

RESEARCH ARTICLE

Results of a Randomized Clinical Trial of Speech After Early Neurostimulation in Parkinson's Disease

Serge Pinto, PhD,^{1†*} Adelheid Nebel, CCC-SLP,^{2†} Jörn Rau, MSc,^{3†} Robert Espesser, PhD,^{1†} Pauline Maillochon, CCC-SLP,¹ Oliver Niebuhr, PhD,⁴ Paul Krack, MD, PhD,⁵ Tatiana Witjas, MD,⁶ Alain Ghio, PhD,¹ Marie-Charlotte Cuartero, PhD,¹ Lars Timmermann, MD, PhD,⁷ Alfons Schnitzler, MD, PhD,⁸ Helke Hesekamp, MD,⁹ Niklaus Meier, MD,⁹ Julia Müllner, MD,⁹ Thomas D. Hälbig, MD,⁹ Bettina Möller, MD,² Steffen Paschen, MD,² Laura Paschen, MD,² Jens Volkmann, MD, PhD,² Michael T. Barbe, MD, PhD,⁷ Gereon R. Fink, MD, PhD,⁷ Johannes Becker, MD,⁷ Paul Reker, MD,⁷ Andrea A. Kühn, MD, PhD,¹⁰ Gerd-Helge Schneider, MD, PhD,¹⁰ Valérie Fraix, MD, PhD,⁵ Eric Seigneuret, MD,⁵ Andrea Kistner, PhD,⁵ Olivier Rascol, MD, PhD,¹¹ Christine Brefel-Courbon, MD,¹¹ Fabienne Ory-Magne, MD,¹¹ Christian J. Hartmann, MD,⁸ Lars Wojtecki, MD, PhD,⁸ Anne Fradet, MSc,¹² David Maltête, MD,¹³ Philippe Damier, MD, PhD,¹⁴ Séverine Le Dily, MSc,¹⁴ Friederike Sixel-Döring, MD,¹⁵ Petra Benecke, CCC-SLP,¹⁵ Daniel Weiss, MD, PhD,¹⁶ Tobias Wächter, MD, PhD,¹⁶ Marcus O. Pinski, MD,¹⁷ Jean Régis, MD, PhD,⁶ Stéphane Thobois, MD, PhD,¹⁸ Gustavo Polo, MD,¹⁸ Jean-Luc Houeto, MD,¹² Andreas Hartmann, MD, PhD,⁹ Karina Knudsen, MD,² Marie Vidailhet, MD, PhD,⁹ Michael Schüpbach, MD, PhD,⁹ Günther Deuschl, MD, PhD,^{2†*} and for the EARLYSTIM Study Group

¹Aix-Marseille Univ, CNRS, LPL, Aix-en-Provence, France

²Department of Neurology, University Medical Center Schleswig-Holstein, Christian-Albrechts-University, Kiel, Germany

³Coordinating Centre for Clinical Trials of the Philipps-University of Marburg, Marburg, Germany

⁴Department of Scandinavian Studies, Frisian, and General Linguistics, Kiel University, Kiel, Germany

⁵Department of Neurology or Neurosurgery, Grenoble University Hospital, Grenoble Alpes University, Grenoble Institut des Neurosciences, Grenoble, France

⁶Aix Marseille Univ, APHM, La Timone, Neurology Department or Department of Functional and Stereotactic Neurosurgery and Radiosurgery, Marseille, France

⁷Department of Neurology, University Hospital Cologne, Cologne, Germany

⁸Department of Neurology and Institute of Clinical Neuroscience and Medical Psychology, Medical Faculty, Heinrich-Heine University, Düsseldorf, Germany

⁹Department of Neurology, Pitié-Salpêtrière Hospital, Assistance Publique Hôpitaux de Paris (APHP), INSERM, Institut du Cerveau et de la Moelle Epinière, and Centre d'Investigation Clinique (CIC), Paris, France

¹⁰Department of Neurology, Charité Hospital, Berlin University, Berlin, Germany

¹¹Department of Neurology and Centre Expert Parkinson, and INSERM U1214, Toulouse University Hospital, Toulouse NeuroImaging Centre, Toulouse, France

¹²Department of Neurology, CIC-INSERM 1402, CHU de Poitiers, Université de Poitiers, Poitiers, France

¹³Department of Neurology, Rouen University Hospital, INSERM U1073, Rouen Faculty of Medicine, Rouen, France

¹⁴CHU Nantes, INSERM, CIC1413, Hôpital Laënnec, Nantes, France

¹⁵Department of Neurology, Paracelsus-Elena-Klinik, Kassel, Germany

¹⁶Department for Neurodegenerative Diseases, Centre for Neurology, and Hertie Institute for Clinical Brain Research, University of Tübingen, Tübingen, Germany

¹⁷Department of Neurosurgery, University Hospital, Freiburg, Germany

¹⁸Hôpital Neurologique Pierre Wertheimer, Centre Expert Parkinson, Hospices Civils de Lyon, Université Claude Bernard Lyon 1 Lyon, France, and Centre de Neurosciences Cognitives, Bron, France

This is an open access article under the terms of the [Creative Commons Attribution](#) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

***Correspondence to:** Serge Pinto, Laboratoire Parole et Langage (LPL), Aix-Marseille Université, CNRS, Aix-en-Provence, France; E-mail: serge.pinto@univ-amu.fr; Günther Deuschl, Department of Neurology, University Medical Center Schleswig-Holstein, Christian-Albrechts-University, Kiel, Germany; E-mail: g.deuschl@neurologie.uni-kiel.de

†Serge Pinto, Adelheid Nebel, Jörn Rau, Robert Espesser, and Günther Deuschl contributed equally to this study.

Relevant conflicts of interest/financial disclosures: None.

Funding agencies: The EARLYSTIM study ([ClinicalTrials.gov](https://clinicaltrials.gov) number NCT00354133) was funded by the German Ministry of Research (Klinische Studien 01KG0502), French Programme Hospitalier de Recherche Clinique National (P050909), and Medtronic.

Received: 29 July 2022; **Revised:** 3 November 2022; **Accepted:** 6 November 2022

Published online 3 December 2022 in Wiley Online Library ([wileyonlinelibrary.com](https://www.wileyonlinelibrary.com)). DOI: 10.1002/mds.29282

ABSTRACT: Background: The EARLYSTIM trial demonstrated for Parkinson's disease patients with early motor complications that deep brain stimulation of the subthalamic nucleus (STN-DBS) and best medical treatment (BMT) was superior to BMT alone.

Objective: This prospective, ancillary study on EARLYSTIM compared changes in blinded speech intelligibility assessment between STN-DBS and BMT over 2 years, and secondary outcomes included non-speech oral movements (maximum phonation time [MPT], oral diadochokinesis), physician- and patient-reported assessments.

Methods: STN-DBS ($n = 102$) and BMT ($n = 99$) groups underwent assessments *on/off* medication at baseline and 24 months (in four conditions: *on/off* medication, ON/OFF stimulation—for STN-DBS). Words and sentences were randomly presented to blinded listeners, and speech intelligibility rate was measured. Statistical analyses compared changes between the STN-DBS and BMT groups from baseline to 24 months.

Results: Over the 2-year period, changes in speech intelligibility and MPT, as well as patient-reported outcomes,

were not different between groups, either *off* or *on* medication or OFF or ON stimulation, but most outcomes showed a nonsignificant trend toward worsening in both groups. Change in oral diadochokinesis was significantly different between STN-DBS and BMT groups, *on* medication and OFF STN-DBS, with patients in the STN-DBS group performing slightly worse than patients under BMT only. A signal for clinical worsening with STN-DBS was found for the individual speech item of the Unified Parkinson's Disease Rating Scale, Part III.

Conclusion: At this early stage of the patients' disease, STN-DBS did not result in a consistent deterioration in blinded speech intelligibility assessment and patient-reported communication, as observed in studies of advanced Parkinson's Disease. © 2022 The Authors. *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: deep brain stimulation; dysarthria; Parkinson's disease; speech; subthalamic nucleus

Similar to subthalamic nucleus deep brain stimulation (STN-DBS) at the advanced stage of Parkinson's disease (PD),¹ stimulation in patients with relatively early-stage motor complications (ie, fluctuations and dyskinesia present for less than 4 years) has demonstrated robust improvements in clinical, behavioral, and quality-of-life outcomes that exceeded the effects of best medical treatment (BMT) alone over a 2-year period.² It is currently assumed that although speech loudness or voice quality can be improved in patients at an advanced stage of PD,³⁻⁶ a decrease in speech intelligibility within a year^{5,6} is a typical marker of worsening functional communication after STN-DBS.⁷ Predictive factors of such worsening have been identified, including longer disease duration, lower speech intelligibility *on* medication before surgery, more medial stimulation location, and left-lateralized non-optimal parameters.^{6,8,9} When designing the EARLYSTIM trial,² a specific prospective assessment of speech intelligibility and nonspeech oral movements was included in the randomized, multicenter, parallel-group trial of PD patients with early motor complications. To date, changes in such speech outcomes following earlier STN-DBS have not been reported. The main objective of the present EARLYSTIM-SPEECH ancillary study was to identify whether any changes in speech 24 months after baseline were specifically identified in patients with STN-DBS and BMT compared to patients receiving BMT alone. Well-established standard methodology was applied to measure changes in speech intelligibility and nonspeech oral movements as primary outcomes,¹⁰ whereas

speech-related items of clinical assessments and patient self-reported questionnaires were considered secondary and were taken from the standard assessments of the EARLYSTIM trial. We expected that for patients with STN-DBS and BMT, the outcomes would worsen to the same extent as for those with BMT alone according to disease progression, while no additional stimulation-induced worsening was expected.

Patients and Methods

Study Design

EARLYSTIM-SPEECH, an ancillary study of the EARLYSTIM trial ([ClinicalTrials.gov](https://clinicaltrials.gov) number: NCT00354133), was conducted between July 2006 and March 2012. It was based on a standard, parallel-group design STN-DBS versus BMT; followed the CONSORT statements (Fig. 1)¹¹; and was in accordance with the Declaration of Helsinki.¹² EARLYSTIM was approved by the local ethical committee for each of the 15 participating centers (seven in Germany: Berlin, Cologne, Düsseldorf, Freiburg, Kassel, Kiel, and Tübingen; eight in France: Grenoble, Lyon, Marseille, Nantes, Toulouse, Paris, Poitiers, and Rouen), with patients providing written informed consent before randomization. PD patients at an early disease stage with fluctuations or dyskinesia for less than 4 years were randomized to STN-DBS plus BMT or BMT alone in the EARLYSTIM trial. The detailed study design has been published previously.^{2,13,14}

Patients were asked to read aloud words and sentences, to produce sustained vowels, and to repeatedly vocalize syllables. All patients performed the assessments at baseline, without (*off*) and with (*on*) medication. Both groups repeated the same protocol at the second time point of the study, 24 months after baseline: the BMT group *off* and *on* medication; the STN-DBS group *off* and *on* medication, OFF and ON STN-DBS. The primary outcome of the study was intelligibility assessed by “blinded observers” (hereafter referred to as “listeners”) with randomized word and sentence presentation. Secondary outcomes were maximum phonation time (MPT), oral diadochokinesis, and clinical scores.

Speech Recordings

Speech intelligibility and nonspeech articulation tasks¹⁰ were selected to provide comparability between the two languages as much as possible, short completion time, and ease of application by nonprofessionals in multiple centers. In particular, these tasks did not demand any linguistic or higher cognitive effort and are classically used for the assessment of motor speech disorders. All patients performed the assessments at baseline, without (*off*) and with (*on*) medication. Both groups repeated the same protocol at the second time point, 24 months after baseline: the BMT group *off* and *on* medication; the STN-DBS group *off* and *on* medication, OFF and ON STN-DBS.

Speech was digitized (.wav encoded format, 44.1kHz sampling, 24-bit resolution) using portable devices (Microtrack 24/92, M-Audio-Avid, Burlington, VT, USA) similar across all centers, connected with a head-mounted microphone (AKG, model K440, Vienna, Austria) that places the microphone at 5 cm from the mouth. This distance has been controlled for and ensured by all centers. Recordings were made in quiet, non-soundproofed rooms at the hospitals. For additional information, see Supplementary Material APPENDIX S1.

Speech Intelligibility

Patients were asked to read aloud 10 words and 10 sentences provided by the French¹⁵ or German¹⁶ adaptations of the intelligibility section of the Frenchay Dysarthria Assessment, version 1.¹⁷ A manual labeling of the stimuli within the audio files allowed the automatic extraction of the stimuli, one audio file corresponding to one stimulus. Such preprocessing of the audio files was required to segment and label all productions separately, to be further presented in two speech labs in Aix-en-Provence and Kiel: Two groups of university students were selected as “blinded listeners” for assessing speech intelligibility in Germany (Kiel, $n = 54$) and France (Aix-en-Provence, $n = 30$) by listening to the patients’ production from the audio-

recordings. Selection criteria were (1) French- or German-native speakers, without any history of auditory and/or visual deficit; (2) unfamiliar with speech modifications associated with any neurodegenerative disease; (3) naive regarding the aim of the experiment; and (4) blind to the patients’ characteristics and treatment conditions. Using a specific computerized environment (PERCEVAL, <http://www2.lpl-aix.fr/~lpldev/perceval/>), each stimulus (ie, a word or a sentence) was heard and literally transcribed using a keyboard by three different listeners. Because the sentences were all constructed on the same pattern (same noun + verb; for examples: in French, “*L’enfant parle*” [the child speaks], “*L’enfant court*” [the child runs]; in German, “*Der Mann hat gesummt*” [the man hummed], “*Der Mann hat gebaut*” [the man built]), the listeners were asked to write down only the verb that differed across sentences. Listeners were asked not to transcribe anything if stimuli were unintelligible. No second listening was possible. All stimuli were presented randomly via headphones to listeners whatever the patient, treatment condition (*off* or *on* medication, OFF or ON STN-DBS), or assessment time (baseline or 24 months). The outcome measure was the percentage of words and sentences correctly understood, that is, intelligibility rates (IR).¹⁰ Inter-rater reliability kappa was calculated according to Fleiss.¹⁸

Maximum Phonation Time

Patients were instructed to take a deep breath and sustain the vowel /a/ thrice at a comfortable pitch and loudness for as long and as steadily as they could. The MPT (in seconds), which is an efficient indicator of the coordination between the respiratory and laryngeal systems, was estimated as the average of the durations of each production per patient from the audio recordings.¹⁰

Oral Diadochokinesis

A syllable alternating motion rate, or articulatory diadochokinesis,^{10,19,20} was completed by patients while repeating the syllables /ba/, /da/, and /ga/ separately in 30 seconds, as fast as possible at comfortable pitch and loudness. The syllables chosen allowed the assessment of three major supralaryngeal articulatory organs, that is, the lips (/ba/), the tongue tip (/da/), and the tongue dorsum (/ga/). These three syllables are composed of voiced consonants (/b/, /d/, and /g/), which, unlike their unvoiced pairs (/p/, /t/, and /k/), involve vibration of the vocal folds while being pronounced. This task allows the calculation of a single “diadochokinesis index” (ddki) from the audio recordings per patient, defined as the ratio between the speech proportion during the 30 seconds and the number of expirations required to complete the task. Thus, ddki

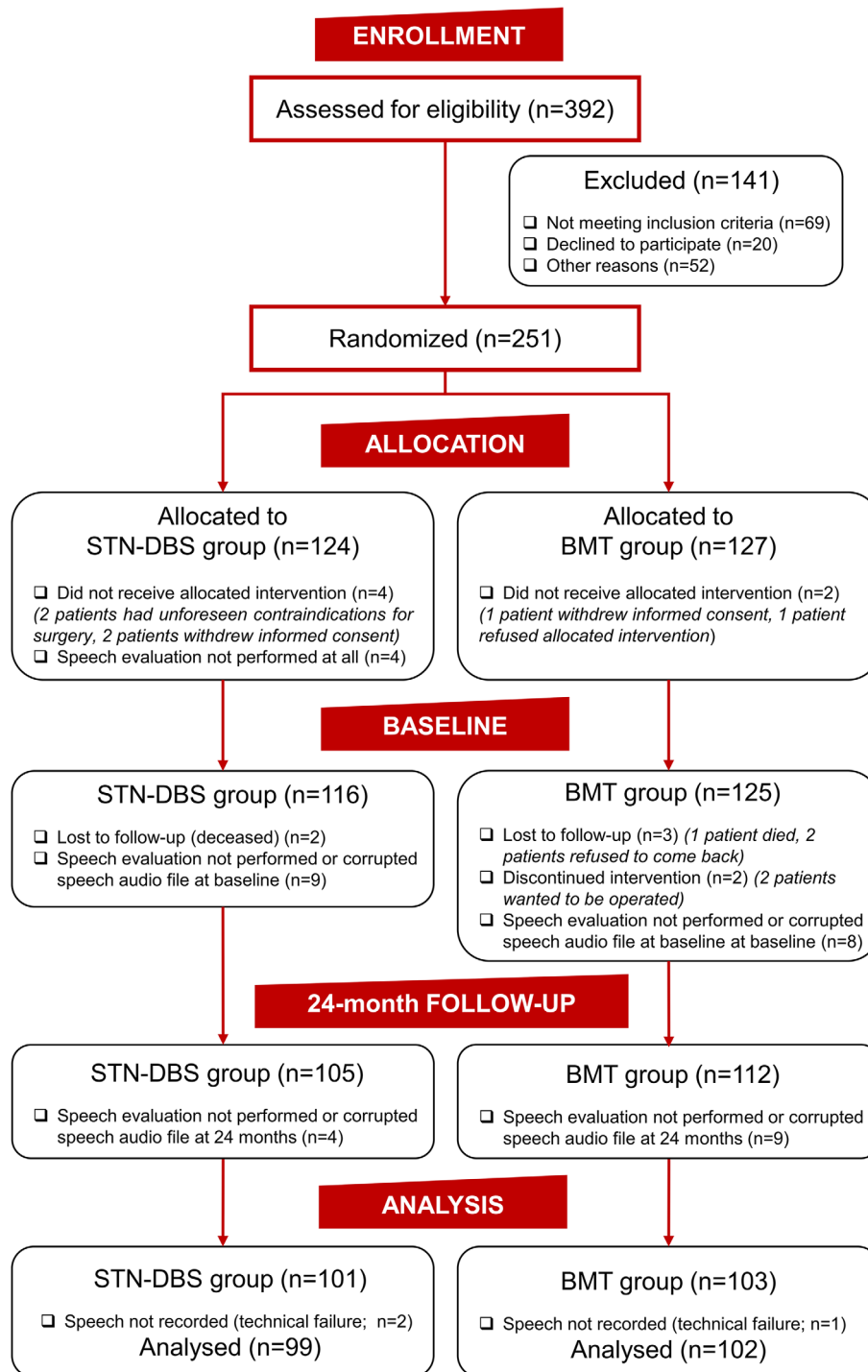


FIG. 1. EARLYSTIM-SPEECH study CONSORT flow diagram. [Color figure can be viewed at wileyonlinelibrary.com]

accounted for both speech duration and fragmentation, as well as the control of the breathing cycle across speech production.

Clinical and Patient Self-Reported Outcomes

As reported previously,² patients underwent clinical evaluation using the Unified Parkinson's Disease Rating

Scale (UPDRS),^{21,22} and speech/voice impairment was estimated using items 5 (activities of daily living, *on* medication and *ON* STN-DBS only) and 18 (motor examination, *off* and *on* medication, *ON* STN-DBS only). Patients self-reported changes in their quality of life during the last month by completing the 39-item Parkinson's Disease Questionnaire (PDQ-39),²³ *on* medication/*ON* STN-DBS; items 34 (speech) and 35

(communication) allowed for the self-estimation of the impact of speech/communication impairment in quality of life. Scoring of individual items is detailed in the legend of Table 1. At baseline, groups were also assessed using the Mattis Dementia Rating Scale and the Beck Depression Inventory (Table 1).^{24,25} These data were retrospectively obtained from the standard EARLYSTIM assessments.

Statistical Analyses

For speech and nonspeech articulation tasks, linear mixed effect models²⁶⁻²⁸ were estimated to compare the differences from baseline to 24 months between the BMT and STN-DBS groups, either *off* or *on* medication and *on* or *off* STN-DBS, and including random intercepts for center and patient. The dependent variables were the IR, MPT, and log-transformed ddki at 24 months (log transformation enhanced the normality of the distribution of the residuals). Intelligibility (logit) models were estimated separately for the French and

German languages. MPT and ddki (linear) models were estimated pooling all data, because no effect of language was expected²⁹; however, a sex variable was added for these models. All *n* nonsignificant interactions were removed using a backward stepwise procedure based on model comparison. To test for the group effect, the final model obtained was compared to the same model without the group factor (BMT, STN-DBS). Model comparisons used a parametric bootstrap method, based on 10,000 simulations, which uses the obtained data via resampling to find a better approximation of the reference distribution for the computation of the *P*-values of test hypotheses.^{30,31} This provides a better control for type 1 error. The nominal statistical significance level was set at 0.05, not correcting for multiple testing. For clinical (UPDRS) and self-reported quality of life (PDQ-39) assessments, non-parametric Wilcoxon rank-sum test for the differences between BMT and STN-DBS groups was performed to compare the differences from baseline to 24 months. The nominal statistical significance level for each *P*-

TABLE 1 Baseline characteristics

	STN-DBS (n = 99)	BMT (n = 102)	<i>P</i> -value*
Age (y)	52.4 ± 6.9	52.3 ± 6.0	0.913
Sex (women/men), n	24/75	32/70	0.260
Disease duration (y)	7.2 ± 2.9	7.6 ± 2.6	0.305
LED dose (mg/d)	938 ± 426	976 ± 403	0.517
UPDRS ^{22,23} motor III			
UPDRS III, <i>off</i> medication	33.5 ± 11.0	32.0 ± 10.5	0.324
UPDRS III, <i>on</i> medication	12.9 ± 7.79	11.9 ± 7.13	0.344
UPDRS III speech item 18, <i>off</i> medication	1.14 ± 0.67	1.10 ± 0.71	0.754
UPDRS III speech item 18, <i>on</i> medication	0.55 ± 0.63	0.60 ± 0.62	0.546
UPDRS II item 5, <i>on</i> medication, best condition	0.60 ± 0.74	0.54 ± 0.61	0.786
UPDRS II Item 5, <i>on</i> medication, worst condition	1.29 ± 0.81	1.20 ± 0.83	0.526
PDQ-39 ²⁴			
Summary index	29.1 ± 13.1	30.0 ± 11.0	0.599
Item 34 (difficulties with speech)	1.29 ± 1.03	1.51 ± 1.00	0.154
Item 35 (ability to communicate properly)	1.00 ± 1.01	1.06 ± 1.06	0.747
Mattis Dementia Rating Scale ²⁵	140.1 ± 0.3	140.8 ± 0.3	0.998
Beck Depression Inventory ²⁶	9.71 ± 5.3	10.0 ± 5.3	0.672

All values are mean ± standard deviation (SD), unless otherwise indicated. The specific subset of patients involved in our study was comparable to the entire population of the EARLYSTIM trial with regard to the PDQ-39 summary index, which was the primary outcome of the primary analysis.² Details of the LED calculation were reported in the supplementary material of the primary analysis.² The maximal score of the motor examination (UPDRS III, from items 18 to 31) of the UPDRS is 108. The scoring of the speech item 18 (from “motor examination” UPDRS III) ranges from 0 (normal), 1 (slight loss of expression, diction, and/or volume), 2 (monotone, slurred but understandable), 3 (marked impairment, difficult to understand), and 4 (unintelligible). The scoring of the speech item 5 (from “activities of daily living” UPDRS II) ranges from 0 (normal), 1 (mildly affected—no difficulty being understood), 2 (moderately affected—sometimes asked to repeat statements), 3 (severely affected—frequently asked to repeat statements), and 4 (unintelligible most of the time). The maximal score (worst) of the summary index of the PDQ-39 is 100. Communication items 34 and 35 of the PDQ-39 are scored as 0 (never), 1 (occasionally), 2 (sometimes), 3 (often), or 4 (always). The patients are required to answer whether they have had difficulty with their speech (item 34, “Due to having Parkinson’s disease, how often during the last month have you had difficulty with your speech?”) and/or felt unable to communicate (item 35: “Due to having Parkinson’s disease, how often during the last month have you felt unable to communicate with people properly?”).

Abbreviations: STN-DBS, subthalamic nucleus deep brain stimulation; BMT, best medical treatment; LED, levodopa equivalent dose; UPDRS, Unified Parkinson’s Disease Rating Scale; PDQ-39, Parkinson’s Disease Questionnaire.

**t* Test was used for the summary scores of PDQ-39 and UPDRS-III, Wilcoxon rank-sum test for their single items, and χ^2 test for the dichotome item.

TABLE 2 Changes in speech intelligibility and nonspeech parameters between baseline and 24-month follow-up

Medication	BMT group						STN-DBS group													
	Off medication			On medication			Off medication			On medication										
	Baseline	Change	N	Baseline	Change	N	Baseline	Change	N	Baseline	Change	N	Baseline	Change	N					
Primary outcome																				
Intelligibility rate—FR (%)	90.28 ± 10.45	-1.35 ± 10.04	95	90.45 ± 9.94	+0.33 ± 10.54	95	90.59 ± 11.97	-0.60 ± 11.16	95	90.33 ± 12.05	+1.29 ± 12.17	93	90.45 ± 9.26	-1.98 ± 10.28	94	91.66 ± 9.55	-1.19 ± 10.27	94	0.355	0.123
Intelligibility rate—GER (%)	87.83 ± 10.46	-0.48 ± 10.25	95	90.35 ± 9.87	-1.95 ± 10.89	95	90.27 ± 10.08	-2.30 ± 11.93	96	90.19 ± 10.00	-3.45 ± 12.60	93	91.48 ± 9.33	-5.69 ± 11.38	94	91.16 ± 9.74	-4.12 ± 11.22	94	0.123	0.586
Secondary outcomes																				
Maximum phonation time (s)	14.17 ± 7.14	-0.31 ± 6.25	201	16.42 ± 7.02	-1.85 ± 6.46	201	12.97 ± 6.61	-0.29 ± 6.82	201	12.95 ± 6.61	+1.91 ± 7.24	200	15.27 ± 6.93	-1.34 ± 7.30	201	15.45 ± 7.04	-0.86 ± 7.12	201	0.559	0.999
Diadochokinesis index	0.25 ± 0.13	-0.02 ± 0.13	194	0.29 ± 0.13	-0.03 ± 0.13	194	0.24 ± 0.12	-0.03 ± 0.12	198	0.24 ± 0.12	-0.01 ± 0.12	194	0.29 ± 0.13	-0.04 ± 0.13	190	0.28 ± 0.13	-0.03 ± 0.14	190	0.040	0.316

Baseline and within-treatment changes from baseline to 24 months are reported as mean ± standard deviations. Negative and positive changes refer to worsening and improvement at 24 months, respectively.

Abbreviations: BMT, best medical treatment; STN-DBS, subthalamic nucleus deep brain stimulation; FR, French; GER, German; N, number of patients (patients nonnative from either France or Germany were excluded from the intelligibility analyses).

*Mixed models were used to compare the change from baseline to 24-months between groups.

value was $\alpha = 0.05$, not correcting for multiple testing. For additional information, see Supplementary Material APPENDIX S1.

Results

Participants

The majority of patients of the EARLYSTIM trial participated in the EARLYSTIM-SPEECH study: Of the 124 patients assigned to the STN-DBS group, 101 completed the EARLYSTIM-SPEECH per protocol; of these, 99 patients performed the speech intelligibility/nonspeech oral movement tasks *both* at baseline and after 24 months. Of the 127 patients originally assigned to the BMT group, 103 completed the EARLYSTIM-SPEECH per protocol, and 102 patients of this group were successfully recorded *both* at baseline and after 24 months (Fig. 1, including details of reasons for exclusion). Baseline characteristics did not differ significantly between the two patient groups in terms of age, sex, PD duration, medication (levodopa equivalent dose [LED]), total motor and speech item scores of the UPDRS, and PDQ-39 summary index, which was the primary outcome parameter of the EARLYSTIM trial. Notably, no disabling dysarthria clinically estimated using the speech scores of the UPDRS was present at baseline in the BMT or STN-DBS groups (Table 1). Twenty-four months after surgery, LED in the BMT group increased up to 1214 ± 438 mg, whereas it decreased to 545 ± 438 mg in the STN-DBS group. For operated patients, mean STN-DBS parameters were as

follows: 2.74 ± 0.70 V (amplitude), 65.57 ± 13.75 μ s (pulse width), 143.12 ± 26.83 Hz (frequency), and 1246.93 ± 317.08 Ω (impedance). For these subsets of the EARLYSTIM-SPEECH study, values were comparable with those reported for the EARLYSTIM trial.

Speech Intelligibility

Inter-rater reliability kappa for the listeners of each country was calculated,¹⁸ with $P = 0.35$ and $P = 0.45$ for the French ($n = 30$) and the German ($n = 54$) groups, respectively, corresponding to a fair-to-moderate agreement.³² No statistically significant differences between the BMT and STN-DBS groups were noted from baseline to 24 months (Table 2) regarding IR of whatever the status of medication (*off*, *on*) or stimulation (OFF, ON).

MPT and Oral Diadochokinesis

The MPT did not differ significantly for the change between baseline and 24 months (Table 2). This was also the case for diadochokinesis, except in condition OFF stimulation/*on* medication, where the mean difference between groups was significant ($P = 0.04$), the STN-DBS group performing worse. Regarding this comparison, the medication change in levodopa equivalents between both groups was significant ($P < 0.0001$) between baseline and 2 years, reduced in the STN group (>300 mg), and increased in the BMT group (>200 mg).

TABLE 3 Changes in clinical (UPDRS) and patient self-reported (PDQ-39) assessments between baseline and 24-months follow-up

Outcome	Condition of assessment	Baseline		Within-treatment change from baseline to 24 months	
		BMT group	STN-DBS group	BMT group	STN-DBS group
UPDRS III—motor score	<i>Off</i> medication/ON STN-DBS	32.0 ± 10.5	33.5 ± 11.0	-1.00 ± 9.98	$-17.03 \pm 11.13^{**}$
	<i>On</i> medication/ON STN-DBS	11.9 ± 7.13	12.9 ± 7.79	$+1.11 \pm 6.35$	$-2.82 \pm 8.11^{**}$
UPDRS III—item 18	<i>Off</i> medication/ON STN-DBS	1.10 ± 0.71	1.14 ± 0.67	$+0.04 \pm 0.73$	-0.08 ± 0.80
	<i>On</i> medication/ON STN-DBS	0.60 ± 0.62	0.55 ± 0.63	$+0.03 \pm 0.76$	$+0.31 \pm 0.74^*$
UPDRS II—item 5	“Best” functional condition	0.54 ± 0.61	0.60 ± 0.74	$+0.05 \pm 0.62$	$+0.07 \pm 0.89$
	“Worst” functional condition	1.20 ± 0.83	1.29 ± 0.81	$+0.14 \pm 0.97$	$-0.33 \pm 0.85^{**}$
PDQ-39—summary index	<i>On</i> medication/ON STN-DBS	30.0 ± 11.0	29.1 ± 13.1	$+0.59 \pm 11.9$	$-7.12 \pm 13.8^{**}$
PDQ-39—item 34	<i>On</i> medication/ON STN-DBS	1.51 ± 1.00	1.29 ± 1.03	$+0.06 \pm 0.94$	$+0.41 \pm 1.27$
PDQ-39—item 35	<i>On</i> medication/ON STN-DBS	1.06 ± 1.06	1.00 ± 1.01	$+0.03 \pm 0.98$	$+0.05 \pm 1.38$

Baseline and within-treatment changes from baseline to 24 months are presented as mean \pm standard deviations. Negative and positive changes refer to improvement and worsening at 24 months, respectively. Wilcoxon rank-sum statistics were performed to compare the difference from baseline to 24 months between the BMT and the STN-DBS groups; the statistical significance was set as $*P < 0.05$ and $**P < 0.0001$.

Abbreviations: BMT, best medical treatment; STN-DBS, subthalamic nucleus deep brain stimulation; UPDRS, Unified Parkinson's Disease Rating Scale; PDQ-39, Parkinson's Disease Questionnaire on self-reported quality of life.

Clinical and Patient Self-Reported Outcomes

The UPDRS motor score improved significantly more in the STN-DBS group than in the BMT group both in the *off* ($P < 0.0001$) and *on* medication ($P < 0.0001$) conditions (Table 3). The between-group comparison of the changes in the UPDRS speech item 18 was found to be significantly worse ($P = 0.012$) for *on* medication and ON stimulation conditions in the STN-DBS group. In contrast, there was no significant difference between groups *off* medication ($P = 0.171$). The “patient-reported” speech item 5 of UPDRS II in the *worst functional* condition differed between groups *on* medication ($P = 0.001$), in favor of STN-DBS patients displaying a lower (better) score on average. There was no significant difference between groups with regard to the *best functional* condition ($P = 0.830$). Changes in PDQ-39 scores from baseline to 24 months are also summarized in Table 3. The between-group comparison of the changes from baseline to 24 months differed significantly ($P < 0.0001$) for the PDQ-39 summary index in favor of the STN-DBS group. The changes in the PDQ-39 single speech and communication items were not significantly different between the two groups (item 34: $P = 0.089$; PDQ-39 item 35: $P = 0.816$).

Discussion

Speech worsening after STN-DBS for PD may be attributed to disease progression, surgical lesions, stimulation itself, or a combination of these factors. Research over more than two decades related to speech outcomes after STN-DBS in advanced patients with PD has so far provided variable results due to different assessments, protocols, tasks, and stimulation-related parameters; furthermore, cohorts were relatively small and heterogeneous regarding the duration of disease or severity dysarthria.³³ Some speech dimensions can be improved by STN-DBS at optimal parameters^{5,6} or can be worsened by suboptimal stimulation.⁵ Misplacement of electrodes or wrongly chosen stimulation parameters may also compromise speech.³ Nevertheless, STN-DBS with or without medical treatment at advanced disease stages has been reported to deteriorate speech intelligibility,^{6,9} particularly in patients presenting with dysarthria before surgery.³⁴

The EARLYSTIM-SPEECH study included speech intelligibility as the primary outcome and parameters of nonspeech oral movements and clinical and patient self-reported scores as secondary outcomes. Speech intelligibility was assessed by naive listeners, and over the 2-year period, changes in speech intelligibility were not different between the STN-DBS and BMT patient groups, either *off* or *on* medication or OFF or ON stimulation. Much care was taken to measure this as objectively as possible, and the blinding and randomization

was well controlled. In addition, parameters of non-speech oral movements were used to monitor the effects of STN-DBS and BMT on speech organ functioning in PD. These parameters capture typical motor deficiencies of oral production in PD, such as the MPT and articulatory diadochokinesis used in our study, and have proven sensitivity in earlier studies (eg, Pinto and colleagues,¹⁰ Lowit and colleagues,¹⁹ Rusz and colleagues²⁹). Overall, few differences were found between our patient groups. In particular, changes in MPT were not significantly different between patient groups, either *off* or *on* medication or OFF or ON stimulation. Oral diadochokinesis was significantly different between groups *on* medication only, when STN-DBS was turned OFF. In this treatment condition, there was significant worsening in STN-DBS patients, which was compensated when stimulation was turned ON. The most likely explanation for this worsening was the significant reduction in dopaminergic medication down to 545-mg levodopa equivalent daily dose after 24 months in the STN-DBS group.²

Our additional approach to capture stimulation-related changes was to take also at hand further secondary outcomes that address speech and communication issues, namely the individual items of PDQ-39 (items 34 and 35) and UPDRS II (item 5) self-assessments (patient perspective), and the clinical assessment (physician perspective) of the UPDRS III (item 18). Interesting differences in opposite directions could be observed: The two items of the PDQ-39 on patient-reported speech and communication did not show significant differences between groups, and UPDRS item 5 (patient-reported) improved even slightly in the STN-DBS patient group compared to the BMT group. On the contrary, the between-group comparison of the changes in the UPDRS speech item 18, *on* medication and ON STN-DBS, was statistically significant ($P = 0.012$) in favor of the BMT group. To quantify the meaning of this finding statistically, the “relative effect”³⁵ was determined, which is 0.59 (95% confidence interval: 52.1; 66.4). A relative effect of 0.59 corresponds statistically to a “small,” although significant effect, as a value of 0.50 would have been attributed to chance alone. In addition, the psychometric quality of these individual items has its limitations and is not comparable with a comprehensive patient- or expert-assessed multidimensional speech questionnaire. Our clinical interpretation of the changes induced by STN-DBS at this disease stage is that they are still questionable and, when they occur, they seem to be small. Anyway, these results contrast strongly with the results of speech performance in patients with advanced PD. At these later stages, speech intelligibility assessments^{36,37} and other subjective/perceptual measures^{5,38} usually show an undisputable deterioration over such a long observation period.

Why does STN-DBS cause no or only minimal deterioration of speech at this disease stage? Unlike other studies comparing STN-DBS and BMT (eg, Tripoliti and colleagues⁹), one crucial difference is that the PD patients in our EARLYSTIM-SPEECH study had a much shorter disease duration; they were less affected by the disease itself, experienced less dyskinesia and motor fluctuations, and did not present with disabling dysarthria at baseline (Table 1).² This may be the most important factor contributing to the benign speech outcome after early STN-DBS. A possibility for speech deterioration is that unintended lesions during surgery may cause worsening of speech, as it has been shown for some neuropsychological outcomes.³⁹ This is unlikely in our study, as we would have observed more pronounced differences between the two groups. Furthermore, the stimulation strength has been identified as one of the factors deteriorating speech.³ It is well known that stimulation parameters need to be increased to alleviate the progressive worsening of symptoms at later disease stages. In parallel, speech impairment changes from typical PD dysarthria to a specific dysarthria combining both disease and STN-DBS-induced features, with perceptual and acoustic characteristics that are not only hypokinetic but also spastic, ataxic, or even flaccid.⁵ This confirms a current diffusion toward the cortico-bulbar tract^{3,5} and/or the cerebello-thalamo-cortical fibers.^{9,40} Thus, stimulation strength could be a factor, but the total electrical energy delivered in 1 second ($V^2 \times \mu s \times Hz/Ohm$) was found to be lower (61.9) for EARLYSTIM⁴¹ than for studies with more advanced patients.^{42,43} This may be another important factor why stimulation-induced deterioration of speech was lower in our patient group compared to more advanced patients.

Our study has limitations. Methodological issues impeded strict calibration of sound pressure levels for the majority of recordings. As a normalized loudness calculation was not possible, we decided not to consider any loudness measure in our analyses, which explains why we did not include assessment of hypophonia and monoloudness in the present study. Also, our tasks were not designed for monopitch calculation. Extended speech assessments would have been desirable, and we acknowledge that several other acoustic variables could have been calculated, as recommended in most recent guidelines.⁴⁴ However, our protocol was designed to be performed quickly by non-experts of speech, and the EARLYSTIM protocol included numerous assessments of other important functions resulting in time constraints. But, consistent with a previous trial,¹⁰ our study was based on validated methodology and had as its main objective the evaluation of the three specific and predetermined parameters of intelligibility, MPT, and oral diadochokinesis. In the future, correlation analysis in

the STN-DBS group between the location of the active stimulated contact and outcomes of speech intelligibility and nonspeech oral movements would be interesting. The strength of our study is the large patient population (more than 200 PD patients) compared in a prospective study and randomized to a powerful intervention or BMT with a rigid design over a 2-year period, which is unprecedented in the literature so far.

Conclusion

This largest interventional study of speech intelligibility and nonspeech oral movements to date has compared STN-DBS against medication-only treatment in PD patients with early fluctuations. We showed that STN-DBS maintained these outcomes at the same level as BMT, as well as patient-reported outcomes, but unlike the neurologist speech evaluation. Weighting the result, our overall clinical conclusion is that at this early disease stage—in contrast to later disease stages—neither changes in objectively measured nonspeech movement parameters and speech intelligibility nor limitations of self-assessed communication caused by STN-DBS can be found. Longer-term results from the EARLYSTIM-SPEECH study will provide insight into whether progressive deterioration of speech in STN-DBS at earlier stage PD is different or similar to that observed in patients with advanced PD. ■

Acknowledgments: We acknowledge the substantial contribution of Val Stoker (Medtronic) to the organization of the EARLYSTIM trial and the EARLYSTIM-SPEECH ancillary study. We thank all patients who participated, monitors, and project managers, study nurses charged with the recordings and documentations. We also thank L. Jankowski and A. Purson for their helpful assistance in French speech data labeling and preprocessing and BA students E. Jäger, M. Meiners, and S. Martin for organizing the German assessment of the intelligibility. Open Access funding enabled and organized by Projekt DEAL.

Data Availability Statement

Data will be available for further use based on a consent by the steering committee of the study.

References

1. Krack P, Volkman J, Tinkhauser G, Deuschl G. Deep brain stimulation in movement disorders: from experimental surgery to evidence-based therapy. *Mov Disord* 2019;34(12):1795–1810. <https://doi.org/10.1002/mds.27860>
2. Schuepbach WMM, Rau J, Knudsen K, et al. Neurostimulation for Parkinson's disease with early motor complications. *N Engl J Med* 2013;368(7):610–622. <https://doi.org/10.1056/NEJMoa1205158>
3. Pinto S, Gentil M, Krack P, et al. Changes induced by levodopa and subthalamic nucleus stimulation on parkinsonian speech. *Mov Disord* 2005;20(11):1507–1515. <https://doi.org/10.1002/mds.20601>
4. Karlsson F, Olofsson K, Blomstedt P, Linder J, van Doorn J. Pitch variability in patients with Parkinson's disease: effects of deep brain stimulation of caudal zona incerta and subthalamic nucleus. *J Speech Lang Hear Res* 2013;56(1):150. [https://doi.org/10.1044/1092-4388\(2012/11-0333\)](https://doi.org/10.1044/1092-4388(2012/11-0333))

5. Tsuboi T, Watanabe H, Tanaka Y, et al. Distinct phenotypes of speech and voice disorders in Parkinson's disease after subthalamic nucleus deep brain stimulation. *J Neurol Neurosurg Psychiatry* 2015;86(8):856–864. <https://doi.org/10.1136/jnnp-2014-308043>
6. Tripoliti E, Zrinzo L, Martinez-Torres I, et al. Effects of subthalamic stimulation on speech of consecutive patients with Parkinson disease. *Neurology* 2011;76(1):80–86. <https://doi.org/10.1212/WNL.0b013e318203e7d0>
7. Wertheimer J, Gottuso AY, Nuno M, et al. The impact of STN deep brain stimulation on speech in individuals with Parkinson's disease: the patient's perspective. *Parkinsonism Relat Disord* 2014;20(10):1065–1070. <https://doi.org/10.1016/j.parkreldis.2014.06.010>
8. Tripoliti E, Zrinzo L, Martinez-Torres I, et al. Effects of contact location and voltage amplitude on speech and movement in bilateral subthalamic nucleus deep brain stimulation. *Mov Disord* 2008;23(16):2377–2383. <https://doi.org/10.1002/mds.22296>
9. Tripoliti E, Limousin P, Foltynie T, et al. Predictive factors of speech intelligibility following subthalamic nucleus stimulation in consecutive patients with Parkinson's disease: speech intelligibility after STN-DBS. *Mov Disord* 2014;29(4):532–538. <https://doi.org/10.1002/mds.25816>
10. Pinto S, Ferraye M, Espesser R, et al. Stimulation of the pedunculopontine nucleus area in Parkinson's disease: effects on speech and intelligibility. *Brain* 2014;137(10):2759–2772. <https://doi.org/10.1093/brain/awu209>
11. Hopewell S, Clarke M, Moher D, et al. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008;371(9609):281–283. [https://doi.org/10.1016/S0140-6736\(07\)61835-2](https://doi.org/10.1016/S0140-6736(07)61835-2)
12. World Medical Association Declaration of Helsinki. Ethical principles for medical research involving human subjects. *JAMA* 2013;310(20):2191. <https://doi.org/10.1001/jama.2013.281053>
13. Deuschl G, Schüpbach M, Knudsen K, et al. Stimulation of the subthalamic nucleus at an earlier disease stage of Parkinson's disease: concept and standards of the EARLYSTIM-study. *Parkinsonism Relat Disord* 2013;19(1):56–61. <https://doi.org/10.1016/j.parkreldis.2012.07.004>
14. 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(3):223–231. [https://doi.org/10.1016/S1474-4422\(18\)30035-8](https://doi.org/10.1016/S1474-4422(18)30035-8)
15. Auzou P, Ozsancak C, Jan M, et al. Clinical assessment of dysarthria: presentation and validation of a method. *Rev Neurol* 1998;154(6–7):523–530.
16. Enderby PM. Frenchay-Dysarthrie-Untersuchung: Handanweisung. Munich: Urban & Fischer in Elsevier; 1991.
17. Enderby P. Frenchay dysarthria assessment. *Int J Lang Commun Disord* 1980;15(3):165–173. <https://doi.org/10.3109/13682828009112541>
18. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76(5):378–382. <https://doi.org/10.1037/h0031619>
19. Lowit A, Marchetti A, Corson S, Kuschmann A. Rhythmic performance in hypokinetic dysarthria: relationship between reading, spontaneous speech and diadochokinetic tasks. *J Commun Disord* 2018;72:26–39. <https://doi.org/10.1016/j.jcomdis.2018.02.005>
20. Wang Y, Kent RD, Duffy JR, Thomas JE, Weismer G. Alternating motion rate as an index of speech motor disorder in traumatic brain injury. *Clin Linguist Phon* 2004;18(1):57–84. <https://doi.org/10.1080/02699200310001596160>
21. Fahn S, Elton R, UPDRS Program Members. Unified Parkinsons disease rating scale. In: Fahn S, Marsden CD, Goldstein M, Calne DB, eds. *Recent Developments in Parkinson's Disease*. Vol. 2. Macmillan Healthcare Information; Florham Park, New Jersey. 1987:153–163.
22. Goetz CG, Stebbins GT, Chmura TA, Fahn S, Klawans HL, Marsden CD. Teaching tape for the motor section of the unified Parkinson's disease rating scale. *Mov Disord* 1995;10(3):263–266. <https://doi.org/10.1002/mds.870100305>
23. Peto V, Jenkinson C, Fitzpatrick R, Greenhall R. The development and validation of a short measure of functioning and well being for individuals with Parkinson's disease. *Qual Life Res* 1995;4(3):241–248. <https://doi.org/10.1007/BF02260863>
24. Mattis S. Dementia Rating Scale: DRS: Professional Manual. PAR Psychological Assessment Resources; 1988.
25. Beck AT. An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4(6):561. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
26. Bates D, Maechler M, Bolker B. lme4: linear mixed-effects models using S4 classes. 2013. <http://CRAN.R-project.org/package=lme4>
27. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015;67(1):1–48. <https://doi.org/10.18637/jss.v067.i01>
28. Diggle P, Diggle P, eds. *Analysis of Longitudinal Data*. 2nd ed. Oxford: Oxford University Press; 2002.
29. Ruz J, Hlavnička J, Novotný M, et al. Speech biomarkers in rapid eye movement sleep behavior disorder and Parkinson disease. *Ann Neurol* 2021;90:62–75. <https://doi.org/10.1002/ana.26085>
30. Halekoh U, Højsgaard S. pbkrtest: Parametric bootstrap and Kenward Roger based methods for mixed model comparison. 2013. <http://CRAN.R-project.org/package=pbkrtest>
31. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Soft* 2017;82(13):553. <https://doi.org/10.18637/jss.v082.i13>
32. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33(1):159–174.
33. