Original research

Stratifying quality of life outcome in subthalamic stimulation for Parkinson's disease

Alexandra Gronostay,[1] Stefanie Theresa Jost ,[1] Monty Silverdale,[2] Alexandra Rizos,[3] Philipp Alexander Loehrer ,[4] Julian Evans,[2] Anna Sauerbier,[1,3] Donya Indi,[1] Valentina Leta ,[3] Paul Reker,[1] Gereon Rudolf Fink,[1,5] Keyoumars Ashkan,[3] Angelo Antonini ,[6] Christopher Nimsky,[7] Veerle Visser-Vandewalle,[8] Pablo Martinez-Martin,[9] K Ray Chaudhuri,[3,10] Lars Timmermann,[4] Haidar S Dafsari [1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jnnp-2023-332272).

For numbered affiliations see end of article.

Correspondence to
Dr Haidar S Dafsari, Department of Neurology, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany; haidar.dafsari@uk-koeln.de

Received 24 July 2023
Accepted 3 December 2023
Published Online First 19 December 2023

To cite: Gronostay A, Jost ST, Silverdale M, *et al*. *J Neurol Neurosurg Psychiatry* 2024;**95**:630–638.

ABSTRACT

Background Subthalamic nucleus deep brain stimulation (STN-DBS) for Parkinson's disease (PD) improves quality of life (QoL), motor and non-motor symptoms (NMS). However, in previous studies, 43%–49% of patients did not experience clinically relevant postoperative QoL improvement. To inform individualised prediction of postoperative QoL improvement, we developed a stratification analysis of QoL outcomes based on preoperative non-motor total burden, severity of motor progression and motor response in levodopa challenge tests.

Methods This was a prospective, open-label, multicentre, international study with a 6-month follow-up. A distribution-based threshold identified 'QoL responders' in the PDQuestionnaire-8 Summary Index (PDQ-8 SI). After baseline stratification based on the NMS Scale, Hoehn and Yahr Scale and levodopa response assessed with the Unified PD Rating Scale-III, we compared postoperative QoL response between these strata. To assess the clinical usefulness and statistical feasibility of stratifications, we compared cumulative distribution function curves, respectively PDQ-8 within-stratum variation.

Results All main outcomes improved postoperatively. Based on the 8.1 points threshold for clinically meaningful PDQ-8 SI improvement, only 80/161 patients were classified as 'QoL responders'. The absolute risk reductions for QoL non-response among respective non-motor, motor progression and levodopa response strata were 23%, 8% and 3%, respectively. Only non-motor stratification reduced PDQ-8 within-stratum variation compared with the overall cohort.

Conclusions Non-motor stratification, but not motor progression or levodopa response stratification, is clinically useful and statistically feasible for personalised preoperative prediction of postoperative QoL outcome of STN-DBS for PD. Our findings highlight that non-motor assessments are necessary components of a case-based, holistic approach of DBS indication evaluations geared towards optimising postoperative QoL outcomes.

Trial registration number GermanClinicalTrialsRegister: #6735.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ In previous studies, 43%–49% of patients undergoing subthalamic nucleus deep brain stimulation for Parkinson's disease (PD) did not experience clinically relevant postoperative improvement of quality of life (QoL). To inform individualised prediction of postoperative QoL improvement, we developed a stratification analysis of QoL outcomes based on preoperative non-motor total burden, severity of motor progression and motor response in levodopa challenge tests.

WHAT THIS STUDY ADDS

⇒ Non-motor stratification, but not motor progression or levodopa response stratification, is clinically useful and statistically feasible for personalised preoperative prediction of postoperative QoL outcome of subthalamic stimulation for PD.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our findings highlight that non-motor assessments are necessary components of a case-based, holistic approach of deep brain stimulation indication evaluations geared specifically towards optimising postoperative QoL outcomes.

INTRODUCTION

Subthalamic nucleus deep brain stimulation (STN-DBS) is a safe and effective treatment of advanced Parkinson's disease (PD), improving quality of life (QoL),[1] motor and non-motor symptoms (NMS).[2] However, considerable interindividual variability has been observed for postoperative QoL outcome, as approximately 43%–49% of patients do not experience clinically relevant QoL improvement after STN-DBS surgery.[3–5] Several preoperative clinical predictors of greater postoperative QoL improvement have been reported: more severe non-motor total burden and motor impairment,[6] and greater response of motor symptoms in the levodopa challenge test.[7] However, these studies report general relationships between postoperative QoL outcome and predictor variables on a group level, but do not stratify QoL outcomes based on specific cut-offs of preoperative clinical parameters,

which could be used for individual patients. Therefore, the findings of these studies are of limited use in clinical practice.

To investigate the probability of postoperative QoL improvement of individual subjects, we conducted a stratification analysis based on preoperative NMS total burden, severity of motor progression and motor response in the levodopa challenge test. Conceptually, a stratification scheme is considered *clinically useful* if it provides a high discriminatory ability of clinically meaningful postoperative QoL change and *statistically feasible* if it results in small within-stratum variation of QoL.[8]

As previous studies reported that non-motor aspects of PD predict postoperative QoL outcome better than motor aspects,[5 9] we tested the hypothesis that preoperative non-motor total burden allows better stratification of postoperative QoL outcome than severity of motor progression and motor response in the levodopa challenge test.

MATERIALS AND METHODS

Design and ethical approval

This prospective, observational, international multicentre study in patients with advanced PD was conducted in four DBS centres (Cologne, Marburg, London and Manchester; GermanClinicalTrialsRegister: #6735).[10–12] The study was carried out following the Declaration of Helsinki. All patients gave written and informed consent before participation.

Patients

PD was diagnosed according to the Movement Disorder Society (MDS) criteria.[13] DBS treatment was performed in accordance to the MDS guidelines.[14] DBS was not performed if clinically relevant cognitive and mood disorders were observed in multidisciplinary assessments including neuropsychological and neuropsychiatric specialist assessments.[15 16]

Clinical assessment

Clinical assessments were conducted in the on-medication state before and in the on-medication and on-stimulation state at 6-month follow-up after DBS surgery:

Main outcome parameter:

- The PDQuestionnaire (PDQ) is recommended by the Movement Disorders Society Scales Committee for QoL assessments in patients with PD.[17 18] We used the 8-item version (PDQ-8), which contains the domains mobility, activities of daily living (ADL), emotional well-being, social support, cognition, communication, bodily discomfort and stigma.[17] The abbreviated PDQ-8 correlates highly with the PDQ-39 and is commonly used in DBS studies.[19–22] The results are reported as Summary Index (PDQ-8 SI) ranging from 0 (no impairment) to 100 (maximum impairment).

