10%	DYSTONIA	DYSTONIA	DYSTONIA
69	-8%	NDG PD	VASCULAR	VASCULAR
72	-6%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
38	-5%	NDG PD	FUNCTIONAL	FUNCTIONAL
60	-4%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
54	-3%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
37	-2%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
30	-2%	NDG PD	FUNCTIONAL	FUNCTIONAL
53	2%	DYSTONIA	DYSTONIA	NDG PD
67	5%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
63	5%	NDG PD	FUNCTIONAL	FUNCTIONAL
28	7%	NDG PD	UNKNOWN	UNKNOWN
25	19%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
51	41%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
51	67%	NDG PD	DRUG INDUCED	DRUG INDUCED
43	74%	DYSTONIA	FUNCTIONAL	FUNCTIONAL

Green background denotes cases where clinical diagnosis was concordant with scan result.

Red background denotes cases where clinical diagnosis was discordant with scan result.

* Putamen binding on the most affected side.

** Visual assessment of scan demonstrated significant and asymmetrical reduction in posterior putamen binding compared to caudate binding



NDG = Neurodegenerative

PD = Parkinson's disease

MSA = multiple system atrophy

CBS = Corticobasal syndrome

AD = Alzheimer's disease

MRT = Monosymptomatic resting tremor

UNK = Unknown

ET = Essential tremor

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4, 5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	4, 5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4
	9	Whether participants formed a consecutive, random or convenience series	4
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6, 7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6, 7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6, 7
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6, 7
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	7
	15	How indeterminate index test or reference standard results were handled	7, 8
	16	How missing data on the index test and reference standard were handled	7, 8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7, 8
	18	Intended sample size and how it was determined	7
RESULTS			
Participants	**19**	Flow of participants, using a diagram	N/A
	20	Baseline demographic and clinical characteristics of participants	9
	21a	Distribution of severity of disease in those with the target condition	9
	21b	Distribution of alternative diagnoses in those without the target condition	9
	22	Time interval and any clinical interventions between index test and reference standard	9
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	10
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	10
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	14
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	7
	30	Sources of funding and other support; role of funders	16

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.

BMJ Open

Diagnostic Accuracy of Imaging Brain Vesicular Monoamine Transporter type-2 (VMAT2) in Clinically Uncertain Parkinsonian Syndrome (CUPS): a 3 year follow-up study in community patients.

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2018-025533.R1
Article Type:	Research
Date Submitted by the Author:	30-Aug-2018
Complete List of Authors:	Xu, San San; Austin Health, Department of Molecular Imaging and Therapy; Austin Health, Department of Neurology Alexander, Paschal; Austin Health, Department of Molecular Imaging and Therapy Lie, Yenni; Austin Health, Department of Molecular Imaging and Therapy Dore, Vincent; Austin Health, Department of Molecular Imaging and Therapy; The Commonwealth Scientific and Industrial Research Organisation, Health & Biosecurity Flagship, The Australian eHealth Research Centre Bozinovski, Svetlana; Austin Health, Department of Molecular Imaging and Therapy Mulligan, Rachel; Austin Health, Department of Molecular Imaging and Therapy Young, Kenneth; Austin Health, Department of Molecular Imaging and Therapy Villemagne, Victor; Austin Health, Department of Molecular Imaging and Therapy; Florey Institute of Neuroscience and Mental Health Rowe, Christopher C.; Austin Health, Department of Molecular Imaging and Therapy; Florey Institute of Neuroscience and Mental Health
Primary Subject Heading:	Neurology
Secondary Subject Heading:	Radiology and imaging, Diagnostics
Keywords:	Diagnostic accuracy, Molecular imaging, Parkinson-s disease < NEUROLOGY, PET, VMAT2

TITLE PAGE

Diagnostic Accuracy of Imaging Brain Vesicular Monoamine Transporter type-2 (VMAT2) in Clinically Uncertain Parkinsonian Syndrome (CUPS): a 3 year follow-up study in community patients.

Authors:

San San Xu[1,2], Paschal K Alexander[1], Yenni Lie[1], Vincent Dore[1, 3, 4], Svetlana Bozinovski[1], Rachel S. Mulligan[1], Kenneth Young[1], Victor L. Villemagne[1,3], Christopher C. Rowe[1,3].

1. Department of Molecular Imaging and Therapy, Austin Health, Melbourne, Victoria, Australia.
2. Department of Neurology, Austin Health, Melbourne Victoria, Australia
3. Florey Institute of Neuroscience and Mental Health, University of Melbourne, Melbourne, Victoria, Australia.
4. Health & Biosecurity Flagship, The Australian eHealth Research Centre, The Commonwealth Scientific and Industrial Research Organisation, Melbourne, Victoria, Australia

Corresponding author:

San San Xu,

Neurologist, Department of Neurology, Austin Heath,

145 Studley Road, Heidelberg, Victoria, 3084, Australia

Telephone: +61400851823, +61-3-9496 5529

Fax: +61-3-9496 4065

Email: sansan.xu@austin.org.au

Word count: 2909

Abstract word count: 300

Number of references: 38

Keywords: Diagnostic accuracy; molecular imaging; Parkinson's disease; PET; VMAT2

ABSTRACT

Objectives: To further validate the diagnostic utility of ^{18}F-AV-133 Vesicular monoamine transporter type-2 (VMAT2) PET in patients with Clinically Uncertain Parkinsonian Syndromes (CUPS) by comparison to clinical diagnosis at 3 years follow-up.

Design, setting and participants: In a previous study, we reported that ^{18}F-AV-133 PET in community CUPS patients changed diagnosis and management and increased diagnostic confidence. The current diagnosis of this cohort was obtained from the patient and treating specialist and compared to the diagnosis suggested three years earlier by the ^{18}F-AV-133 PET. A second ^{18}F-AV-133 PET was available in those with a discordant or inconclusive final diagnosis.

Study outcome measures: The primary endpoint was the proportion of patients who had a follow-up clinical diagnosis, which was concordant with their initial ^{18}F-AV-133 PET scan. Secondary endpoints were the proportion of patients who had the same diagnosis at follow-up as that reached after the initial scan and the stability of diagnostic changes made after the first scan.

