Original research

Cognitive and psychiatric outcomes in the GALAXY trial: effect of anaesthesia in deep brain stimulation

Rozemarije A Holewijn ,[1] Thomas J C Zoon ,[2] Dagmar Verbaan,[1] Isidoor O Bergfeld ,[2] Esmée Verwijk,[3,4] Gert J Geurtsen,[4] Geeske van Rooijen,[2] Pepijn van den Munckhof,[1] Maarten Bot,[1] Damiaan A J P Denys,[2,3] Rob M A De Bie,[5] P Rick Schuurman[1]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/jnnp-2023-331791).

[1]Department of Neurosurgery, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[2]Department of Psychiatry, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[3]Amsterdam Neuroscience, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[4]Department of Medical Psychology, Amsterdam University Medical Centers, Amsterdam, The Netherlands
[5]Department of Neurology, Amsterdam University Medical Centers, Amsterdam, The Netherlands

Correspondence to
Thomas J C Zoon, Department of Psychiatry, University of Amsterdam, Amsterdam, The Netherlands; t.j.zoon@amsterdamumc.nl

RH and TZ contributed equally.

RAH and TJCZ are joint first authors.

Received 4 May 2023
Accepted 23 August 2023

To cite: Holewijn RA, Zoon TJC, Verbaan D, *et al*. *J Neurol Neurosurg Psychiatry* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/jnnp-2023-331791

ABSTRACT

Background This study aims: (1) To compare cognitive and psychiatric outcomes after bilateral awake versus asleep subthalamic nucleus (STN) deep brain stimulation (DBS) surgery for Parkinson's disease (PD). (2) To explore the occurrence of psychiatric diagnoses, cognitive impairment and quality of life after surgery in our whole sample. (3) To validate whether we can predict postoperative cognitive decline.

Methods 110 patients with PD were randomised to receive awake (n=56) or asleep (n=54) STN DBS surgery. At baseline and 6-month follow-up, all patients underwent standardised assessments testing several cognitive domains, psychiatric symptoms and quality of life.

Results There were no differences on neuropsychological composite scores and psychiatric symptoms between the groups, but we found small differences on individual tests and cognitive domains. The asleep group performed better on the Rey Auditory Verbal Learning Test delayed memory test (f=4.2, p=0.04), while the awake group improved on the Rivermead Behavioural Memory Test delayed memory test. (f=4.4, p=0.04). The Stroop III score was worse for the awake group (f=5.5, p=0.02). Worse scores were present for Stroop I (Stroop word card) (f=6.3, p=0.01), Stroop II (Stroop color card) (f=46.4, p<0.001), Stroop III (Stroop color-word card) (f=10.8, p=0.001) and Trailmaking B/A (f=4.5, p=0.04). Improvements were seen on quality of life: Parkinson's Disease Questionnaire-39 (f=24.8, p<0.001), and psychiatric scales: Hamilton Depression Rating Scale (f=6.2, p=0.01), and Hamilton Anxiety Rating Scale (f=5.5, p=0.02).

Conclusions This study suggests that the choice between awake and asleep STN DBS does not affect cognitive, mood and behavioural adverse effects, despite a minor difference in memory. STN DBS has a beneficial effect on quality of life, mood and anxiety symptoms.

Trial registration number NTR5809.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ In the GALAXY study, a single-centre, randomised clinical trial, the incidence of a composite score expressing cognitive, mood and behavioural effects after subthalamic nucleus (STN) deep brain stimulation (DBS) surgery under local anaesthesia was not higher than after DBS surgery under general anaesthesia.

WHAT THIS STUDY ADDS

⇒ This in-depth analysis of the neuropsychological and psychiatric data of the GALAXY study reinforces the conclusion of the primary analysis that the anaesthesia method does not affect cognitive, mood and behavioural adverse effects.

⇒ Both STN DBS performed under local (awake) and general anaesthesia (asleep) did have a strong beneficial effect on quality of life, mood and anxiety symptoms.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study endorses the current development in clinical practice to replace awake DBS surgery with asleep DBS surgery for Parkinson's disease. Abandoning awake DBS surgery, which can be considered as a burdensome surgical procedure, contributes to a more patient-friendly surgical treatment of patients with Parkinson's disease.

INTRODUCTION

Deep brain stimulation (DBS) in the subthalamic nucleus (STN) is an effective treatment for patients with Parkinson's disease (PD) who experience response fluctuations despite optimal medical treatment.[1] In current practice, DBS surgery is often performed under local anaesthesia (LA) to enable intraoperative neurological testing. This is burdensome for patients who are awake during frame fixation and burr hole drilling, and have to endure clinical evaluations throughout the procedure while being restricted from their Parkinsonian medication.[2–4]

Neurological testing is only one of three methods that are used to guide optimal electrode placement, in addition to imaging of the target nucleus and microelectrode recordings to confirm positioning of the electrode in the nucleus. Due to advancements in MR imaging direct visualisation of the STN is of sufficient quality to guide electrode placement directly. Furthermore, during surgery microelectrode recordings can confirm specific STN neuronal activity in the preoperatively image-based defined target area. Finally, introduction of

Table 1 Baseline demographic and clinical characteristics

	Local anaesthesia (N=56)	General anaesthesia (N=54)
Age—years, mean (SD) (range)	60.0 (7.4) (36–73)	61.3 (7.9) (41–75)
Age at onset of Parkinson's disease—years, mean (SD)	49.1 (7.2)	50.7 (8.8)
Male sex, no. (%)	40 (71)	38 (70)
Duration of Parkinson's disease—years, mean (SD)	10.8 (5.3)	10.6 (5.0)
Duration of use of medication for Parkinson's disease—years, mean (SD)	10.4 (5.1)	10.3 (4.7)
On-drug phase Hoehn and Yahr stage—no. (%)		
1	1 (2)	0 (0)
2	47 (84)	43 (80)
3	5 (9)	10 (19)
4	3 (5)	1 (2)
5	0 (0)	0 (0)
Levodopa equivalent daily dose—mean (SD)	1567.6 (555.2)	1550.6 (599.4)
Difference in MDS-UPDRS ME score in ON-drug vs OFF-drug phase >40%, no. (%)	50 (89)	48 (89)
Mattis Dementia Rating Scale—mean (SD)	139.7 (3.1)	139.9 (2.5)
National Adult Reading Test IQ	107.25 (14.7)	105.59 (19.5)
PD-CRS	99.4 (14.7)	100.0 (12.8)

MDS-UPDRS ME, Movement Disorder Society Unified Parkinson's Disease Rating Scale Motor Examination; PD-CRS, Parkinson's Disease-Cognitive Rating Scale.

intraoperative imaging facilitates direct confirmation of adequate electrode placement. These advancements in the workflow of DBS surgery obviate the requirement of neurological testing for target determination, allowing for surgery under general anaesthesia (GA).[2 5 6] In the recent General Anesthesia versus Local Anesthesia in stereotaXY (GALAXY) trial, we compared bilateral STN DBS under LA and bilateral STN DBS under GA, demonstrating that there is no difference between DBS surgery under LA and STN DBS under GA with respect to symptomatic and functional improvement 6 months after surgery and on a composite score for cognition, mood and behaviour.[7] In the current report, we will describe the cognitive and psychiatric outcomes of the patients 6 months after STN DBS surgery under either LA or GA in the GALAXY trial. Our objectives are to compare cognitive and psychiatric outcomes 6 months after bilateral STN DBS surgery under either LA or GA for PD and to explore the occurrence of psychiatric symptoms, cognitive impairment, quality of life and dopaminergic medication reduction 6 months after STN DBS in our whole sample. Additionally, we try to validate whether a select set of neuropsychological tests can predict cognitive decline.[8] We expect that the burden of undergoing awake surgery (ie, LA) could contribute to the risk of adverse effects concerning psychiatric outcome and hypothesise that STN DBS under GA would reduce cognitive and psychiatric adverse effects.[5 9]

METHODS

Trial design

The GALAXY trial was a prospective, randomised, open-label, blinded endpoint study comparing STN DBS surgery either under LA following the current standard practice (n=56) or under GA (n=54) and assessed the cognitive, mood and behavioural adverse effects in addition to the functional and symptomatic effectiveness of DBS. Patients were included if they suffered from idiopathic PD with bradykinesia, tremor and/or rigidity, and at least one of the following symptoms despite optimal pharmacological treatment (1) severe motor response fluctuations, (2) dyskinesias or (3) painful dystonia. Exclusion criteria were(1) previous PD-related neurosurgery or (2) contraindications for DBS surgery, such as severe cognitive impairment indicated by a Mattis Dementia Rating Scale score of 120 or lower, current depression or psychosis in psychiatric evaluation or a physical disorder making surgery hazardous.[7 10] The trial design and primary outcome (composite score for cognitive, mood and behavioural adverse effects) and serious adverse events were reported in the primary manuscripts.[7 10] The trial was registered with the Netherlands Trial Register. This secondary analysis was designed following the Consolidated Standards of Reporting Trials guidelines.[11] First, the cognitive tests, psychiatric scales and clinical outcomes of the GA and LA groups are compared in a pre-test/post-test control group design. Second, the cognitive and psychiatric outcomes at 6 months were compared with baseline while omitting anaesthesia as a factor.

Surgical methods

DBS electrodes were placed bilaterally in the dorsolateral part of the STN. Dopaminergic medication was stopped in the evening before surgical procedure in all patients.

Surgery under LA. The patient underwent frame fixation, microelectrode recordings and macroelectrode stimulation under LA. Following the implantation of the permanent electrodes the stereotactic frame was removed and the patient was immediately placed under GA for implantation of extension cables and the implantable pulse generator.

Surgery under GA. The patient was placed under GA using propofol and remifentanil. Propofol was stopped for 20 min prior to microelectrode recordings. Propofol cessation lasted maximally 45 min, while high-dose remifentanil was continued. No macroelectrode stimulation was performed. The patients remained under GA for implantation of extension cables and the implantable pulse generator.

A more detailed description of the surgical procedure in both study groups has been published elsewhere.[7]

Cognitive assessment

Cognitive assessment was done at baseline in the on-drug phase and at 6 month follow-up in the on-drug phase and DBS on. Language was assessed with the Boston Naming Test (BNT) and Wechsler Adult Intelligence Scale IV (WAIS)-Similarities. Memory was tested with the Dutch version of the Rey Auditory Verbal Learning Test (RAVLT) and logical memory from the 'Story' subtest of the Rivermead Behavioural Memory Test (RBMT), both tests have immediate and delayed recall scores, with an index score. Attention and psychomotoric functioning was tested with the Trail Making Test (TMT)-A, the Stroop Color-Word Test (Stroop)-I and Stroop-II. Executive function was measured by the TMT-B, TMT-B/A, Stroop-III, Stroop interference and letter fluency. Visuospatial function was assessed by using the Judgement of Line Orientation (JOLO).[12–19] The outcomes of the TMT, Stroop, letter fluency, WAIS-IV Similarities, RBMT, JOLO, BNT and the RAVLT were converted into T-scores by age and education level correction.[20] A higher score indicates a better performance.