Outcomes used for stratification:

1. The Non-Motor Symptom Scale (NMSS) surveyed NMS in nine domains: cardiovascular, sleep/fatigue, mood/apathy, perceptual problems/hallucinations, attention/memory, gastrointestinal, urinary, sexual function and miscellaneous. The NMSS is commonly used in DBS studies, was assessed in the on-state and ranges from 0 (no impairment) to 360 points (maximum impairment).[23–26] Based on published cut-offs, we stratified non-motor total burden as 'mild to moderate' (≤40), 'severe' (41–70) or 'very severe' (≥71).[23 27]
2. The Hoehn and Yahr (H&Y) Scale assessed the severity of motor symptom progression ranging from 1 to 5[28]: 1 (unilateral involvement only usually with minimal or no functional disability), 2 (bilateral or midline involvement without impairment of balance), 3 (bilateral disease: mild to moderate disability with impaired postural reflexes; physically independent), 4 (severely disabling disease; still able to walk or stand unassisted) and 5 (confinement to bed or wheelchair unless aided). Based on published cut-offs for H&Y, we stratified the severity of motor progression as 'minimal to mild' (≤2), 'moderate' (3) or 'severe' (≥4).[29]
3. Motor response in the preoperative levodopa challenge test was quantified using the Unified PD Rating Scale-III (UPDRS-III), which ranges from 0 (no motor impairment) to 108 (maximum motor impairment). Off-medication state was assessed after long-acting dopamine agonists were halted for >72 and levodopa for >12 hours. Best on-medication state was assessed 45–60 min after administration of 250 mg levodopa.[30 31] Based on published cut-offs, we stratified levodopa responses as 'low' (<33%), 'partial' (33%–50%), or 'full' (>50%).[32]

The resulting groups were termed (1) non-motor, (2) motor and (3) levodopa response strata, respectively.

Secondary outcome parameters:

- ADL and motor complications were assessed with the Scales for Outcomes in PD-motor scale (SCOPA-M), which is an abbreviated form of the UPDRS and was used because of its time-efficiency.[33] The two scales highly correlate in respective sections.[34] The SCOPA-M ADL section ranges from 0 (no impairment) to 21 (maximum impairment) and the motor complications section ranges from 0 (no impairment) to 12 (maximum impairment).
- Antiparkinsonian medication requirements were quantified as levodopa equivalent daily dose (LEDD) according to the Jost method.[35]

Statistical analysis

Shapiro-Wilk tests assessed the normality of distribution. We tested outcome changes from baseline to 6-month follow-up using Wilcoxon signed-rank tests and the Benjamini-Hochberg correction for multiple comparisons due to the use of multiple tests. We present adjusted p values at the 0.05 significance threshold throughout the manuscript. We calculate relative change from baseline to follow-up ((mean $test_{baseline}$–mean $test_{follow\text{-}up}$)/mean $test_{baseline}$×100) and quantified effect size with Cohen's d.

We analysed the proportion of patients with a clinically meaningful QoL improvement using a distribution-based threshold (>1/2 SD of PDQ-8 $SI_{baseline}$).[36–38] Previous studies have employed the same method to derive a distribution-based cut-off value for QoL improvement.[5 39–41] Subsequently, we analysed postoperative PDQ-8 SI outcome changes of aforementioned non-motor (NMSS), motor (H&Y) and levodopa response strata.

We tested between-strata differences of baseline parameters and change scores using Kruskal-Wallis tests. We calculated change scores for clinical outcome parameters (change score=$test_{baseline}$–$test_{6\text{-}month\ follow\text{-}up}$) and analysed longitudinal within-stratum outcome changes with Wilcoxon signed-rank tests. Benjamini-Hochberg correction accounted for multiple comparisons (multiple scales and three stratifications).

Furthermore, to assess the *statistical feasibility* of our stratification analysis, that is, the ability to identify homogenous groups of patients, we estimated PDQ-8 SI within-stratum variation based on its baseline SD. Moreover, to assess the *clinical utility* of our stratification analysis, we assessed discriminatory ability of clinically meaningful QoL outcomes by comparing

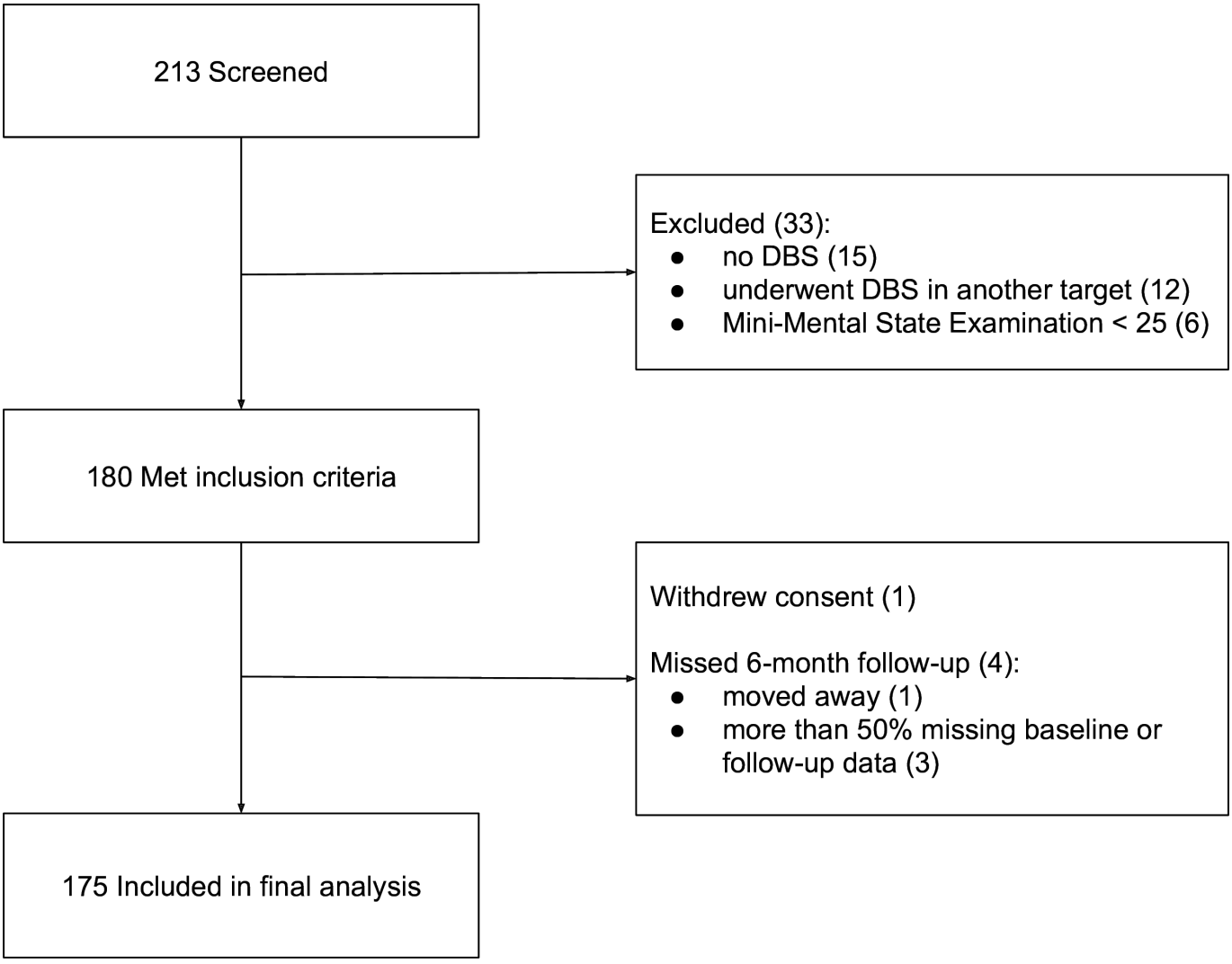

Figure 1 The flow chart describes the enrolment of patients. DBS, deep brain stimulation.

cumulative distribution function curves, which provided information on the complete distribution of QoL changes in each stratum.