Results: 81 of the 85 patients previously recruited to the CUPS study had follow-up of which 79 had a clinical diagnosis and two remained CUPS. The diagnosis was in agreement with the initial ^{18}F-AV-133 PET scan result in 74 cases. Five patients had a discordant diagnosis; one patient with rubral tremor had a severely abnormal scan that had worsened when rescanned; four cases with normal initial and repeat scans had a clinical diagnosis of Parkinson's disease. Two patients with suspected genetic disorders remained classified as CUPS and both had normal scans. In the 24 CUPS cohort patients where ^{18}F-AV-133 PET initially changed diagnosis, this change was supported by follow-up diagnosis in all but the one rubral tremor case.

Conclusion: ^{18}F-AV-133 PET is a useful tool in improving diagnostic accuracy in CUPS providing results and diagnostic changes that remain robust after 3 years follow-up.

STRENGTHS AND LIMITATIONS

- This is a 3 year follow up study evaluating the current clinical diagnosis of patients with Clinically Uncertain Parkinsonian Syndromes (CUPS) who have previously had a ^{18}F-AV-133 PET.
- The aim of the study is to evaluate the diagnostic accuracy and validate the optimal binding reduction threshold of 50% for abnormal scans.
- The final diagnosis was nominated by the treating specialist and patient after a period of clinical follow up.
- The follow-up clinical diagnosis was compared to the diagnosis suggested by the ^{18}F-AV-133 PET and the initial clinical diagnosis reached after the scan.
- A repeat ^{18}F-AV-133 PET was reviewed in patients who had a current clinical diagnosis that was discordant with the scan result or those with an unknown diagnosis.

MANUSCRIPT

INTRODUCTION

Parkinson's disease (PD) is a common neurodegenerative condition, second only to Alzheimer's disease and the most prevalent of the Parkinsonian syndromes. Diagnostic certainty of Parkinson's disease and the other Parkinsonian syndromes (multiple system atrophy (MSA)[1], progressive supranuclear palsy (PSP)[2] and cortico-basal syndrome (CBS)[3] can only be confirmed by histopathological demonstration of the characteristic pathology and resultant nigrostriatal degeneration. In clinical practice, diagnosis of Parkinson's disease relies on the presence of bradykinesia and at least one of rest tremor, rigidity or postural instability[4]. Atypical or mild clinical features may delay diagnosis and introduction of appropriate therapies. In a tertiary movement disorders centre, the diagnostic accuracy of clinical assessment in post-mortem clinicopathologically confirmed Parkinson's disease did improve from 79% to 90% over a 10 year period[5 6]. However, the rate of misdiagnosis is likely higher in early disease and in the primary care setting. In community patients with Parkinsonian features or on anti-parkinsonian medications, only 53% - 83% of patients fulfilled the criteria for probable Parkinson's disease[7 8]. Other conditions that may mimic Parkinson's disease include essential tremor, dystonia, drug induced parkinsonism (DIP), vascular parkinsonism and functional movement disorder. Misdiagnosis of these disorders can have significant prognostic and management implications.

Positron emission tomography (PET) and single photon emission computerised tomography (SPECT) imaging can accurately evaluate the nigrostriatal system and aid in early diagnosis of Parkinson's disease[9]. Vesicular monoamine transporter type 2 (VMAT2) plays an integral role in pre-synaptic dopamine uptake and storage and is a reliable marker of nigrostriatal terminal integrity [10 11]. ^{18}F-AV-133 is a novel ^{18}F–labelled dihydrotetrabenazine analogue that selectively binds to VMAT2 with high affinity and allows for in vivo evaluation of VMAT2

density. Several clinical studies have demonstrated the feasibility of ^{18}F-AV-133 PET technique to assist in the diagnosis of Parkinson's disease and dementia with Lewy bodies[10-14]. ^{18}F-AV-133 PET has multiple potential advantages compared to dopamine transporter SPECT, including improved image quality and quantification, reduced tracer administration to scan interval time and reduced scan duration. Further, there is no requirement for pre-scan thyroid blockade in contrast to the iodine-123 labelled SPECT dopamine transporter tracers.

In a previous study, we investigated the management impact of ^{18}F-AV-133 PET imaging in patients with Clinically Uncertain Parkinsonian Syndromes (CUPS)[14]. The results of the ^{18}F-AV-133 PET altered diagnosis in 23% of participants (11 of 47) and changed management in more than half of the cases (53%; 25 of 47). Furthermore, diagnostic confidence in clinicians increased in 74% of the participants after the scan, regardless of whether the result was normal or abnormal. Total enrolment in the CUPS study subsequently reached 85 and the present study aims to further confirm the diagnostic accuracy of ^{18}F-AV-133 PET by comparing the results of the initial scan with the clinical diagnosis at 3 years follow-up in the total cohort.

METHODS

Study subjects

All patients previously recruited in our CUPS study were eligible for the current study[14]. Patients with CUPS were recruited from the private and public clinics of movement disorders specialists from across the city of Melbourne, Australia. The criteria for uncertainty of diagnosis was at the discretion of the referring clinician and included the presence of atypical features of parkinsonism including poor levodopa responsiveness, lack of disease progression, dystonia and young age of onset. Participants were excluded if they had a

history of malignancy within the last 5 years or if they were unable to provide informed written consent.

Patient and public involvement

Patients and the public were not involved in the design or analysis of this study. Once published, the results of the study will be summarised in a letter and disseminated to the participants and doctors involved in patient recruitment.

Study design

This was a single-centre, prospective experimental study with a mean follow-up interval of 3 years ± 6 months (range: 18 – 68 months). The study was conducted in accordance with local and international standards and approved by the Austin Health Human Research Ethics Committee. All participants provided written consent prior to their inclusion in the study. At the time of follow-up, phone calls were made to the treating movement disorders specialist and study participant to establish the current diagnosis for the patient. If the treating neurologist had changed during the follow-up period, the diagnosis was made by the most recent clinician involved in the care of the participant. The clinicians had access to the initial ^{18}F-AV-133 PET scan results. The diagnostic categories were classified into parkinsonism with nigrostriatal degeneration (including idiopathic Parkinson's disease, multiple system atrophy, progressive supranuclear palsy and cortico-basal syndrome), other (including essential tremor, dystonia, drug induced parkinsonism, functional (psychogenic), monosymptomatic resting tremor) or an unclear diagnosis i.e. remained CUPS. Follow-up diagnosis was considered in agreement with the initial PET scan diagnosis if it remained in the same binary diagnostic category i.e. parkinsonism with nigrostriatal degeneration or other. For example, if a participant had a change of diagnosis from Parkinson's disease to multiple system atrophy during the follow-up period, this was still considered a concordant

result. All participants had a follow-up AV-133 PET but only those for patients who had a clinical diagnosis which was discordant to the initial PET result are reported here.