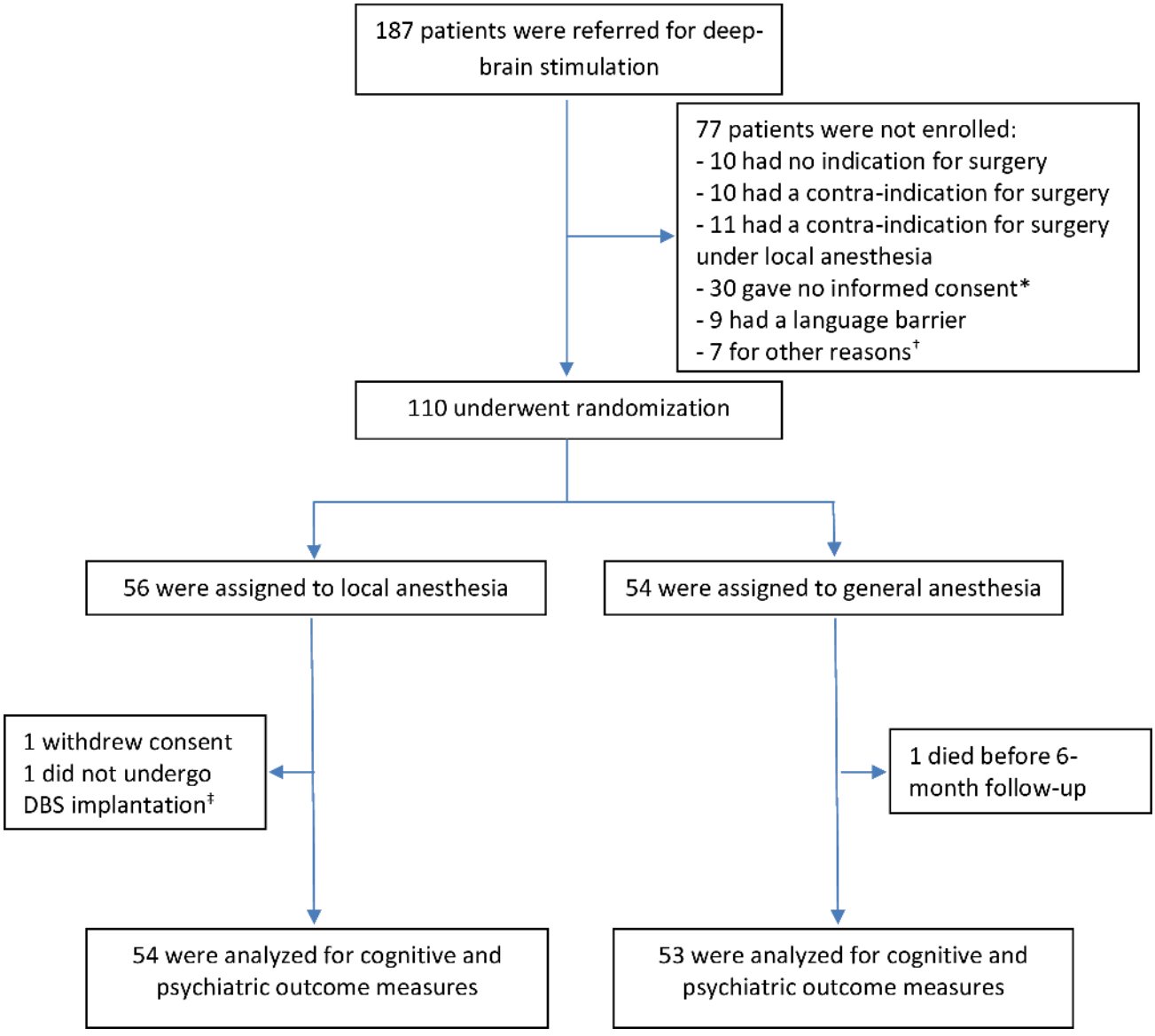

Figure 1 CONSORT Flow Diagram *All 30 patients received deep-brain stimulation (DBS); 26 had a preference for local anaesthesia; 4 had a preference for general anaesthesia. †Five patients had previous unilateral subthalamic nucleus or bilateral globus pallidus internus deep-brain stimulation, one patient lived abroad, one patient decided against surgery. ‡One patient was not eligible for deep-brain stimulation and withdrew from follow-up after randomisation due to new comorbidity. Two patients of each group refused to undergo cognitive examination after 6 months follow-up.

Clinically relevant cognitive worsening was defined as a worse score on three or more cognitive tests based on a Reliable Change Index of −1.645 or less in more than one domain (language, memory, executive function, visuospatial function, attention, psychomotoric functioning) of the neuropsychological examination 6 months after surgery compared with baseline using the corrected T-score.[8] A risk assessment predicting increased chance of cognitive decline after DBS was based on the results of the weighted average on the preoperative Trailmaking B and Stroop Color-Word Card scores.[8] Patients with a mean average T-score less than 40 were indicated as having a higher than average risk for postoperative cognitive decline.[8]

Psychiatric scales

The Hamilton Depression Rating Scale (HAM-D) consists of 17 questions with a score between 0 and 52, a score of 8 or more is indicative of depressive symptoms. The Hamilton Anxiety Rating Scale (HAM-A) consists of 14 questions with a score between 0 and 56, with a score of 18 is indicative for the presence of an anxiety disorder. Suicidal ideation was assessed using the Columbia-Suicide Severity Rating Scale consists of 20 questions, the number of questions answered with 'yes' was rated as outcome with a higher score indicating more suicidal behaviour.[21] The Starkstein Apathy Scale consists of 14 questions, with a score between 0 and 42 and the cut-off for apathy is 14 or more. The Young Mania Rating Scale consists of 11 questions, the score ranges from 0 to 60 with a cut-off of 13 or more indicating a manic episode.[22–26] A higher score on all psychiatric instruments indicate more severe symptoms.

Quality of life and symptomatic outcome

The Parkinson's Disease Questionnaire-39 (PDQ-39) measures disease specific quality of life, consisting of 39 questions with a score between 0 and 100, with 100 indicating the most severe problems. The symptomatic outcome was measured by the Movement Disorder Society Unified Parkinson's Disease Rating

Scale (MDS-UPDRS) motor score, a higher score indicates more severe symptoms. Dopaminergic medication was converted to Levodopa Equivalent Daily Dosage (LEDD).[27 28]

Statistical analysis

Baseline assessments and outcome parameters will be presented in table 1. For normally distributed continuous data, a robust linear mixed effects model with a diagonal structure will be selected to analyse the GA and LA groups and to allow for baseline value adjustment. A careful step-by-step process is followed to first achieve normal distribution of the residuals and when this could not be achieved, non-parametric tests were conducted. This process is explained in detail in online supplemental description 1.[29] Regression analysis will be performed for impacting clinical variables. The three impacting variables that we choose to analyse are: (1) Changes in LEDD to account for hyperdopaminergic or hypodopaminergic symptoms, (2) Comparison of best ON preoperative and best ON postoperative (DBS on and medication on) to account for the best possible functioning of the participant and its possible effect on daily life and social functioning and (3) Comparison of worst off preoperatively and worst off postoperatively (DBS off and medication off) as an approximation of motor disease progression. Statistical analyses are performed with IBM SPSS Statistics software (IBM Corporation, New York, USA, V.25).

Results

A total of 110 patients were enrolled, between March 2015 and January 2020. Fifty-six patients were randomised to the LA group and 54 patients were randomised to the GA group (figure 1). The groups were balanced with respect to baseline characteristics.[7] There were no differences on neuropsychological composite scores and psychiatric symptoms between the groups. The proportion of patients with a predicted increased vulnerability for cognitive deterioration was 17/51 in the LA group and 11/51 in the GA group (nonsignificant). Neuropsychological follow-up data was available for 103 patients; 2 participants withdrew from the LA group (1 withdrew consent due to personal circumstances, 1 was not eligible for DBS after randomisation due to a new comorbidity), and 1 participant was excluded from follow-up in the GA group (death unrelated to treatment before follow-up). Four patients did not undergo complete repeated neuropsychological examination after 6 months, but participated in a few measurements. One hundred and two patients were analysed for the risk assessment predicting increased chance of cognitive decline, due to a missing prediction value in one patient. All completed tests and scales were included using the mixed model analysis, which resulted in differing numbers of participants per test reported (tables 1–4). Psychiatric scale outcome measures were available for 107 patients.

Occurrence of cognitive deterioration

In 8/52 (15%) participants of the LA group and in 4/51 (8%) participants of the GA group cognitive deterioration was measured as defined by 3 ≥worse scores in >1 domain, which did not statistically differ between the groups (χ^2 1.78, p=0.18).

Worse cognitive performance was predicted based on the potential risk score with a sensitivity of 0.636, specificity of 0.769 and diagnostic accuracy of 0.755, with a positive predictive value of 0.25 and negative predictive value of 0.946 (χ^2 8.11, p<0.01) (table 2).

Table 2 Prediction and observed cognitive performance

Predicted vs observed cognitive deterioration	Stable performance	Worse performance
At risk for deterioration	21	7
Not at risk for deterioration	70	4
	χ^2-statistic: 8.11	p<0.01*
Risk and course after baseline prediction	**N**	**Value**
Sensitivity	7/11	0.64
Specificity	70/91	0.77
Diagnostic accuracy	77/102	0.76
At risk and worse performance=PPV	7/28	0.25
At risk and stable performance	21/28	0.75
No risk and worse performance	4/74	0.05
No risk and stable performance=NPV	70/74	0.95

*; Statistically significant, P<0.05
NPV, negative predictive value; PPV, positive predictive value.

Between-group comparisons

Cognitive outcome

Analyses of change scores showed between-group differences in the RAVLT delayed score, with an improvement for both groups but a better score for the GA group (f=4.2, p=0.04). This effect was not present for the other memory test, the RBMT delayed score improved for LA but worsened for GA (f=4.4, p=0.04. The Stroop III score was significantly worse for both groups, but more so for the LA group (f=5.5, p=0.02) (table 3). There was no difference in changes in MDS-UPDRS motor scores and in LEDD between the LA and the GA groups (online supplemental table 1). There was no influence found of MDS-UPDRS ON/OFF scores or LEDD on any of the cognitive tests in the whole sample (online supplemental table 2).

Psychiatric outcome

There were no differences between the groups on any of the psychiatric scales as presented in table 3.

Whole sample longitudinal results

Cognitive and psychiatric outcomes at 6 months compared with baseline are presented in table 4, omitting anaesthesia as a factor. Worse scores after 6 months were present for the Stroop I (f=6.3, p=0.01), Stroop II (f=46.4, p<0.001), Stroop III (f=10.8, p=0.001), and Trailmaking B/A (f=4.5, p=0.04). Improvements were measured on the individual quality of Life scale PDQ-39 (f=24.8, p<0.001), and psychiatric HAM-D (f=6.2, p=0.01), and HAM-A (f=5.5, p=0.02).

Discussion

In this study we showed that in Parkinson's disease there is no significant difference in cognitive outcome between STN DBS surgery under LA and STN DBS surgery under GA. Only a small number of participants (10.8%) scored lower on three cognitive tests in two or more domains at 6 months after surgery. The prediction of postoperative cognitive decline based on the preoperative neuropsychological screening showed a good diagnostic accuracy and an excellent negative predictive value to identify patients who are most likely to preserve cognitive function after DBS.

It is important to note that these results are difficult to interpret for clinical practice, because of the uncertainty of the

Table 3 Cognitive outcomes LA STN DBS versus GA STN DBS

	Baseline				6 months				Analysis				
	Local anaesthesia		General anaesthesia		Local anaesthesia		General anaesthesia		Mean difference				
	N	Mean (SD) 95% CI	N	Mean (SD) 95% CI	N	Mean (SD) 95% CI	N	Mean (SD) 95% CI	LA	GA	*F*	P value	MM
Language													
BNT	53	52.5 (7.2) 50.5 to 54.5	53	52.6 (6.8) 50.7 to 54.4	53	52.3 (7.2) 50.3 to 54.2	51	52.7 (7.4) 50.6 to 54.8	0.2	+0.1	0.1	0.73	2
WAIS IV similarities	52	50.5 (9.3) 47.9 to 53.1	53	52.6 (9.3) 50.0 to 55.1	51	51.4 (10.6) 48.5 to 54.4	51	51.5 (9.3) 48.9 to 54.2	+0.9	1.0	0.3	0.60	1
Memory													
RAVLT immediate	54	43.7 (14.3) 39.8 to 47.6	54	45.9 (10.7) 43.0 to 48.8	53	43.6 (14.5) 39.6 to 47.6	51	49.5 (11.6) 46.2 to 52.7	0.1	+3.6	3.5	0.07	1
RAVLT delayed	54	42.9 (13.0) 39.4 to 46.5	54	47.5 (12.4) 44.1 to 50.9	53	43.6 (12.3) 40.2 to 47.0	51	52.3 (11.7) 49.0 to 55.6	+0.7	+4.8	4.3	0.04*	1
RAVLT delayed\immediate	53	45.6 (10.7) 42.7 to 48.6	54	48.9 (11.9) 45.7 to 52.2	53	46.8 (11.4) 43.6 to 49.9	51	53.2 (8.6) 50.8 to 55.6	+1.1	+4.2	1.4	0.24	1
RBMT immediate	54	45.0 (11.3) 41.9 to 48.1	54	47.9 (9.7) 45.2 to 50.5	52	46.9 (10.9) 43.9 to 49.9	51	45.8 (10.7) 42.8 to 48.8	+1.9	2.1	3.6	0.06	1
RBMT delayed	54	46.8 (11.8) 43.5 to 50.0	54	49.2 (10.0) 46.4 to 51.9	52	49.0 (11.8) 45.8 to 52.3	51	47.2 (10.3) 44.3 to 50.1	+2.3	2.0	4.4	0.04*	1
RBMT delayed\immediate	54	50.9 (10.2) 48.1 to 53.7	54	51.2 (10.3) 48.4 to 54.0	52	51.7 (12.6) 48.2 to 55.2	51	51.0 (12.0) 47.7 to 54.4	+0.8	0.2	0.1	0.71	1
Attention and psychomotoric functioning													
Trailmaking A	56	48.8 (13.1) 45.3 to 52.3	54	47.3 (11.3) 44.2 to 50.4	53	49.5 (17.4) 44.7 to 54.3	51	49.7 (12.9) 46.0 to 53.3	+0.6	+2.3	0.5	0.49	5
Stroop I	56	45.1 (11.7) 41.9 to 48.2	54	45.0 (10.6) 42.1 to 47.9	53	39.6 (12.9) 36.1 to 43.2	51	42.4 (12.0) 39.0 to 45.7	5.4	2.6	0.0	0.97	5
Stroop II	56	43.4 (10.8) 40.5 to 46.3	53	45.7 (9.2) 43.2 to 48.3	53	35.7 (12.6) 32.2 to 39.2	51	40.9 (8.9) 38.4 to 43.4	7.7	4.9	2.9	0.09	1
Letter fluency	54	47.1 (12.3) 43.8 to 50.5	54	50.8 (10.1) 48.0 to 53.5	52	44.4 (11.5) 41.2 to 47.7	50	49.5 (11.5) 46.3 to 52.8	2.7	1.3	0.5	0.48	1
Executive function													
Trailmaking B	54	43.1 (13.5) 39.4 to 46.8	54	44.6 (12.1) 41.3 to 47.9	48	44.0 (14.4) 39.8 to 48.2	51	42.6 (14.2) 38.6 to 46.6	+0.9	2.0	0.7	0.42	5
Trailmaking B/A	54	42.4 (12.9) 39.0 to 45.7	54	45.3 (10.8) 42.4 to 48.3	50	40.0 (14.8) 35.8 to 44.2	51	41.9 (12.7) 38.3 to 45.5	2.3	3.4	0.2	0.64	5
Stroop III	56	45.0 (11.1) 42.1 to 48.0	53	47.1 (10.1) 44.3 to 49.9	52	37.8 (11.9) 34.4 to 41.1	52	44.3 (10.7) 41.2 to 47.3	7.3	2.9	5.5	0.02*	5
Stroop interference	56	50.4 (10.6) 47.6 to 53.2	53	51.1 (11.0) 48.1 to 54.2	52	46.9 (10.7) 43.9 to 49.8	51	51.1 (11.8) 47.8 to 54.4	3.5	0.1	2.6	0.11	2
Visuospatial function													
JOLO	53	54.6 (10.2) 51.8 to 57.4	54	54.7 (9.1) 52.2 to 57.2	53	54.5 (9.2) 52.0 to 57.0	51	54.2 (11.3) 51.0 to 57.4	0.1	0.5	0.2	0.81	
Psychiatric scales and quality of life													
HAM-D	55	8.4 (6.0) 6.8 to 10.1	54	7 (6.0) 6.0 to 9.3	54	7.0 (6.1) 5.3 to 8.5	53	5.3 (5.5) 3.8 to 6.9	2.3	1.4	1.7	0.20	2
HAM-A	55	4.4 (4.1) 3.3 to 5.5	54	2.8 (3.3) 1.9 to 3.7	54	2.9 (2.8) 2.2 to 3.7	53	2.6 (3.3) 1.6 to 3.5	0.3	1.5	1.0	0.32	2
SAS	55	6.8 (4.2) 5.7 to 8.0	54	6.1 (5.0) 4.7 to 7.5	54	7.6 (6.8) 5.8 to 9.5	53	6.6 (5.8) 5.0 to 8.2	+0.5	+0.8	0.3	0.62	2
YMRS	55	0.9 (2.1) 0.3 to 1.5	54	0.6 (1.3) 0.3 to 1.0	54	0.3 (1.5) –0.1 to 0.7	53	0.4 (1.1) 0.1 to 0.8	0.2	0.6			0
C-SSRS	50	0.5 (1.2) 0.2 to 0.9	51	0.8 (1.8) 0.3 to 1.3	52	0.3 (0.9) 0.1 to 0.6	50	0.8 (2.3) 0.2 to 0.9	0.2	+0.1	0.0	0.91	5
PDQ-39	46	48.7 (21.3) 42.4 to 55.1	47	48.5 (16.8) 43.6 to 53.5	45	35.5 (23.7) 28.4 to 42.6	42	31.1 (20.0) 24.9 to 37.3	17.5	13.3	1.7	0.20	1

Cognitive test scores are expressed by the normed T-values adjusted to age.
Stroop word/color/interference; Stroop I; Stroop word card. Stroop II; Stroop color card. Stroop III: Stroop color-word card.
MM; mixed model calculation following online supplemental description 1 handling of normal/non-normally distributed data.
*; Statistically significant, P<0.05
BNT, Boston Naming Test; C-SSRS, Columbia-Suicide Severity Rating Scale; DBS, deep brain stimulation; *F*, F-score; GA, general anaesthesia; HAM-A, Hamilton Anxiety Rating Scale; HAM-D, Hamilton Depression Rating Scale; JOLO, Judgement of Line Orientation; LA, local anaesthesia; PDQ-39, Parkinson's Disease Questionnaire-39; RAVLT, Rey Auditory Verbal Learning Test; RBMT, Rivermead Behavioural Memory Test; SAS, Starkstein Apathy Scale; STN, subthalamic nucleus; WAIS-IV, Wechsler Adult Intelligence Scale, 4th edition; YMRS, Young Mania Rating Scale.

Table 4 Whole sample longitudinal cognitive outcomes

	Baseline		6 months		Mean Difference	*F*	P value	MM
	N	Mean (SD) 95% CI	N	Mean (SD) 95% CI				
Language								
BNT	107	52.5 (7.0) 51.2 to 53.9	102	52.6 (7.3) 51.1 to 53.9	0.1	0.1	0.79	5
WAIS IV similarities	105	51.5 (9.3) 49.8 to 53.3	102	51.5 (10.0) 49.5 to 53.4	0.1	0.0	0.92	5
Memory								
RAVLT immediate	108	44.8 (12.6) 42.4 to 47.2	104	46.5 (13.4), 43.9 to 49.1	+1.7	0.9	0.35	1
RAVLT delayed	108	45.2 (12.9) 42.8 to 47.7	104	47.9 (12.7) 45.4 to 50.3	+2.7	2.3	0.13	1
RAVLT delayed/immediate	107	47.3 (11.4) 45.1 to 49.5	104	49.9 (10.6) 47.9 to 52.0	+2.6	3.0	0.09	1
RBMT delayed	108	48.0 (11.0) 45.9 to 50.1	103	48.1 (11.1) 45.9 to 50.3	+0.2	0.0	0.92	1
RBMT immediate	108	46.4 (10.6) 44.4 to 48.5	103	46.3 (10.8) 44.2 to 48.5	0.1	0.0	0.92	2
RBMT delayed\immediate	108	51.1 (10.2) 49.1 to 53.0	103	51.4 (12.2) 49.0 to 53.8	+0.3	0.1	0.80	5
Attention and psychomotoric functioning								
Trailmaking A	110	48.1 (12.2) 45.8 to 50.4	104	49.6 (15.3) 46.6 to 52.5	+1.5	1.4	0.24	5
Stroop I	110	45.0 (11.1) 42.9 to 47.1	104	41.0 (12.4) 38.5 to 43.4	4.1	6.3	0.01*	1
Stroop II	109	44.5 (10.1) 42.6 to 46.5	104	38.2 (11.2) 36.1 to 40.4	6.3	46.4	<0.001*	5
Letter fluency	108	49.0 (11.3) 46.8 to 51.1	103	46.9 (11.8) 44.6 to 49.2	2.0	1.6	0.21	1
Executive function								
Trailmaking B	108	43.9 (12.7) 41.4 to 46.3	99	43.3 (14.3) 40.4 to 46.1	0.6	1.0	0.33	5
Trailmaking B/A	108	43.8 (11.6) 41.6 to 46.0	101	41.0 (13.8) 38.3 to 43.7	2.9	4.5	0.04*	5
Stroop III	109	46.1 (10.6) 44.0 to 48.1	103	41.0 (11.7) 38.7 to 43.3	5.1	10.8	0.001*	1
Stroop interference	109	50.8 (10.8) 48.7 to 52.8	103	48.9 (11.4) 46.7 to 51.2	1.8	2.2	0.14	5
Visuospatial function								
JOLO	106	54.6 (9.6) 52.8 to 56.5	104	54.3 (12.2) 52.3 to 56.3	0.3	0.3	0.57	5
Psychiatric scales and quality of life								
HAM-D	109	8.1 (6.0) 6.9 to 9.2	107	6.2 (5.9) 5.1 to 7.3	1.9	6.2	0.01*	5
HAM-A	109	3.6 (3.8) 2.9 to 4.4	107	2.8 (3.1) 2.2 to 3.3	0.9	5.5	0.02*	5
SAS	109	6.5 (4.6) 5.6 to 7.3	107	7.1 (6.3) 5.9 to 8.3	+0.7	1.9	0.17	5
YMRS	109	0.8 (1.8) 0.4 to 1.1	107	0.4 (1.3) 0.1 to 0.6	0.4	–	–	0
C-SSRS	101	0.7 (1.5) 0.4 to 1.0	102	0.6 (1.7) 0.2 to 0.9	0.1	0.19	0.66	5
PDQ-39	93	48.6 (19.0) 44.7 to 52.6	87	33.3 (21.9) 28.7 to 38.0	15.3	24.8	<0.001*	1

Cognitive test scores are expressed by the normed T-values adjusted to age.
Stroop word/color/interference; Stroop I; Stroop word card. Stroop II; Stroop color card. Stroop III: Stroop color-word card.
MM; mixed model calculation following online supplemental description 1 handling of normal/non-normally distributed data.
*; Statistically significant, P<0.05
BNT, Boston Naming Test; C-SSRS, Columbia-Suicide Severity Rating Scale; *F*, F-score; HAM-A, Hamilton Anxiety Rating Scale; HAM-D, Hamilton Depression Rating Scale; JOLO, Judgement of Line Orientation; PDQ-39, Parkinson's Disease Questionnaire-39; RAVLT, Rey Auditory Verbal Learning Test; RBMT, Rivermead Behavioural Memory Test; SAS, Starkstein Apathy Scale; WAIS-IV, Wechsler Adult Intelligence Scale, 4th edition; YMRS, Young Mania Rating Scale.

predicted outcome and its clinical relevance. We found significant differences on the RBMT delayed change scores favouring LA STN DBS. However, this effect was contradicted by the RAVLT delayed score favouring GA STN DBS on the delayed recall. These effects were not statistically significant on the immediate recall portions of the tests. The Stroop test outcomes were worse for LA STN DBS, but the difference in attention and executive function was not reproduced for the TMT-A and TMT-B outcomes.[30 31] There were no differences at all in language and complex visual perception. The statistically significant discerning outcomes of two memory tests and one executive test are not consistently in favour of either form of anaesthesia. The lack of a harmonious set of differences between LA and GA STN DBS on cognitive and psychiatric outcomes correspond with the primary results of this study, where no significant differences were found in cognitive decline.[7]

The occurrence of Post-Traumatic Stress Disorder (PTSD) after awake surgery has been of interest recently, and while PTSD symptoms were no outcome measure in this study, we did measure depression and anxiety which would likely be impacted during severe PTSD symptoms.[32–34] We expected that awake surgery might have been more traumatic for patients with PD suffering from frailty, with a higher risk of adverse outcomes.[35] However, differences between LA and GA STN DBS in depression and anxiety scores were not indicative for traumatic experiences in patients with LA STN DBS, who were awake during part of the surgery. A recent study found that the HAM-A scores after 1 month were lower in the GA group, but this effect disappeared after several months, which could mean that stressful experiences during DBS surgery usually do not develop into PTSD.[36]

The individual tests comparing baseline with 6-month follow-up suggest that some cognitive functions might worsen after STN DBS.[37] Notably, the speed tasks Stroop I, II and TMT-A showed worsening at 6-month follow-up. This effect was persistent despite increased motor function after STN DBS, and might well be a sign of cognitive decline as part of disease progression.[38] There were no indications for a learning effect. The PD-Cognitive Rating Scale, letter fluency, RAVLT, RBMT and trail making do have multiple versions, which should minimise the learning effect. Of these, none improved and the TMT-B/A score worsened at 6-month follow-up. While for the neuropsychological tests without alternative versions (ie, Stroop, JOLO, BNT, WAIS-IV) and therefore a higher likelihood for the learning effect, all three of the Stroop subtest scores worsened.