All statistical analyses were conducted with SPSS V.29.0.0.0.

RESULTS

We included 175 patients (64 women) undergoing STN-DBS (see figure 1). In our cohort, the distribution-based threshold for clinically meaningful PDQ-8 SI improvement was 8.1 points. In four patients, baseline PDQ-8 SI was <8.1 points and thus by definition could not clinically relevantly improve at follow-up. Therefore, these four patients were excluded from further analyses.

The remaining 171 patients were mean aged 61.2 years±10.7 with 10.5 years mean disease duration±4.9 (see Table 1). All main outcomes improved from baseline to follow-up (all p≤0.001). Effect sizes were 'small' for H&Y, 'moderate' for PDQ-8 SI, NMSS, UPDRS-III and SCOPA-M ADL and 'large' for motor complications and LEDD.

PDQ-8 SI change scores were available for 161 patients. We observed a considerable interindividual variability of PDQ-8 SI outcomes as 80/161 patients experienced a clinically relevant postoperative PDQ-8 SI improvement (49.4% 'QoL responders').

In the subsequent stratification analysis based on preoperative non-motor total burden, severity of motor progression, and levodopa response, the non-motor stratification resulted in the strata 'mild to moderate' in 56, 'severe' in 50 and 'very severe' in 60 patients. Motor progression stratification resulted in the strata 'minimal to mild' in 96, 'moderate' in 48 and 'severe' in 19 patients. The levodopa response stratification resulted in the response strata 'low' in 43, 'partial' in 55 and 'full' in 66 patients. The percentage of missing values was <5% for each scale and, therefore, required no imputations.[42] The 'low' levodopa response stratum contained patients, in which the clinical indication for DBS treatment was case based on the following indications: medication-refractory tremor (38 %), troublesome dyskinesia (10 %) and on-off motor fluctuations (42 %) and others (10 %).

QoL at baseline in non-motor, motor and levodopa response strata

At baseline, PDQ-8 SI increased with more severe non-motor total burden, from 'mild to moderate' through 'severe' to 'very severe' (p<0.001, see table 2). In contrast, baseline PDQ-8 SI did not differ significantly between the motor or the levodopa challenge test response strata (p=0.105, respectively p=0.947).

Online supplemental table 1 presents preoperative NMSS total, H&Y, levodopa challenge test response, UPDRS-III, SCOPA-M

Table 1 Demographic characteristics and clinical outcomes at baseline and change scores at 6-month follow-up

	Baseline			6-month follow-up			Relative change	Effect size	P value
	n	Mean	SD	n	Mean	SD	(%)		
Age	167	61.2	10.7						
Disease duration	166	10.5	4.9						
Sex (female/male) (%)	167	(105/62)	(62.9/37.1%)						
Levodopa challenge test response (%)	164	0.5	0.2						
PDQ-8 Summary Index	167	33.8	15.1	161	25.7	16	−24.0	0.52	**<0.001**
NMSS total	166	62.4	36.9	162	42.4	27.7	−32.1	0.61	**<0.001**
Hoehn and Yahr	163	2.4	0.9	159	2.2	0.7	−8.3	0.25	**0.001**
UPDRS-III	167	26	12.9	158	20	10.1	−23.1	0.52	**<0.001**
SCOPA-motor ADL	160	7.3	3.2	155	5.4	3.1	−26.0	0.60	**<0.001**
SCOPA-motor complications	160	5.1	2.9	154	2.7	2.6	−47.1	0.87	**<0.001**
LEDD	166	1107.1	515.8	162	566.0	346.7	−48.9	1.23	**<0.001**

We analysed outcome changes from baseline to 6-month follow-up with Wilcoxon signed-rank tests and used Benjamini-Hochberg correction for multiple comparisons due to the use of multiple tests (seven clinical outcomes). We calculate relative change from baseline to follow-up ((mean $test_{baseline}$−mean $test_{follow\text{-}up}$)/mean $test_{baseline}$×100) and quantified effect size with Cohen's d.
Bold font highlights significant results.
ADL, activities of daily living; LEDD, levodopa equivalent daily dose; NMSS, Non-motor Symptom Scale; PDQ-8, Parkinson's Disease Questionnaire-8; SCOPA, Scales for Outcomes in Parkinson's Disease; UPDRS-III, Unified Parkinson's Disease Rating Scale-Motor Examination.

ADL and motor complications, and LEDD for each stratum in the non-motor, motor and levodopa response stratification. As expected, non-motor, motor and levodopa response stratification resulted in a significant difference in baseline NMSS total, H&Y and levodopa challenge test response in respective strata (all p<0.001). Notably, a difference in baseline H&Y was also observed in the non-motor stratification (p=0.041), a difference in baseline NMSS total in the motor progression stratification (p=0.020) and baseline motor complications differed in all three stratifications (all p<0.005).

In the non-motor strata, we observed between-strata differences for all baseline PDQ-8 domains (communication and bodily discomfort domains: p<0.001, emotional well-being, bodily discomfort and stigma: p<0.05) except for the social support domain (p=0.095), whereas motor progression stratification resulted in significant between-strata differences only in the mobility (p=0.031) and ADL (p=0.014) domains, and levodopa response stratification only in the bodily discomfort domain (p=0.026).

Change scores of clinical outcome parameters at 6-month follow-up in non-motor, motor and levodopa response strata

At 6-month follow-up, non-motor stratification resulted in between-strata differences of PDQ-8 SI change scores (p=0.003, see table 3), whereas we observed no significant between-strata differences in the motor or levodopa response stratification (both p>0.05). Post hoc comparisons showed that postoperative PDQ-8 SI changes of the 'very severe' non-motor stratum differed from the 'mild-to-moderate' (p=0.004) and the 'severe' (p=0.028) non-motor strata.

Furthermore, explorative analyses showed that these between-strata differences were driven by the PDQ-8 domains cognition (p=0.033) and communication (p=0.007; see online

Table 2 Baseline clinical parameters in non-motor, motor and levodopa response strata

	Non-Motor Symptom Scale									
	'Mild to moderate'			'Severe'			'Very severe'			
	(≤40)			(41–70)			(≥71)			
	n	Mean	SD	n	Mean	SD	n	Mean	SD	P value
PDQ-8 Summary Index	56	26.2	12.3	50	33.7	13.3	60	41.4	15.5	**<0.001*** **
	Hoehn and Yahr Scale									
	'Minimal to mild'			'Moderate'			'Severe'			
	(≤ 2.5)			(3)			(≥ 4)			
	n	Mean	SD	n	Mean	SD	n	Mean	SD	P value
PDQ-8 Summary Index	96	31.3	14.5	48	37.2	14.7	19	37.3	16.7	0.105
	Levodopa challenge test response									
	'Low'			'Partial'			'Full'			
	(<33%)			(33%–50%)			(>50%)			
	n	Mean	SD	n	Mean	SD	n	Mean	SD	P value
PDQ-8 Summary Index	43	34.2	14.1	55	34.1	15.5	66	33.7	15.8	0.947

Kruskal-Wallis tests were performed for independent samples for between-strata analyses of baseline parameters. Multiple comparisons (three strata) were corrected with the Bonferroni method.
Bold font highlights significant results.
***p<0.001.
PDQ-8, 8-item Parkinson's Disease Questionnaire.

Table 3 Changes of clinical parameters at 6-month follow-up in non-motor, motor and levodopa response stratification

	Non-Motor Symptom Scale												
	'Mild to moderate'				'Severe'				'Very severe'				
	(≤40)				(41–70)				(≥71)				
	n	Mean	SD	P value†	n	Mean	SD	P value†	n	Mean	SD	P value†	P value‡
PDQ-8 Summary Index Change Score	52	7.9	12.5	**<0.001*** **	49	2.7	16.7	0.173	60	13.0	16.3	**<0.001*** **	**0.003****
	Hoehn and Yahr Scale												
	'Minimal to mild'				'Moderate'				'Severe'				
	(≤2.5)				(3)				(≥4)				
	n	Mean	SD	P value†	n	Mean	SD	P value†	n	Mean	SD	P value†	P value‡
PDQ-8 Summary Index Change Score	93	7.9	13.3	**<0.001*** **	47	7.6	19.9	**0.010***	18	10.9	16.3	**0.006****	0.961
	Levodopa challenge test response												
	'Low'				'Partial'				'Full'				
	(<33%)				(33%–50%)				(>50%)				
	n	Mean	SD	P value†	n	Mean	SD	P value†	n	Mean	SD	P value†	P value‡
PDQ-8 Summary Index Change Score	42	8.2	16.0	**0.007****	51	8.0	16.7	**0.001****	66	8.9	15.2	**<0.001*** **	0.993

Bold font highlights significant results.
*p < 0.05, ** p < 0.01, *** p < 0.001.
†Wilcoxon signed-rank tests were used for within-group analyses of change scores between baseline and 6-month follow-up. Multiple comparisons (three strata and seven scales) were corrected with the Benjamini-Hochberg method.
‡Kruskal-Wallis tests were performed for independent samples for between-group analyses of change scores between baseline and 6-month follow-up.
PDQ-8, 8-item Parkinson's Disease Questionnaire.

supplemental figure 1). In contrast, no significant differences of PDQ-8 domain changes were observed between the motor and levodopa response strata. We also explored within-group PDQ-8 domain changes in the non-motor strata and observed that in the 'very severe' stratum all PDQ-8 domains improved (all p≤0.013) except for social support (p>0.05). In the 'very severe' non-motor stratum, effect sizes were 'moderate' in the ADL, cognition and bodily discomfort domains and 'small' in all other domains (see online supplemental table 2). In contrast, only two domains, mobility (p=0.028) and ADL (p=0.014), improved in the 'severe' non-motor stratum and only one domain, cognition (p=0.004), in the 'mild-to-moderate' non-motor stratum, and the effect sizes were generally smaller than in the 'very severe' non-motor stratum.

Online supplemental table 3 presents postoperative change scores of NMSS total, H&Y, UPDRS-III, SCOPA-M ADL and motor complications, and LEDD for each stratum in the non-motor, motor and levodopa response stratification. Non-motor and motor progression stratification resulted in significant between-strata differences of postoperative change scores of NMSS total, respectively H&Y Scale (both p<0.001). Notably, only non-motor stratification resulted in a significant between-strata difference of motor complication change scores (p=0.032). The postoperative improvement of motor complications increased from the 'mild to moderate' through the 'severe' to the 'very severe' non-motor stratum.

Within-stratum SDs of PDQ-8 SI at baseline and discriminatory ability of PDQ-8 SI change

In the overall cohort, the SD of PDQ-8 SI was 16.1 points. Table 2 shows that non-motor stratification resulted in a reduction of within-stratum SD of baseline PDQ-8 SI (range: from 12.3 in the 'mild-to-moderate' stratum to 15.5 points in the 'severe' non-motor stratum). In contrast, levodopa response stratification reduced SD to a smaller degree (range: 14.1–15.8) and motor progression stratification did not reduce SD in all strata (range: 14.5–16.7). Therefore, non-motor stratification resulted in a greater reduction of SD of baseline PDQ-8 SI than motor and levodopa response stratifications. Compared with the overall cohort (16.1 points), the SD was reduced in all non-motor strata with a maximum reduction of 3.8 points (approximately 24%) in the 'mild-to-moderate stratum'.

Furthermore, as illustrated in the cumulative distribution curves (see figure 2), in the non-motor stratification, approximately 62% of patients in the 'very severe' stratum experienced a clinically relevant PDQ-8 SI improvement compared with approximately 46% and 39% of patients in the 'mild-to-moderate' and 'severe' strata, respectively. In the motor progression stratification, approximately 56% of patients in the 'severe' stratum experienced a clinically relevant PDQ-8 SI improvement compared with approximately 48% in the 'minimal-to-mild' strata and approximately 49% in the 'moderate' strata. In the levodopa response stratification, the proportion of patients who experienced a clinically relevant PDQ-8 SI improvement ranged from approximately 49% in the 'partial' to approximately 50% in the 'full' and approximately 52% in the 'low' response strata. Therefore, the discriminatory ability for clinically relevant PDQ-8 SI improvement was greater in the non-motor stratification (approximately 23% difference between non-motor strata) than for motor progression and levodopa response stratification (approximately 8% and 3% difference between motor progression, respectively levodopa response strata).

DISCUSSION

In this prospective, observational, international multicentre study, we observed that preoperative non-motor total burden is a suitable stratification parameter for postoperative QoL outcome. Specifically, non-motor stratification was *clinically useful* because of its ability to discriminate QoL responders and QoL non-responders as the absolute risk of QoL non-response was reduced by 24% from the 'very severe' to 'severe' non-motor stratum. In contrast, the clinical usefulness of motor progression and levodopa response stratifications of QoL outcome was considerably smaller (8%, respectively 3% between-strata absolute risk reduction for QoL non-response). Moreover, in contrast to motor progression and levodopa response stratification, non-motor stratification reduced within-stratum variation of PDQ-8 SI compared with the overall cohort, which demonstrates its

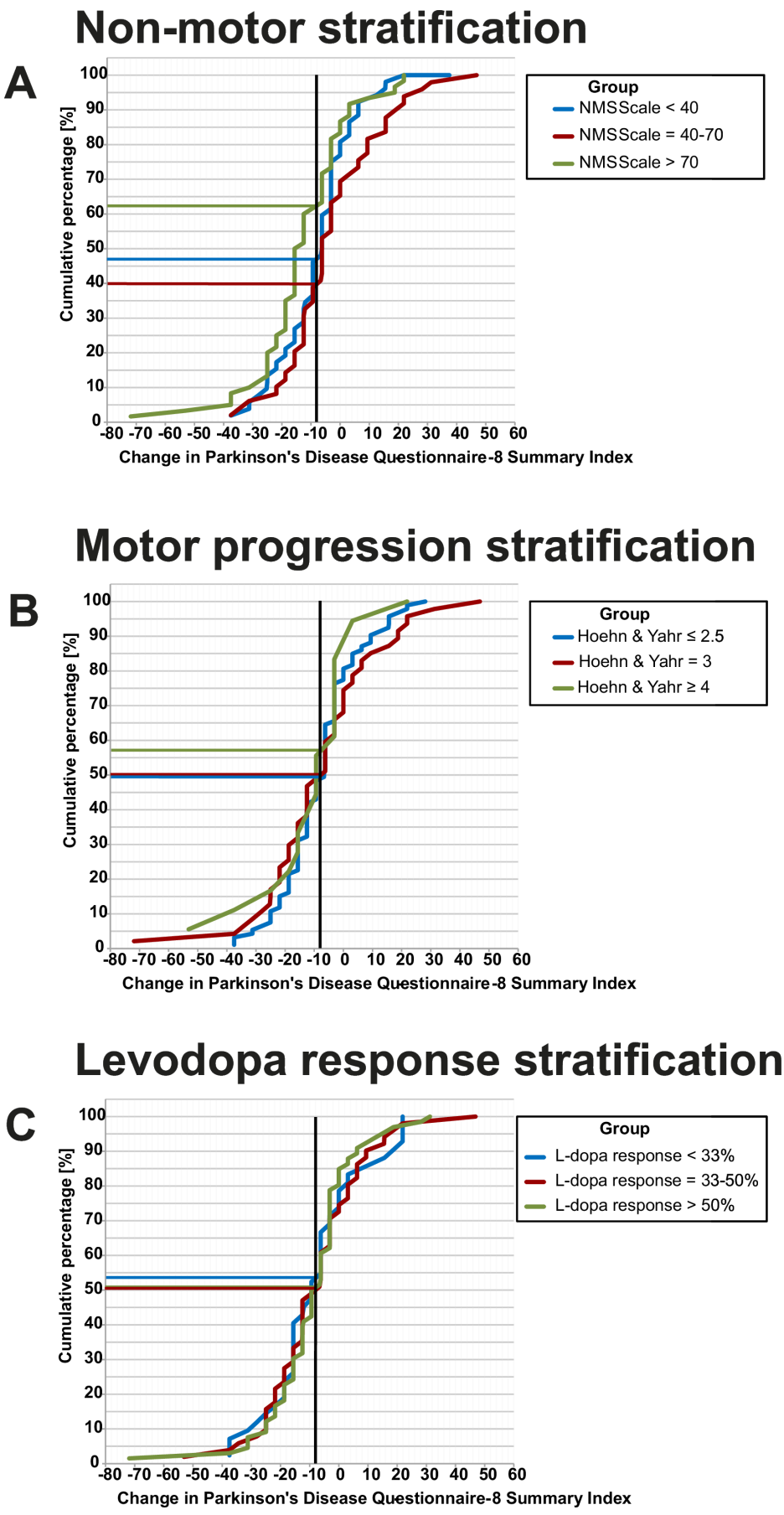

Figure 2 Cumulative distribution curves for patients stratified by Non-Motor Symptom (NMS) Scale total burden, Hoehn and Yahr motor stage and levodopa challenge test response.

statistic feasibility for the identification of homogenous groups of patients.

Clinical efficacy of subthalamic stimulation in the overall cohort

In line with the previous literature, STN-DBS improved QoL, non-motor total burden and motor aspects of PD, and reduced dopaminergic medication requirements.[1] The 46.2% mean preoperative levodopa challenge test response was well within the range of previous DBS studies.[1 43] We confirm a considerable interindividual variability of postoperative QoL outcome as approximately 50% of our cohort experienced a clinically relevant postoperative QoL improvement, which is in line with previous studies.[3–5]

QoL impairments in non-motor, motor progression and levodopa response strata

We observed that non-motor stratification resulted in between-strata differences of baseline PDQ-8 SI as patients with more severe non-motor total burden expressed worse preoperative QoL. Data from the Earlystim trial and other cohorts indicates that patients with more severe baseline QoL impairment experience greater postoperative QoL improvement[4 39] and our study demonstrates that worse preoperative non-motor total burden, not severity of motor progression or motor response in the levodopa challenge test, is associated with greater preoperative QoL impairment.

A meta-analysis of 22 studies showed that greater preoperative levodopa challenge test response is associated with better postoperative outcome of UPDRS subscales for ADL and motor examination.[44] The relationship between the preoperative levodopa challenge test response and postoperative QoL outcome is less clear. We confirm the findings of an analysis of two randomised, controlled studies, which reported similar QoL improvements between patients with low, partial and full preoperative levodopa challenge test response.[32]

Indication evaluations for deep brain stimulation: levodopa challenge test response matters, but only as a supportive criterion for the diagnosis of PD

A recent survey among 207 movement disorders experts from 59 countries reports that a minimum levodopa challenge test response of >50% or >33% is required in 15.2%, respectively 37.3% of respondents.[45] This finding is surprising as there is no hard evidence that patients with a clinically established diagnosis of PD and <33% motor response in the levodopa challenge test per se experience negative outcomes following DBS.[46] In light of the emerging evidence from several cohorts that postoperative QoL improvements are similar across levodopa response strata,[32] the application of an absolute cut-off of levodopa challenge test response likely denies access to DBS treatment for patients with PD and low levodopa response (26.2% of our cohort). The partial levodopa response threshold implemented in some centres would have denied DBS treatment in 59.8% of our cohort.

The cut-off of >33% levodopa challenge test response supports the diagnosis of idiopathic PD as its positive predictive value for a clinical diagnosis of PD is 88.6%.[47] In this context, we advocate applying the Movement Disorders Society (MDS) Clinical Diagnostic Criteria for PD as a first step in a case-based approach in DBS indication evaluations.[13] The MDS criteria open up the possibility to clinically establish diagnosis of PD, even if patients show <30% motor response in the levodopa challenge test, which can occur, for example, due to gastrointestinal barriers to levodopa transport and absorption. Within the framework of these MDS criteria, the clinically established diagnosis of PD requires the absence of absolute exclusion criteria and red flags and, in patients with <30% levodopa response, the presence of at least two of the following criteria: (1) rest tremor in one limb, which may or may not be medication refractory, (2) troublesome dyskinesia, (3) unequivocal and marked on-off fluctuations, which must have at some point in time included predictable end-of-dose wearing-off and (4) positive results from at least one ancillary test for idiopathic PD with a specificity of 80% or higher. Once PD has been diagnosed according to the MDS criteria, we advocate an individualised case-based evaluation of the expected DBS outcome and benefit. This second step should include, but not be limited to, an assessment of QoL, non-motor and motor aspects of PD, a consideration of antiparkinsonian medication regimens, effects and side effects, an assessment of comorbidities and a careful reflection on the psychosocial and occupational situation of patients. Patients with PD who experience worse preoperative QoL impairments and worse non-motor total burden experience greater postoperative QoL improvements and, closely connected to this point, have a higher chance of experiencing a clinically relevant postoperative QoL improvement.