PET scan protocol and image analysis

As previously described, a 20-minute emission PET scan was obtained two hours after intravenous injection of approximately 250 MBq of ^{18}F-AV133[11 14]. For attenuation-correction purposes, a rotation transmission sinogram in 3D mode with a single ^{137}Cs point source was acquired prior to radiotracer injection. The final images were reconstructed using a 3D row-action maximum-likelihood algorithm. The regional tracer binding of the caudate nucleus, anterior and posterior putamen were calculated using the ratio of regional activity to primary visual cortex, the latter being a region devoid of monoaminergic terminals, and therefore suitable as a reference standard. Each individual image was spatially normalised to a normal AV-133 template using statistical parametric mapping software (SPM8; Wellcome Department of Cognitive Neurology). A standard region of interest (ROI) template was previously constructed manually over 13 slices for the caudate and 8 slices for the putamen (each slice 2mm thick). The putamen ROI was bisected to give anterior and posterior putamen binding. Abnormal images were determined quantitatively and visually. Quantitatively, abnormal images were defined as those with a greater than 50% reduction in binding in the most affected posterior putamen, which corresponds to 4 standard deviations below the mean of the healthy control reference group[11 14] that consisted of 16 healthy controls; 9 males and 7 females, with a mean age of 72 +/- 5.1 (range 64 – 78 years). Scans were called visually abnormal when there was significant asymmetry in the posterior putamen or marked reduction in uptake in the putamen relative to the caudate nucleus. In two cases, visual analysis was abnormal when quantitative results were not but otherwise all classifications were concordant. The binding percentage for each region was calculated by subtracting the regional control group mean binding ratio from the patient result, then dividing this by the control group mean and then multiplying by one hundred. This threshold

is based on multiple histopathological studies which suggest that motor symptoms of Parkinson's disease only emerge after at least 50% loss of dopaminergic neurons in the substantia nigra [15 16]. The 3 year follow-up data was also used to test the validity of the 50% binding reduction threshold by comparison to ROC curve using the Youden criterion.

Outcome variables

The primary endpoint of the study was the proportion of patients who had a clinical diagnosis at 3 years follow-up, which was in agreement with the results of their initial ^{18}F-AV-133 PET scan. Secondary endpoints were the proportion of patients who had the same diagnosis in the follow-up period as that reached after the initial scan and the stability of diagnostic changes made after the first scan.

Statistical analysis

The results of the study are expressed as a mean ± standard deviation with accuracy figures derived from two by two contingency tables. The diagnostic accuracy of ^{18}F-AV-133 PET was further investigated using a receiver operator characteristics (ROC) analysis and area under the curve (AUC)[17], with binding reduction threshold determined using the Youden criterion[18]. Data processing and statistical analysis was conducted using Microsoft ® Excel ® 2016 software, Minitab 18 (Minitab Inc., Pennsylvania, USA) and R Version 3.4.3[19].

Data sharing

No additional data is available for sharing.

RESULTS

Study population

81 of the 85 participants (95%) enrolled in our CUPS study (47 previously described in Alexander et al, 2017[11 14]) were reviewed after 3 years. Four patients were lost to follow-up. Patient characteristics are outlined in Table 1. There were no significant differences when comparing participants involved in the study with those who were lost to follow-up.

Table 1: Patient characteristics at baseline in the follow-up and lost to follow-up groups

	Follow-up	Lost to follow-up
Demographics		
n	81	4
Age (mean ± SD)	57 ± 13.1	57 ± 17.0
Female	44 (54%)	2 (50%)
UPDRS Motor score mean ± SD	10.3 ± 4.2	8.4 ± 5.9
Hoehn and Yahr Score stage average	1.6 ± 0.76	1.1 ± 0.25
MMSE mean ± SD	28.9 ± 1.6	29.0 ± 1.4
Scan results		
Abnormal ^{18}F-AV-133 Scan	42 (52%)	2 (50%)
Baseline Pre-scan Diagnosis		
Neurodegenerative conditions	**45**	**2**
Parkinson's disease	31	1
Multiple system atrophy	1	
Progressive supranuclear palsy	1	
Corticobasal syndrome	1	
Alzheimer's disease	1	
Undefined	10	1
Non-degenerative conditions	**36**	**2**
Functional	13	2
Dystonia	10	0
Drug induced parkinsonism	5	0
Essential tremor	3	0
Monosymptomatic resting tremor	3	0
Vascular parkinsonism	1	0
Rubral tremor	1	0

UPDRS = Unified Parkinson's Disease Rating Scale

SD = Standard Deviation

Diagnosis summary

Of the 81 participants followed up, 79 (98%) had a specific clinical diagnosis and two cases had an inconclusive diagnosis i.e. remained CUPS. When a specific clinical diagnosis was available, the follow-up diagnosis was concordant with that suggested by the initial ^{18}F-AV-133 PET scan in 74 cases (94%), with a positive predictive value of 98%, a negative predictive value of 89% and a sensitivity and specificity of 91% and 97% respectively (derived from Table 2). The diagnostic agreement rate was slightly lower in the parkinsonism associated with nigrostriatal degeneration category compared to the "other" category (91% vs 97%). The agreement rate of clinical diagnosis at follow-up was 67% when compared to the most likely pre-scan clinical diagnosis and was 97% when compared to the initial post-scan diagnosis. The diagnostic accuracy of ^{18}F-AV-133 PET in predicting Parkinsonism with nigrostriatal degeneration was further evaluated with a ROC curve (Figure 1). The calculated AUC was 0.94 (95% confidence interval 0.88 - 0.99) with an optimal binding reduction threshold of 50% as per Youden criterion. The clinical diagnosis of all participants in the pre-scan, post-scan and follow-up period are listed individually in supplementary Table 1.