Quality of life, depression and anxiety scores all improved after 6 months of STN DBS in our sample, with an exceptional increase in quality of life scores. These findings are important for many patients who will become dependent on STN DBS for the management of refractory PD and are in line with other studies suggesting a relation between DBS of basal ganglia and improved functional performance and subsequently expanded social activities.[39–41] Successful LEDD reduction after STN DBS could also have contributed to better perceived quality of life after STN DBS due to less severe side-effects of the medication. While suicidal ideation and behaviour have been observed following DBS surgery, the participants in our sample experienced no increase of suicidal symptoms after 6 months of STN DBS.[42] An important limitation to our findings is the multiple analyses that we have performed, which increases the chance of a type I error.

There are some limitations to this study. As described in the methods, in both study groups microelectrode recordings were executed. Therefore our asleep procedure is only minimally less invasive compared with the awake procedure, due to omitting the macroelectrode stimulation. Other DBS centres also omit microelectrode recordings during GA STN DBS, resulting in less surgical passes through the brain. It is hypothesised that surgical microlesions can cause postoperative cognitive decline following STN DBS.[43 44] In this regard, a greater level of trajectories may cause a bigger microlesion effect. However, the number of microelectrode recording trajectories is not directly associated with postoperative cognitive decline.[45 46]

Furthermore, the GALAXY study was not powered on finding a difference in cognitive decline alone. Therefore it might be due to the relatively small sample size that the difference found, 15% cognitive decline in the LA group versus 8% cognitive decline in the GA group, was not statistically significant.

In summary, this in-depth analysis of the neuropsychological and psychiatric data of the GALAXY trial shows minor and inconsistent differences between STN surgery under LA and GA for PD, which reinforces the conclusion of the primary analysis that the anaesthesia method does not affect cognitive, mood and behavioural adverse effects. Both STN DBS performed under LA and GA did have a strong beneficial effect on quality of life, mood, and anxiety symptoms.

Other information

Registered in the Netherlands Trial registry, on 23 April 2016, with the protocol published in Trials.[10] Funding was reported in the primary article.[7]

Contributors RH and TZ are co-first authors, contributing equally to this manuscript and the revision, listed alphabetically. All authors contributed to the writing of the manuscript and approved the latest version. DV was involved in the experimental design and implementation of the study. IOB was involved in the statistical analysis and data interpretation. EV and GG were contacted as experts on neuropsychological testing and provided intellectual input to the revision. GvR, PvdM and MB participated in drafting the article and revised it critically with intellectual input. DD and RdB were involved in the critical revision of the final version of the manuscript. RS was principal investigator of this study and participated in the final version of the manuscript. RS is the guarantor of this study, accepting full responsibility for the work and the conduct of the study, had access to the data, and controlled the decision to publish.

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

Competing interests RH and PvdM reported grants from Dutch Brain Foundation during the conduct of the study. DD received grants from ZonMw and Boston Scientific for a trial on deep brain stimulation for depression. RMADB reported grants from Hersenstichting Charitable Organization during the conduct of the study and grants from the Netherlands Organisation for Health Research and Development, Stichting Parkinson Nederland, GE Healthcare, Medtronic, Lysosomal Therapeutics and Neuroderm, all paid to institution, outside the submitted work. RS reported personal fees from Medtronic and Boston Scientific during the conduct of the study.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants and was approved by Medisch Ethische Toetsinsgscommissie (METC) AMC, Amsterdam, The Netherlands. Reference number: 2015_25#B2019820. 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. Data available: Yes. Data types: De-identified participant data. How to access data: p.r.schuurman@amsterdamumc.nl. When available: With publication Supporting Documents. Document types: None. Additional Information: Who can access the data: With investigators whose proposed use of the data has been approved by an independent review committee from our institution that will be formed for this purpose upon request. Types of analyses: For any purpose approved by the independent review committee. Mechanisms of data availability: To gain access, data requestors will need to sign a data access agreement.

of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Rozemarije A Holewijn http://orcid.org/0000-0001-6112-411X
Thomas J C Zoon http://orcid.org/0000-0002-9198-4476
Isidoor O Bergfeld http://orcid.org/0000-0002-0601-7271

REFERENCES

1 Odekerken VJJ, van Laar T, Staal MJ, *et al*. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. *Lancet Neurol* 2013;12:37–44.
2 Liu Z, He S, Li L. General anesthesia versus local anesthesia for deep brain stimulation in Parkinson's disease: a meta-analysis. *Stereotact Funct Neurosurg* 2019;97:381–90.
3 Santos-García D, de Deus Fonticoba T, Suárez Castro E, *et al*. Quality of life and non-motor symptoms in Parkinson's disease patients with subthreshold depression. *J Neurol Sci* 2020;418:117109.
4 Yamada K, Goto S, Kuratsu J-I, *et al*. Stereotactic surgery for subthalamic nucleus stimulation under general anesthesia: a retrospective evaluation of Japanese patients with Parkinson's disease. *Parkinsonism Relat Disord* 2007;13:101–7.
5 Ho AL, Ali R, Connolly ID, *et al*. Awake versus asleep deep brain stimulation for Parkinson's disease: a critical comparison and meta-analysis. *J Neurol Neurosurg Psychiatry* 2018;89:687–91.
6 Sheshadri V, Rowland NC, Mehta J, *et al*. Comparison of general and local anesthesia for deep brain stimulator insertion: a systematic review. *Can J Neurol Sci* 2017;44:697–704.
7 Holewijn RA, Verbaan D, van den Munckhof PM, *et al*. General anesthesia vs local anesthesia in microelectrode recording-guided deep-brain stimulation for Parkinson disease: the GALAXY randomized clinical trial. *JAMA Neurol* 2021;78:1212–9.
8 Smeding HMM, Speelman JD, Huizenga HM, *et al*. Predictors of cognitive and Psychosocial outcome after STN DBS in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 2011;82:754–60.
9 Milian M, Tatagiba M, Feigl GC. Patient response to awake craniotomy - a summary overview. *Acta Neurochir (Wien)* 2014;156:1063–70.
10 Holewijn RA, Verbaan D, de Bie RMA, *et al*. General anesthesia versus local anesthesia in stereotaxy (GALAXY) for Parkinson's disease: study protocol for a randomized controlled trial. *Trials* 2017;18:417.
11 von Elm E, Altman DG, Egger M, *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806–8.
12 Stroop JR. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 1935;18:643–62.
13 Reitan RM, Wolfson D. The trail making test as an initial screening procedure for neuropsychological impairment in older children. *Arch Clin Neuropsychol* 2004;19:281–8.
14 Wechsler D. *Wechsler Adult Intelligence Scale -- Fourth Edition (WAIS-IV)*. APA PsycTests, 2008.
15 Kreutzer JS, DeLuca J, Caplan B, eds. E. VJC. National adult reading test. In: *Encyclopedia of Clinical Neuropsychology*. New York, NY: Springer, 2011.
16 Wilson BA, Cockburn J, Baddeley AD. *The Rivermead Behavioural Memory Test*. London: Pearson Assessment, 1985.
17 Spencer RJ, Wendell CR, Giggey PP, *et al*. Judgment of line orientation: an examination of eight short forms. *J Clin Exp Neuropsychol* 2013;35:160–6.
18 Kreutzer JS, DeLuca J, Caplan B, eds. C.R.Boston naming test. In: *Encyclopedia of Clinical Neuropsychology*. New York, NY: Springer, 2011.
19 Kreutzer JS, DeLuca J, Caplan B, eds. J. B. Rey auditory verbal learning test, Rey AVLT. In: *Encyclopedia of Clinical Neuropsychology*. New York, NY: Springer, 2011.
20 de Vent NR, Agelink van Rentergem JA, Schmand BA, *et al*. Advanced neuropsychological diagnostics infrastructure (ANDI): a normative database created from control Datasets. *Front Psychol* 2016;7:1601.
21 Emslie GJ, Ventura D, Korotzer A, *et al*. Escitalopram in the treatment of adolescent depression: a randomized placebo-controlled multisite trial. *J Am Acad Child Adolesc Psychiatry* 2009;48:721–9.
22 HAMILTON M. A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
23 HAMILTON M. The assessment of anxiety states by rating. *Br J Med Psychol* 1959;32:50–5.
24 Posner K, Brown GK, Stanley B, *et al*. The Columbia-suicide severity rating scale: initial validity and internal consistency findings from three multisite studies with adolescents and adults. *Am J Psychiatry* 2011;168:1266–77.
25 Young RC, Biggs JT, Ziegler VE, *et al*. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry* 1978;133:429–35.
26 Starkstein SE, Mayberg HS, Preziosi TJ, *et al*. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *J Neuropsychiatry Clin Neurosci* 1992;4:134–9.
27 Tomlinson CL, Stowe R, Patel S, *et al*. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. *Mov Disord* 2010;25:2649–53.
28 Goetz CG, Tilley BC, Shaftman SR, *et al*. Movement disorder society-sponsored revision of the unified Parkinson's disease rating scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Mov Disord* 2008;23:2129–70.
29 Arnau J, Bono R, Blanca MJ, *et al*. Using the linear mixed model to analyze nonnormal data distributions in longitudinal designs. *Behav Res Methods* 2012;44:1224–38.
30 Lezak MD. *Neuropsychological Assessment*. 3rd edition. New York: Oxford University Press, 1995.
31 Corrigan JD, Hinkeldey NS. Relationships between parts A and B of the trail making test. *J Clin Psychol* 1987;43:402–9.
32 Huguet L, Lohkamp L-N, Beuriat P-A, *et al*. Psychological aspects of awake brain surgery in children-interests and risks. *Childs Nerv Syst* 2020;36:273–9.
33 Milian M, Luerding R, Ploppa A, *et al*. Imagine your neighbor mows the lawn: a pilot study of psychological sequelae due to awake craniotomy: clinical article. *J Neurosurg* 2013;118:1288–95.
34 Goebel S, Nabavi A, Schubert S, *et al*. Patient perception of combined awake brain tumor surgery and intraoperative 1.5-T magnetic resonance imaging: the Kiel experience. *Neurosurgery* 2010;67:594–600.
35 McMillan JM, Michalchuk Q, Goodarzi Z. Frailty in Parkinson's disease: a systematic review and meta-analysis. *Clin Park Relat Disord* 2021;4.
36 Lu Y, Chang L, Li J, *et al*. The effects of different anesthesia methods on the treatment of Parkinson's disease by bilateral deep brain stimulation of the Subthalamic nucleus. *Front Neurosci* 2022;16.
37 Cernera S, Okun MS, Gunduz A. A review of cognitive outcomes across movement disorder patients undergoing deep brain stimulation. *Front Neurol* 2019;10:419.
38 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.
39 Eghlidos Z, Rahimian Z, Vadiee G, *et al*. Effects of Subthalamic deep brain stimulation on non-motor symptoms of Parkinson's disease: a meta-analysis. *Acta Neurol Scand* 2022;146:115–25.
40 Zoon TJC, van Rooijen G, Balm G, *et al*. Apathy induced by subthalamic nucleus deep brain stimulation in Parkinson's disease: a meta-analysis. *Mov Disord* 2021;36:317–26.
41 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.
42 Costanza A, Radomska M, Bondolfi G, *et al*. Suicidality associated with deep brain stimulation in extrapyramidal diseases: a critical review and hypotheses on Neuroanatomical and Neuroimmune mechanisms. *Front Integr Neurosci* 2021;15.
43 Le Goff F, Derrey S, Lefaucheur R, *et al*. Decline in verbal fluency after subthalamic nucleus deep brain stimulation in Parkinson's disease: a microlesion effect of the electrode trajectory? *J Parkinsons Dis* 2015;5:95–104.
44 Lefaucheur R, Derrey S, Martinaud O, *et al*. Early verbal fluency decline after STN implantation: is it a cognitive Microlesion effect? *J Neurol Sci* 2012;321:96–9.
45 Mulders AEP, Temel Y, Tonge M, *et al*. The association between surgical characteristics and cognitive decline following deep brain stimulation of the Subthalamic nucleus in Parkinson's disease. *Clin Neurol Neurosurg* 2021;200.
46 Smith KM, O'Connor M, Papavassiliou E, *et al*. Phonemic verbal fluency decline after Subthalamic nucleus deep brain stimulation does not depend on number of microelectrode recordings or lead tip placement. *Parkinsonism Relat Disord* 2014;20:400–4.