Toward a case-based approach in indication evaluations for deep brain stimulation for PD including non-motor assessments

There is a growing body of evidence which demonstrates the beneficial non-motor effects of STN-DBS: Data from randomised, controlled studies provides class I evidence for beneficial effects up of STN-DBS for up to 2 years on several non-motor aspects of PD, such as anxiety, depression, hyperdopaminergic behavioural disorders, pain/bodily discomfort or urinary symptoms.[48–51] Moreover, class IIb evidence shows that non-motor total burden improves from baseline up to 3-year follow-up, which also correlates with an improvement of QoL.[52] In the current study, we demonstrate that preoperative non-motor total burden is a clinically useful and statistically feasible stratification parameter for postoperative QoL response and that neither preoperative levodopa challenge test response nor preoperative severity of motor progression demonstrated these characteristics in our cohort.

Moreover, we observed between-strata differences of baseline motor complications in non-motor, motor progression and levodopa response stratifications. However, motor complication change scores were only different between non-motor strata as motor complication improvements were greater with worse preoperative non-motor total burden. Stratifications based on motor aspects of PD (severity of motor progression and levodopa response) were not suited to stratify postoperative motor complications outcome.

Continuing the longstanding tradition of clinical research on the refinement of selection criteria for DBS, we advocate that greater weight should be given to the non-motor aspects of PD. In the context of DBS indication evaluations, levodopa challenge test response should be viewed as a supportive criterion for the diagnosis of PD, not as an absolute exclusion criterion for DBS treatment when specific cut-offs of >33% or >50% are not achieved in individual patients. In our view, in DBS indication evaluations, a case-based, holistic approach may be considered best practice and should also include assessments of non-motor aspects of PD.

Limitations

In particular, in the 'severe' motor progression stratum (n=19), stratification analysis resulted in a small stratum size, whereas all

other non-motor, motor progression and levodopa response strata were larger (n>40) and the sample size of the overall cohort was larger than in most DBS studies.[1 51] Studies with longer follow-up periods are needed for the stratification of long-term postoperative QoL outcome. Further studies in other target regions, for example, the Globus Pallidus internus, are needed. Another limitation of our stratification analysis is the use of global non-motor, motor progression and QoL scales (NMSS, H&Y, UPDRS-III response in the in a levodopa challenge test, and PDQ-8). Specific motor and NMS or QoL domains may drive the global effects reported in the present study and further work is needed to identify these aspects in detail. However, the scales we used in the present study are recommended by the MDS clinical outcomes committee and are commonly used for the assessment of non-motor and motor aspects of PD,[17 28 53 54] and we were interested specifically in global disturbances of NMS and motor symptoms, and QoL.

Conclusion

We report that preoperative non-motor total burden, not severity of motor progression or levodopa challenge test response, is clinically useful and statistically feasible to stratify postoperative QoL outcome in patients undergoing STN-DBS for PD. Our findings highlight that, besides the traditional preoperative assessments of motor symptoms of PD including levodopa challenge test response, examinations of QoL disturbances and non-motor total burden are necessary components for a case-based, holistic approach in indication evaluations for DBS.

Author affiliations
[1]Department of Neurology, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[2]Department of Neurology and Neurosurgery, Salford Royal NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, Greater Manchester, UK
[3]Parkinson Foundation International Centre of Excellence, King's College Hospital, London, UK
[4]Department of Neurology, Philipps-University Marburg, Marburg, Germany
[5]Cognitive Neuroscience, Institute of Neuroscience and Medicine (INM-3), Research Center Jülich, Jülich, Germany
[6]Parkinson and Movement Disorders Unit, Department of Neurosciences (DNS), Padova University, Padova, Italy
[7]Department of Neurosurgery, Philipps-University Marburg, Marburg, Germany
[8]Department of Stereotactic and Functional Neurosurgery, University of Cologne, Faculty of Medicine and University Hospital Cologne, Cologne, Germany
[9]Center for Networked Biomedical Research in Neurodegenerative Diseases (CIBERNED), Carlos III Institute of Health, Madrid, Spain
[10]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

Contributors AG: Organisation and execution of the study, execution of the statistical analysis and writing of the first draft of the manuscript. STJ, AR, PAL, JE, AS, DI, VL, PR, GRF, KA, AA, VV-V and KRC: Review and critique of the statistical analysis and of the manuscript. PM-M: Conception and organisation of the study, review and critique of the statistical analysis and of the manuscript. KRC and LT: Conception, organisation and execution of the study, review and critique of the manuscript. HSD: Conception, organisation and execution of the study, design and execution of the statistical analysis and writing of the first draft and review and critique of the manuscript. Guarantor of the study.

Funding This paper is independent research funded by the German Research Foundation (Grant KFO 219).

Competing interests AG reports no financial disclosures. STJ was funded by the Professor KT Foundation and the Brandau-Laibach-Foundation, outside of the submitted work. MS has received honoraria from Bial, Britannia and Medtronic, outside the submitted work. AR has received salary support from National institute of Health Research (NIHR) Clinical Research Network (CRN) South London and the International Parkinson and Movement Disorder Society (MDS), outside the submitted work. PAL was supported by the SUCCESS-Program of the Philipps-University of Marburg, the 'Stiftung zur Förderung junger Neurowissenschaftler' and the Professor KT Foundation in the German Society of Neurology. He received an honorarium for the publication of the book 'Continuous Dopaminergic Stimulation for Parkinson's Disease: where are we'. JE reports no financial disclosures. AS is funded by the Gusyk program and the Advanced Cologne Clinician Scientist program of the Medical Faculty of the University of Cologne and has received funding from the Professor KT Foundation, outside the submitted work. DI reports no financial disclosures. VL received honoraria for speaker-related activities from Bial, Profile and Britannia Pharmaceuticals, outside the submitted work. PR reports no financial disclosures. GRF is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 431549029—SFB 1451. GRF serves as an editorial board member of Cortex, Neurological Research and Practice, NeuroImage: Clinical, Zeitschrift für Neuropsychologie and DGNeurologie; receives royalties from the publication of the books *Funktionelle MRT in Psychiatrie und Neurologie*, *Neurologische Differentialdiagnose* and *SOP Neurologie*; receives royalties from the publication of the neuropsychological tests KAS and Köpps; received honoraria for speaking engagements from Bayer, Desitin, DGN, Ergo DKV, Forum für medizinische Fortbildung FomF GmbH, GSK, Medica Academy Messe Düsseldorf, Medicbrain Healthcare, Novartis, Pfizer and Sportärztebund NRW, outside the submitted work. KA has received honoraria for educational meetings, travel and consultancy from Medtronic, St. Jude Medical and Boston Scientific, outside the submitted work. AA reports personal consultancy fees from UCB, Boehringer Ingelheim, Britannia, AbbVie, Kyowa Kirin, Zambon, Bial, Neuroderm, Theravance Biopharma, Roche; he receives research support from Chiesi Pharmaceuticals, Lundbeck, Horizon 2020—Grant 825785, Horizon2020 Grant 101016902, Ministry of Education University and Research (MIUR) Grant ARS01_01081, Cariparo Foundation. He serves as consultant for Boehringer Ingelheim for legal cases on pathological gambling, outside the submitted work. CN is consultant for Brainlab and received speaker honoraria from BK medical, Brainlab and BB, outside the submitted work. VV-V is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Project-ID 431549029—SFB 1451, is a member of the advisory boards and reports consultancies for Medtronic, Boston Scientific, Abbott, Insightec and LivaNova, outside the submitted work. PM-M has received honoraria from Bial for lecturing in a course and from the International Parkinson and Movement Disorder Society for activities during the International Congress 2022 and for directing the Clinical Outcome Assessment Program, outside the submitted work. KRC has received funding from Parkinson's UK, NIHR, UCB and the European Union; he received honoraria from UCB, Abbott, Britannia, US Worldmeds and Otsuka Pharmaceuticals; and acted as a consultant for AbbVie, UCB and Britannia, outside the submitted work. LT has received payments as a consultant for Boston Scientific; honoraria as a speaker on symposia sponsored by UCB, Desitin, Boston Scientific, AbbVie, Novartis, GlaxoSmithKline, Neuraxpharm, Teva, the Movement Disorders Society und DIAPLAN. The institution of LT, not LT personally, received funding by Boston Scientific, the German Research Foundation, the German Ministry of Education and Research, the Otto Loewi Foundation and the Deutsche Parkinson Vereinigung, outside the submitted work. Neither LT nor any member of his family holds stocks, stock options, patents or financial interests in any of the above-mentioned companies or their competitors. LT serves as the vice president of the German Neurological Society without any payment or any income. HSD was funded by the EU Joint Programme—Neurodegenerative Disease Research (JPND), the Professor Klaus Thiemann Foundation and the Felgenhauer Foundation and has received honoraria by Boston Scientific, Medtronic and Stadapharm, outside the submitted work.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by ethics committee of the University of Cologne, Germany: #0012-145London, UK:National Research Ethics Service (NRES)—South East London research ethics committee (REC), 10084, 10/H0808/141. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request.