Table 2: Correlation between AV-133 PET scan result and diagnosis after 3 years

AV-133 PET scan result	**Diagnosis at 3 year follow-up**		
	Parkinsonism with nigrostriatal degeneration	Other diagnosis	Inconclusive diagnosis
Abnormal	41	1	0
Normal	4	33	2
Total	45	34	2

Discordant clinical diagnosis with imaging results

Five patients (6%) had a follow-up clinical diagnosis that did not concur with the results of their ^{18}F-AV-133 PET scan (Table 3). One participant with an abnormal scan was diagnosed

pre scan as rubral tremor, post scan as PD and at follow-up diagnosis had returned to rubral tremor despite worsening of the scan (Figure 2). This patient has an asymmetrical, resting, action and postural upper limb tremor that is levodopa responsive but has remained relatively stable for 30 years with no bradykinesia or rigidity. Brain MRI did not reveal a structural lesion that accounted for reduced AV-133 uptake in the absence of nigrostriatal degeneration.

Three participants with a follow-up clinical diagnosis of Parkinson's disease had a normal ^{18}F-AV-133 PET scan. In two of these patients the managing clinician now considers them as having PD phenotype due to Symptoms Without Evidence of Dopaminergic Deficit (SWEDD). Follow-up imaging in these two patients 2 years later remained stable (Figure 2 shows one of these cases). One patient had an initial pre-scan and post-scan diagnosis of dystonia. During the follow-up period, the diagnosis was revised to PD in the context of emerging bradykinesia and a good response to levodopa. Follow-up imaging at three years remained normal.

One patient had a follow-up diagnosis of progressive atypical parkinsonian syndrome. The initial and repeat ^{18}F-AV-133 PET scan two years later showed stable and symmetrical binding in the lower range but less than the 50% reduced cut-off (posterior putamen binding of -39% and -34% respectively).

Table 3: Scan results and diagnosis of patients with clinical diagnosis discordant to scan results or unknown

Case	Age	UPDRS	H& Y	Post scan diagnosis	3 year follow-up diagnosis	Posterior putamen binding	Follow-up posterior putamen binding	Putamen to caudate ratio	Left-right asymmetry ratio	Time between scans (months)
1	79	6	1	NDG PD	Rubral tremor	-85%	-94%	0.31	0.46	24
2*	51	9	1	NDG PD	NDG PD	-18%	-12%	1.21	0.98	27
3*	61	7	2	NDG PD	NDG PD	-20%	-16%	1.20	0.91	29
4	53	4	1	Dystonia	NDG PD	2%	-14%	1.12	0.98	36
5	53	5	1	NDG PD	NDG AP	-39%	-34%	1.21	0.97	26
6	20	6	1	UNK	UNK (Neurogenetic)	-16%	N/A	1.14	0.97	N/A
7	28	8	1.5	UNK	UNK (Dystonia parkinsonism syndrome)	9%	-24%	1.18	1.02	19

Putamen binding figures represent most affected side. Putamen to caudate ratio and asymmetry ratio represent posterior putamen results from baseline scans.

*Cases considered consistent with SWEDD (Scans without evidence of dopaminergic deficit) by treating specialist at follow-up.

UPDRS = Unified Parkinson's Disease Rating Scale

H & Y = Hoehn and Yahr

NDG = Neurodegenerative

PD = Parkinson's disease

AP = Atypical parkinsonism

UNK = Unknown

N/A = Not available

Cases with uncertain diagnosis

Two cases continue to be CUPS. One participant, a young male, had a pre-scan diagnosis of an unknown neurodegenerative condition. He had a normal ^{18}F-AV-133 PET scan and the post scan diagnosis remained unclear. The patient's symptoms have progressed but the current diagnosis remains "undefined neurogenetic condition". The patient did not return for repeat ^{18}F-AV-133 PET imaging. The second participant also had a normal initial ^{18}F-AV-133 PET scan. The pre-scan diagnosis was Parkinson's disease and the immediate post scan diagnosis was unclear. At follow-up, the participant clinically presents with a dystonia parkinsonism syndrome. A second ^{18}F-AV-133 PET scan 1.5 years after the initial study remained in the normal range but showed a decline in posterior putamen tracer binding from 9% to-24%. This participant has a sibling with early onset Parkinson's disease who had an abnormal ^{18}F-AV-133 PET.

Follow-up After Pre-scan to Post-scan Diagnostic Change.

Of the 81 CUPS with 3 year follow-up, 24 had a change in binary diagnostic classification due to the initial ^{18}F-AV-133 PET (see supplementary Table 1 for details). Of these, the current clinical diagnosis remained the same as the post scan diagnosis in 23 (96%). The patient with an original diagnosis of rubral tremor was re-classified as Parkinson's disease following an initial abnormal ^{18}F-AV-133 PET but the diagnosis had reverted back to rubral tremor at follow-up.

DISCUSSION

Our study provides further evidence that ^{18}F-AV-133 PET is a feasible adjunctive tool in the diagnosis of degenerative parkinsonism. The 3 year follow-up data validated the 50% binding reduction threshold and clinical diagnosis remained remarkably concordant with the

results of the ^{18}F-AV-133 PET scan (94% agreement rate) with an impressive sensitivity and specificity of 91% and 97% respectively. This is highly comparable to the sensitivity (87-98%) and specificity (80-100%) reported in the widely used ^{123}I-FP-CIT SPECT scans (DaTSCAN) [20-23]. When a follow-up diagnosis had been made, the diagnosis was in agreement with the pre-scan diagnosis in only 67% of cases, but was in agreement with the post-scan diagnosis made with the aid of the ^{18}F-AV-133 PET scan in 97%, highlighting the diagnostic challenges in early Parkinson's disease.

Four patients had a clinical diagnosis of Parkinson's disease or atypical parkinsonism despite a normal ^{18}F-AV-133 PET scan. This discrepancy has been widely reported with other measures of dopaminergic integrity and is referred to as Symptoms Without Dopaminergic Deficit (SWEDD). The number of SWEDD cases in the literature has been described to be between 3.5% to 20% in patients with clinical features of Parkinson's disease undergoing DAT scanning[24]. However, the entity of SWEDD remains controversial and may characterise a heterogenous group of conditions. Some studies have suggested that the tremor dominant presentation may represent adult onset dystonia[25 26]. However, pathologically confirmed cases of multiple system atrophy and corticobasal ganglionic degeneration with normal DaTSCAN have been described in the literature[27-29]. A patient with levodopa responsive parkinsonism and dyskinesia and a G2019S mutation in the LRRK2 gene with a normal ^{18}F-fluorodopa PET scan[30] has been reported, suggesting that imaging of the nigrostriatal pathway may be normal in some cases of early Parkinson's disease[31]. In the current study, an abnormal scan was defined as a $\geq$ 50% reduction in tracer binding in the posterior putamen compared to healthy controls. This threshold is based on post mortem studies suggesting that the clinical features of parkinsonism emerge after > 50% loss of dopaminergic neurons in the substantia nigra[15 16]. Therefore, it is feasible that this prescribed range will miss preclinical or early PD with very mild motor symptoms. This is supported by the observation of progressive clinical and AV-133 binding decline in our patient with a suspected genetic dystonia parkinsonism syndrome. Further studies evaluating ^{18}F-AV-133

PET in at risk patients such as those with REM sleep behaviour disorder[32] may shed light on the reasons for false negatives and could help identify an appropriate threshold for detection of preclinical individuals. Our 2 SWEDD cases had posterior putamen binding of -16.5% and -17% and demonstrated no decline with repeat scans after 2 years. Consistent with the literature of SWEDD, these subjects did not show clinical progression over the follow-up period[24 33]. The one patient who was reclassified from dystonia to Parkinson's disease at follow-up had minimal decline in the scans from a posterior putamen binding of 3% to -14% after 2 years.

There were three patients who had a change in diagnosis from Parkinson's disease to multiple system atrophy during the follow-up period. This is consistent with the literature, that suggests that pre-synaptic dopaminergic imaging cannot differentiate Parkinson's disease from atypical parkinsonian syndromes[9]. Evaluation of the postsynaptic dopaminergic systems with D2 receptor binding ligands or metabolic imaging with ^{18}F-FDG-PET has been suggested to further differentiate idiopathic Parkinson's disease from atypical parkinsonism[34 35].

There are advantages of ^{18}F-AV-133 PET including improved spatial resolution and there is reduced tracer uptake and scan acquisition time in comparison to DaTSCAN, without the need for thyroid blockade or other patient preparation. In addition to its role in diagnosis, the ^{18}F-AV-133 PET may prove to be a valuable tool for disease monitoring and in patient selection and evaluation of the therapeutic impact of interventions in clinical trials.

There are limitations to the current study. Firstly, the definitive diagnosis of Parkinson's disease relies on histopathological evidence[5] and this is not available in this cohort at this time. However, in view of the logistical challenges of post mortem studies, a clinical diagnosis such as that outlined by the UK Brain Bank criteria is commonly accepted as a substitute gold standard[4]. Additionally, a long clinical follow-up period, such as the one

employed in this study, has been reported to improve diagnostic accuracy[36]. Post-mortem studies of autopsy confirmed Parkinson's disease have revealed a correlation between ligand uptake in ^{123}I-FP-CIT SPECT and nigrostriatal neuronal loss[37 38]. Similar histopathological studies would be worthwhile to further validate the diagnostic accuracy of ^{18}F-AV-133 PET.

CONCLUSION

This study validates and extends the findings of our previous CUPS study, providing further evidence of the diagnostic value of ^{18}F-AV-133 PET, with a robust impact after 3 years of follow-up indicating that management change initiated by ^{18}F-AV-133 PET scan findings, was and remains appropriate.

ACKNOWLEDGMENTS

We wish to thank Drs Andrew Hughes, Andrew Evans, David Williams, John Merory, Katya Kotchet, Katrina Reardon, Richard Peppard, Malcolm Horne and Dominic Thyagarajan for their role in patient recruitment.

AUTHOR CONTRIBUTORSHIP STATEMENT

Dr San San Xu was involved in statistical analysis and wrote the manuscript. Dr Paschal K Alexander, Dr Yenni Lie and Dr Vincent Dore were involved in the research project execution and statistical analysis. Ms Svetlana Bozinvski was involved with research project organisation and execution. Ms Rachel S Mulligan and Mr Kenneth Young were involved in research project execution. Dr Victor L. Villemagne was involved in statistical analysis design, execution, review and critique. Professor Christopher C. Rowe was involved in

research project conception, organisation and execution, statistical analysis review and critique and manuscript review and critique.

FINANCIAL DISCLOSUES AND CONFLICT RELATED TO MANUSCRIPT

A grant for this study was provided to Professor Rowe to support this study by Avid Radiopharmaceuticals, who developed and own the rights to ^{18}F-AV-133.

FUNDING SOURCES

This was an investigator-initiated study supported by a grant from Avid Radiopharmaceuticals.

FINANCIAL DISCLOSURES OF ALL AUTHORS UNRELATED TO CURRENT RESEARCH

Professor Christopher Rowe has received research grants for imaging in dementia from Piramal Imaging, GE Healthcare, Cerveau, Astra Zeneca, Biogen and Navidea. He has been a consultant or paid speaker at sponsored conference sessions for Piramal Imaging, GE Healthcare, Astra Zeneca, Roche and Biogen. Dr Victor Villemagne has been a consultant or paid speaker at sponsored conference sessions for Piramal Imaging, GE Healthcare, Astra Zeneca and Novartis.