ORCID iDs
Stefanie Theresa Jost http://orcid.org/0000-0003-0477-2289
Philipp Alexander Loehrer http://orcid.org/0000-0002-5279-2156
Valentina Leta http://orcid.org/0000-0002-5794-7540
Angelo Antonini http://orcid.org/0000-0003-1040-2807
Haidar S Dafsari http://orcid.org/0000-0001-8849-4233

REFERENCES

1 Deuschl G, Schade-Brittinger C, Krack P, *et al*. A randomized trial of deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2006;355:896–908.

2 Leta V, Dafsari HS, Sauerbier A, *et al*. Personalised advanced therapies in Parkinson's disease: the role of non-motor symptoms profile. *J Pers Med* 2021;11:773.
3 Daniels C, Krack P, Volkmann J, *et al*. Is improvement in the quality of life after Subthalamic nucleus stimulation in Parkinson's disease predictable? *Mov Disord* 2011;26:2516–21.
4 Floden D, Cooper SE, Griffith SD, *et al*. Predicting quality of life outcomes after Subthalamic nucleus deep brain stimulation. *Neurology* 2014;83:1627–33.
5 Jost ST, Visser-Vandewalle V, Rizos A, *et al*. Non-motor predictors of 36-month quality of life after Subthalamic stimulation in Parkinson disease. *NPJ Parkinsons Dis* 2021;7:48.
6 Geraedts VJ, Feleus S, Marinus J, *et al*. What predicts quality of life after Subthalamic deep brain stimulation in Parkinson's disease? A systematic review. *Euro J of Neurology* 2020;27:419–28. 10.1111/ene.14147 Available: https://onlinelibrary.wiley.com/toc/14681331/27/3
7 Zheng Z, Yin Z, Zhang B, *et al*. Levodopa challenge test predicts STN-DBS outcomes in various Parkinson's disease motor subtypes: a more accurate judgment. *Neural Plast* 2021;2021:4762027.
8 Yong FH, Tian L, Yu S, *et al*. Optimal stratification in outcome prediction using baseline information. *Biometrika* 2016;103:817–28.
9 Sauerbier A, Loehrer P, Jost ST, *et al*. Predictors of short-term impulsive and compulsive behaviour after Subthalamic stimulation in Parkinson disease. *J Neurol Neurosurg Psychiatry* 2021;92:1313–8.
10 Jost ST, Konitsioti A, Loehrer PA, *et al*. Non-motor effects of deep brain stimulation in Parkinson's disease motor subtypes. *Parkinsonism & Related Disorders* 2023;109:105318.
11 Dafsari HS, Dos Santos Ghilardi MG, Visser-Vandewalle V, *et al*. Beneficial Nonmotor effects of Subthalamic and Pallidal Neurostimulation in Parkinson's disease. *Brain Stimul* 2020;13:1697–705.
12 Jost ST, Ray Chaudhuri K, Ashkan K, *et al*. Subthalamic stimulation improves quality of sleep in Parkinson disease: a 36-month controlled study. *J Parkinsons Dis* 2021;11:323–35.
13 Postuma RB, Berg D, Stern M, *et al*. MDS clinical diagnostic criteria for Parkinson's disease. *Mov Disord* 2015;30:1591–601.
14 Deuschl G, Antonini A, Costa J, *et al*. European Academy of neurology/movement disorder society-European section guideline on the treatment of Parkinson's disease: I. *Mov Disord* 2022;37:1360–74.
15 Florin E, Dafsari HS, Reck C, *et al*. Modulation of local field potential power of the Subthalamic nucleus during Isometric force generation in patients with Parkinson's disease. *Neuroscience* 2013;240:106–16.
16 Reker P, Jost ST, Schiller P, *et al*. Gender differences in referrals for deep brain stimulation for essential tremor. *Parkinsonism Relat Disord* 2023;112:105490.
17 Martinez-Martin P, Jeukens-Visser M, Lyons KE, *et al*. Health-related quality-of-life scales in Parkinson's disease: critique and recommendations. *Mov Disord* 2011;26:2371–80.
18 Jenkinson C, Fitzpatrick R. Cross-cultural evaluation of the short form 8-item Parkinson's disease questionnaire (PDQ-8): results from America. *Parkinsonism Relat Disord* 2007;13:22–8.
19 Storch A, Schneider CB, Wolz M, *et al*. Nonmotor fluctuations in Parkinson disease: severity and correlation with motor complications. *Neurology* 2013;80:800–9.
20 Dafsari HS, Ray-Chaudhuri K, Ashkan K, *et al*. Beneficial effect of 24-month bilateral Subthalamic stimulation on quality of sleep in Parkinson's disease. *J Neurol* 2020;267:1830–41.
21 Dafsari HS, Ray-Chaudhuri K, Mahlstedt P, *et al*. Beneficial effects of bilateral Subthalamic stimulation on Alexithymia in Parkinson's disease. *Eur J Neurol* 2019;26:222–e17.
22 Dafsari HS, Reker P, Silverdale M, *et al*. Subthalamic stimulation improves quality of life of patients aged 61 years or older with short duration of Parkinson's disease. *Neuromodulation* 2018;21:532–40.
23 Chaudhuri KR, Martinez-Martin P, Brown RG, *et al*. The metric properties of a novel non-motor symptoms scale for Parkinson's disease: results from an international pilot study. *Mov Disord* 2007;22:1901–11.
24 Sauerbier A, Bachon P, Ambrosio L, *et al*. The new satisfaction with life and treatment scale (SLTS-7) in patients with Parkinson's disease. *J Parkinsons Dis* 2022;12:453–64.
25 Jost ST, Strobel L, Rizos A, *et al*. Gender gap in deep brain stimulation for Parkinson's disease. *NPJ Parkinsons Dis* 2022;8:47.
26 Lazcano-Ocampo C, van Wamelen D, Samuel M, *et al*. Evaluation of the effect of bilateral Subthalamic nucleus deep brain stimulation on fatigue in Parkinson's disease as measured by the non-motor symptoms scale. *Br J Neurosurg* 2021:1–4.
27 Sauerbier A, Schrag A, Brown R, *et al*. Clinical non-motor Phenotyping of black and Asian minority ethnic compared to white individuals with Parkinson's disease living in the United Kingdom. *J Parkinsons Dis* 2021;11:299–307.