REFERENCES

1. Wenning GK, Colosimo C, Geser F, et al. Multiple system atrophy. *The Lancet Neurology* 2004;3(2):93-103. [published Online First: 2004/01/30]
2. Hauw J-J, Daniel S, Dickson D, et al. Preliminary NINDS neuropathologic criteria for Steele-Richardson-Olszewski syndrome (progressive supranuclear palsy). *Neurology* 1994;44(11):2015-15.
3. Lee SE, Rabinovici GD, Mayo MC, et al. Clinicopathological correlations in corticobasal degeneration. *Annals of neurology* 2011;70(2):327-40.
4. Jankovic J. Parkinson's disease: clinical features and diagnosis. *Journal of Neurology, Neurosurgery & Psychiatry* 2008;79(4):368-76.
5. Hughes AJ, Daniel SE, Kilford L, et al. Accuracy of clinical diagnosis of idiopathic Parkinson's disease: a clinico-pathological study of 100 cases. *Journal of Neurology, Neurosurgery & Psychiatry* 1992;55(3):181-84.
6. Hughes AJ, Daniel SE, Lees AJ. Improved accuracy of clinical diagnosis of Lewy body Parkinson's disease. *Neurology* 2001;57(8):1497-99.
7. Meara J, Bhowmick BK, Hobson P. Accuracy of diagnosis in patients with presumed Parkinson's disease. *Age and ageing* 1999;28(2):99-102.
8. Schrag A, Ben-Shlomo Y, Quinn N. How valid is the clinical diagnosis of Parkinson's disease in the community? *Journal of Neurology, Neurosurgery & Psychiatry* 2002;73(5):529-34.
9. Politis M. Neuroimaging in Parkinson disease: from research setting to clinical practice. *Nature Reviews Neurology* 2014;10(12):708-22.
10. Frey KA, Koeppe RA, Kilbourn MR, et al. Presynaptic monoaminergic vesicles in Parkinson's disease and normal aging. *Annals of neurology* 1996;40(6):873-84.
11. Okamura N, Villemagne VL, Drago J, et al. In vivo measurement of vesicular monoamine transporter type 2 density in Parkinson disease with 18F-AV-133. *Journal of Nuclear Medicine* 2010;51(2):223-28.
12. Villemagne VL, Okamura N, Pejoska S, et al. Differential diagnosis in Alzheimer's disease and dementia with Lewy bodies via VMAT2 and amyloid imaging. *Neurodegenerative Diseases* 2012;10(1-4):161-65.
13. Hsiao IT, Weng YH, Lin WY, et al. Comparison of 99mTc-TRODAT-1 SPECT and 18 F-AV-133 PET imaging in healthy controls and Parkinson's disease patients. *Nuclear medicine and biology* 2014;41(4):322-9. doi: 10.1016/j.nucmedbio.2013.12.017 [published Online First: 2014/02/08]

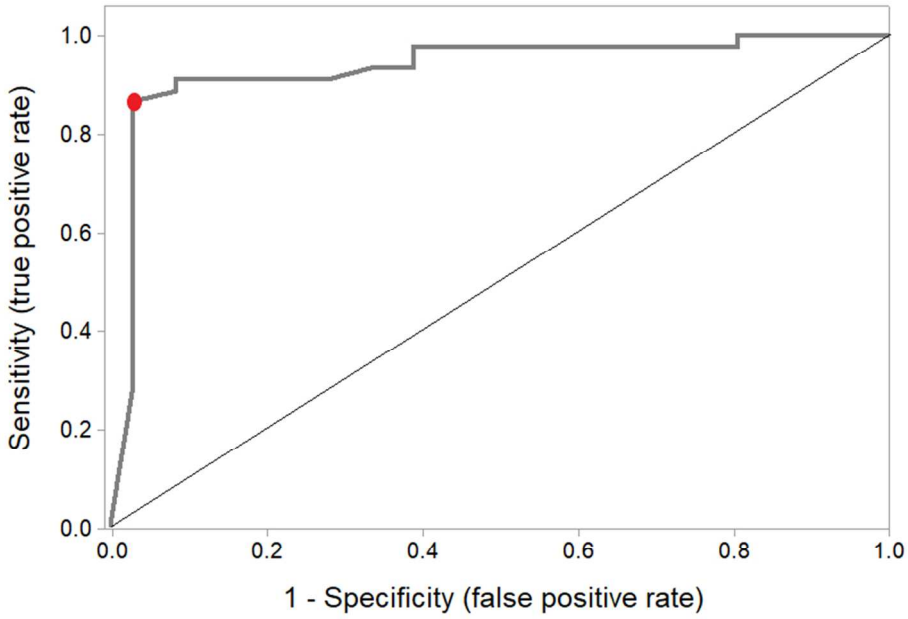
14. Alexander PK, Lie Y, Jones G, et al. Management impact of imaging brain vesicular monoamine transporter type-2 (VMAT2) in clinically uncertain parkinsonian syndrome (CUPS) with 18F-AV133 and PET. *Journal of Nuclear Medicine* 2017:jnumed. 116.189019.
15. Kordower JH, Olanow CW, Dodiya HB, et al. Disease duration and the integrity of the nigrostriatal system in Parkinson's disease. *Brain : a journal of neurology* 2013;136(8):2419-31.
16. Damier P, Hirsch EC, Agid Y, et al. The substantia nigra of the human brain II. Patterns of loss of dopamine-containing neurons in Parkinson's disease. *Brain : a journal of neurology* 1999;122(8):1437-48. doi: 10.1093/brain/122.8.1437
17. Basic principles of ROC analysis. Seminars in nuclear medicine; 1978. Elsevier.
18. Youden WJ. Index for rating diagnostic tests. *Cancer* 1950;3(1):32-35.
19. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2014, 2016.
20. Benamer HT, Patterson J, Grosset DG, et al. Accurate differentiation of parkinsonism and essential tremor using visual assessment of [123I]-FP-CIT SPECT imaging: The [123I]-FP-CIT study group. *Movement Disorders* 2000;15(3):503-10.
21. Group PS. A multicenter assessment of dopamine transporter imaging with DOPASCAN/SPECT in parkinsonism. *Neurology* 2000;55(10):1540-47.
22. Tolosa E, Borght TV, Moreno E. Accuracy of DaTSCAN (123I-ioflupane) SPECT in diagnosis of patients with clinically uncertain parkinsonism: 2-Year follow-up of an open-label study. *Movement Disorders* 2007;22(16):2346-51.
23. Catafau AM, Tolosa E. Impact of dopamine transporter SPECT using 123I-Ioflupane on diagnosis and management of patients with clinically uncertain parkinsonian syndromes. *Movement Disorders* 2004;19(10):1175-82.
24. Erro R, Schneider SA, Quinn NP, et al. What do patients with scans without evidence of dopaminergic deficit (SWEDD) have? New evidence and continuing controversies. *J Neurol Neurosurg Psychiatry* 2015:jnnp-2014-310256.
25. Schneider SA, Edwards MJ, Mir P, et al. Patients with adult-onset dystonic tremor resembling parkinsonian tremor have scans without evidence of dopaminergic deficit (SWEDDs). *Movement Disorders* 2007;22(15):2210-15.
26. Schwingenschuh P, Ruge D, Edwards MJ, et al. Distinguishing SWEDDs patients with asymmetric resting tremor from Parkinson's disease: a clinical and electrophysiological study. *Movement disorders* 2010;25(5):560-69.

27. Kaasinen V, Gardberg M, Röyttä M, et al. Normal dopamine transporter SPECT in neuropathologically confirmed corticobasal degeneration. *Journal of neurology* 2013;260(5):1410.
28. McKinley J, O'Connell M, Farrell M, et al. Normal dopamine transporter imaging does not exclude multiple system atrophy. *Parkinsonism Relat Disord* 2014;20:933-4.
29. O'Sullivan SS, Burn DJ, Holton JL, et al. Normal dopamine transporter single photon-emission CT scan in corticobasal degeneration. *Mov Disord* 2008;23(16):2424-6. doi: 10.1002/mds.22323 [published Online First: 2008/10/01]
30. Whone AL, Watts RL, Stoessl AJ, et al. Slower progression of Parkinson's disease with ropinirole versus levodopa: The REAL-PET study. *Annals of neurology* 2003;54(1):93-101.
31. Stoessl AJ, Halliday GM. DAT-SPECT diagnoses dopamine depletion, but not PD. *Movement Disorders* 2014;29(14):1705-06.
32. Iranzo A, Molinuevo JL, Santamaría J, et al. Rapid-eye-movement sleep behaviour disorder as an early marker for a neurodegenerative disorder: a descriptive study. *The Lancet Neurology* 2006;5(7):572-77.
33. Marek K, Seibyl J, Eberly S, et al. Longitudinal follow-up of SWEDD subjects in the PRECEPT Study. *Neurology* 2014;82(20):1791-97.
34. Hellwig S, Amtage F, Kreft A, et al. [18F] FDG-PET is superior to [123I] IBZM-SPECT for the differential diagnosis of parkinsonism. *Neurology* 2012;79(13):1314-22.
35. Südmeyer M, Antke C, Zizek T, et al. Diagnostic accuracy of combined FP-CIT, IBZM, and MIBG scintigraphy in the differential diagnosis of degenerative parkinsonism: a multidimensional statistical approach. *Journal of Nuclear Medicine* 2011;52(5):733-40.
36. de Lau LM, Breteler MM. Epidemiology of Parkinson's disease. *The Lancet Neurology* 2006;5(6):525-35. doi: 10.1016/S1474-4422(06)70471-9
37. Colloby SJ, McParland S, O'Brien JT, et al. Neuropathological correlates of dopaminergic imaging in Alzheimer's disease and Lewy body dementias. *Brain : a journal of neurology* 2012;135(9):2798-808.
38. Kraemmer J, Kovacs GG, Perju-Dumbrava L, et al. Correlation of striatal dopamine transporter imaging with post mortem substantia nigra cell counts. *Movement Disorders* 2014;29(14):1767-73. doi: 10.1002/mds.25975

FIGURE LEGEND

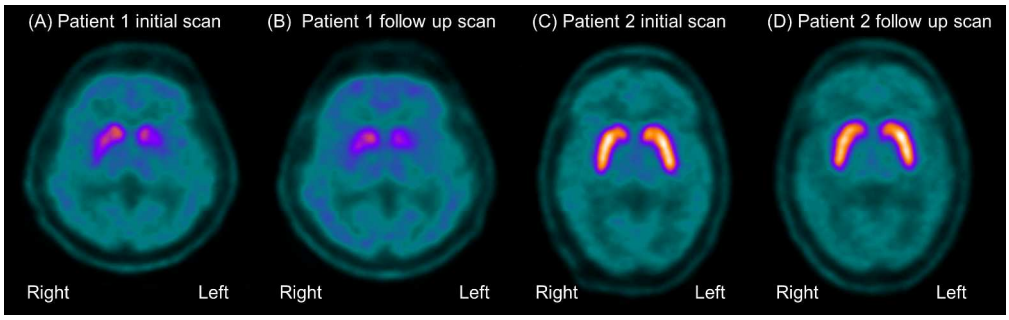
Figure 1. ROC curve of ^{18}F-AV-133 PET for predicting nigrostriatal degeneration in CUPS patients. The red line denotes the optimal binding reduction threshold of -50% as determined by the Youden criterion.

Figure 2. ^{18}F-AV-133 VMAT2 PET in two discordant cases. Patient with "rubral tremor" showing bilateral, asymmetrical reduction in tracer uptake at baseline (A) and follow-up two years later (B) with decline from -85% to –94% in the left posterior putamen. (C) Normal ^{18}F-AV-133 PET scan in a patient diagnosed with Parkinson's disease, unchanged in two year follow-up scan (D).



ROC curve of 18F-AV-133 PET for predicting nigrostriatal degeneration in CUPS patients. The red dot denotes the optimal binding reduction threshold of 50% as determined by the Youden criterion.

74x51mm (300 x 300 DPI)



18F-AV-133 VMAT2 PET in two discordant cases. Patient with "rubral tremor" showing bilateral, asymmetrical reduction in tracer uptake at baseline (A) and follow-up two years later (B) with decline from -85% to –94% in the left posterior putamen. (C) Normal 18F-AV-133 PET scan in a patient diagnosed with Parkinson's disease, unchanged in two year follow-up scan (D).

260x80mm (300 x 300 DPI)

Supplementary table 1. Individual list of cases with baseline putamen binding on the most affected side, and the pre-scan, post-scan and follow-up diagnoses.