28 Goetz CG, Poewe W, Rascol O, *et al*. Movement disorder society task force report on the Hoehn and Yahr staging scale: status and recommendations. *Mov Disord* 2004;19:1020–8.
29 Young A, Home M, Churchward T, *et al*. Comparison of sleep disturbance in mild versus severe Parkinson's disease. *Sleep* 2002;25:573–7.
30 Perez Parra S, McKay JL, Factor SA. Diphasic worsening of freezing of gait in Parkinson's disease. *Mov Disord Clin Pract* 2020;7:325–8.
31 Nettersheim FS, Loehrer PA, Weber I, *et al*. Dopamine substitution alters effective Connectivity of cortical Prefrontal, Premotor, and motor regions during complex Bimanual finger movements in Parkinson's disease. *Neuroimage* 2019;190:118–32.
32 Deuschl G, Follett KA, Luo P, *et al*. Comparing two randomized deep brain stimulation trials for Parkinson's disease. *J Neurosurg* 2019;132:1376–84.
33 Marinus J, Visser M, Stiggelbout AM, *et al*. A short scale for the assessment of motor impairments and disabilities in Parkinson's disease: the SPES/SCOPA. *J Neurol Neurosurg Psychiatry* 2004;75:388–95.
34 Martínez-Martín P, Benito-León J, Burguera JA, *et al*. The SCOPA-motor scale for assessment of Parkinson's disease is a consistent and valid measure. *J Clin Epidemiol* 2005;58:674–9.
35 Jost ST, Kaldenbach M, Antonini A, *et al*. Levodopa dose equivalency in Parkinson's disease: updated systematic review and proposals. *Movement Disorders* 2023;38:1236–52. 10.1002/mds.29410 Available: https://movementdisorders.onlinelibrary.wiley.com/toc/15318257/38/7
36 Sloan JA, Cella D, Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol* 2005;58:1217–9.
37 Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003;41:582–92.
38 Martinez-Martin P, Kurtis MM. Health-related quality of life as an outcome variable in Parkinson's disease. *Ther Adv Neurol Disord* 2012;5:105–17.
39 Dafsari HS, Weiß L, Silverdale M, *et al*. Short-term quality of life after Subthalamic stimulation depends on non-motor symptoms in Parkinson's disease. *Brain Stimul* 2018;11:867–74.
40 Dafsari HS, Reker P, Stalinski L, *et al*. Quality of life outcome after Subthalamic stimulation in Parkinson's disease depends on age. *Mov Disord* 2018;33:99–107.
41 Dafsari HS, Martinez-Martin P, Rizos A, *et al*. Euroinf 2: Subthalamic stimulation, apomorphine, and levodopa infusion in Parkinson's disease. *Mov Disord* 2019;34:353–65.
42 Jakobsen JC, Gluud C, Wetterslev J, *et al*. When and how should multiple imputation be used for handling missing data in randomised clinical trials – a practical guide with Flowcharts. *BMC Med Res Methodol* 2017;17:162.
43 Williams A, Gill S, Varma T, *et al*. Deep brain stimulation plus best medical therapy versus best medical therapy alone for advanced Parkinson's disease (PD SURG trial): a randomised, open-label trial. *Lancet Neurol* 2010;9:581–91.
44 Kleiner-Fisman G, Herzog J, Fisman DN, *et al*. Subthalamic nucleus deep brain stimulation: summary and meta-analysis of outcomes. *Mov Disord* 2006;21 Suppl 14:S290–304.
45 Mahajan A, Butala A, Okun MS, *et al*. Global variability in deep brain stimulation practices for Parkinson's disease. *Front Hum Neurosci* 2021;15:667035.
46 Wolke R, Becktepe JS, Paschen S, *et al*. The role of levodopa challenge in predicting the outcome of Subthalamic deep brain stimulation. *Mov Disord Clin Pract* 2023;10:1181–91. 10.1002/mdc3.13825 Available: https://movementdisorders.onlinelibrary.wiley.com/toc/23301619/10/8
47 Merello M, Nouzeilles MI, Arce GP, *et al*. Accuracy of acute levodopa challenge for clinical prediction of sustained long-term levodopa response as a major criterion for idiopathic Parkinson's disease diagnosis. *Mov Disord* 2002;17:795–8.
48 Lhommée E, Wojtecki L, Czernecki V, *et al*. Behavioural outcomes of Subthalamic stimulation and medical therapy versus medical therapy alone for Parkinson's disease with early motor complications (EARLYSTIM trial): secondary analysis of an open-label randomised trial. *Lancet Neurol* 2018;17:223–31.
49 Witt K, Daniels C, Reiff J, *et al*. Neuropsychological and psychiatric changes after deep brain stimulation for Parkinson's disease: a randomised, Multicentre study. *Lancet Neurol* 2008;7:605–14.
50 Witte LP, Odekerken VJJ, Boel JA, *et al*. Does deep brain stimulation improve lower urinary tract symptoms in Parkinson's disease? *Neurourol Urodyn* 2018;37:354–9.
51 Schuepbach WMM, Rau J, Knudsen K, *et al*. Neurostimulation for Parkinson's disease with early motor complications. *N Engl J Med* 2013;368:610–22.
52 Jost ST, Sauerbier A, Visser-Vandewalle V, *et al*. A prospective, controlled study of non-motor effects of Subthalamic stimulation in Parkinson's disease: results at the 36-month follow-up. *J Neurol Neurosurg Psychiatry* 2020;91:687–94.
53 Movement Disorder Society Task Force on Rating Scales for Parkinson's DiseaseThe unified Parkinson's disease rating scale (UPDRS): status and recommendations. *Mov Disord* 2003;18:738–50.
54 Evatt ML, Chaudhuri KR, Chou KL, *et al*. Dysautonomia rating scales in Parkinson's disease: Sialorrhea, Dysphagia, and constipation--critique and recommendations by movement disorders task force on rating scales for Parkinson's disease. *Mov Disord* 2009;24:635–46.