Age	Putamen binding (%)*	Pre-scan diagnosis	Post-scan diagnosis	Follow-up diagnosis
52	-93%	FUNCTIONAL	NDG PD	NDG PD
71	-91%	NDG PD	NDG PD	NDG PD
69	-90%	NDG PD	NDG PD	NDG PD
45	-89%	NDG UNK	NDG PD	NDG PD
70	-88%	NDG UNK	NDG PD	NDG PD
53	-88%	NDG CBS	NDG PD	NDG PD
66	-87%	NDG PD	NDG PD	NDG PD
54	-87%	NDG PD	NDG PD	NDG PD
53	-86%	NDG PD	NDG PD	NDG PD
63	-85%	NDG PSP	NDG PD	NDG PD
74	-85%	MRT	NDG PD	NDG PD
68	-85%	NDG PD	NDG PD	NDG PD
49	-85%	NDG PD	NDG PD	NDG PD
79	-85%	RUBRAL TREMOR	NDG PD	RUBRAL TREMOR
51	-84%	NDG UNK	NDG PD	NDG PD
64	-84%	FUNCTIONAL	NDG PSP	NDG PSP
65	-83%	NDG PD	NDG PD	NDG MSA
45	-83%	NDG PD	NDG PD	NDG PD
69	-82%	NDG PD	NDG PD	NDG MSA
68	-80%	NDG PD	NDG PD	NDG PD
79	-80%	NDG PD	NDG PD	NDG PD
44	-79%	NDG PD	NDG PD	NDG PD
48	-79%	NDG PD	NDG PD	NDG PD
50	-79%	DRUG INDUCED	NDG PD	NDG PD
65	-79%	VASCULAR	NDG PD	NDG PD
57	-78%	MRT	NDG PD	NDG PD
30	-78%	NDG PD	NDG PD	NDG PD
63	-78%	DYSTONIA	NDG PD	NDG PD
59	-76%	NDG UNK	NDG PD	NDG PD
62	-76%	FUNCTIONAL	NDG PD	NDG PD
46	-72%	NDG PD	NDG PD	NDG PD
58	-72%	DYSTONIA	NDG PD	NDG PD
66	-71%	NDG PD	NDG PD	NDG PD
72	-71%	FUNCTIONAL	NDG PD	NDG PD
48	-69%	NDG PD	NDG PD	NDG PD
67	-69%	DYSTONIA	NDG PD or MSA	NDG MSA
46	-62%	NDG PD	NDG PD	NDG PD
51	-60%	NDG UNK	NDG PD	NDG PD
38	-55%	DYSTONIA	NDG PD	NDG PD

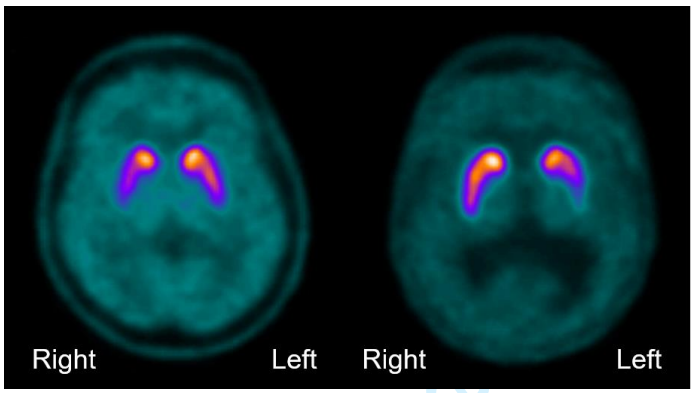
70	-54%	NDG UNK	NDG PD	NDG PD
68	-39%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
66	-39%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
53	-39%	MRT	NDG PD	NDG UNK
73**	-35%	NDG UNK	NDG PD	NDG PD
57	-27%	NDG PD	DYSTONIA	DYSTONIA
61	-25%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
63	-24%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
67	-24%	DYSTONIA	DYSTONIA	DYSTONIA
57	-23%	NDG PD	ET	ET
69	-23%	ET	ET	ET
55	-22%	ET	ET	ET
69	-20%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
51	-20%	NDG UNK	OTHER	FUNCTIONAL
61	-20%	NDG PD	NDG PD (SWEDD)	NDG PD (SWEDD)
69	-19%	DYSTONIA	DYSTONIA	DYSTONIA
47	-19%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
65**	-18%	NDG MSA	NDG MSA	NDG MSA
51	-18%	NDG PD	NDG PD (SWEDD)	NDG PD (SWEDD)
71	-16%	NDG PD	ET	ET
20	-16%	NDG (UNK)	UNK (NEUROGENETIC)	UNK (NEUROGENETIC)
68	-16%	DYSTONIA	LOST TO FOLLOW-UP	ET
61	-16%	NDG AD	NDG AD	NDG AD
68	-13%	ET	ET	ET
63	-11%	NDG PD	ET	FUNCTIONAL
43	-11%	NDG UNK	MRT	MRT
25	-10%	DYSTONIA	DYSTONIA	DYSTONIA
69	-8%	NDG PD	VASCULAR	VASCULAR
72	-6%	DRUG INDUCED	DRUG INDUCED	DRUG INDUCED
38	-5%	NDG PD	FUNCTIONAL	FUNCTIONAL
60	-4%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
54	-3%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
37	-2%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
30	-2%	NDG PD	FUNCTIONAL	FUNCTIONAL
53	2%	DYSTONIA	DYSTONIA	NDG PD
67	5%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
63	5%	NDG PD	FUNCTIONAL	FUNCTIONAL
28	7%	NDG PD	UNKNOWN	UNKNOWN
25	19%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
51	41%	FUNCTIONAL	FUNCTIONAL	FUNCTIONAL
51	67%	NDG PD	DRUG INDUCED	DRUG INDUCED
43	74%	DYSTONIA	FUNCTIONAL	FUNCTIONAL

Green background denotes cases where clinical diagnosis was concordant with scan result.

Red background denotes cases where clinical diagnosis was discordant with scan result.

* Putamen binding on the most affected side.

** Visual assessment of scan demonstrated significant and asymmetrical reduction in posterior putamen binding compared to caudate binding



NDG = Neurodegenerative

PD = Parkinson's disease

MSA = multiple system atrophy

CBS = Corticobasal syndrome

AD = Alzheimer's disease

MRT = Monosymptomatic resting tremor

UNK = Unknown

ET = Essential tremor

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	2
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	3
	4	Study objectives and hypotheses	4
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	4
Participants	**6**	Eligibility criteria	4, 5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	4, 5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	4
	9	Whether participants formed a consecutive, random or convenience series	4
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6, 7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6, 7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	6, 7
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	6, 7
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	7
	15	How indeterminate index test or reference standard results were handled	7, 8
	16	How missing data on the index test and reference standard were handled	7, 8
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	7, 8
	18	Intended sample size and how it was determined	7
RESULTS			
Participants	**19**	Flow of participants, using a diagram	N/A
	20	Baseline demographic and clinical characteristics of participants	9
	21a	Distribution of severity of disease in those with the target condition	9
	21b	Distribution of alternative diagnoses in those without the target condition	9
	22	Time interval and any clinical interventions between index test and reference standard	9
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	10
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	10
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	14
	27	Implications for practice, including the intended use and clinical role of the index test	13
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	7
	30	Sources of funding and other support; role of funders	16

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.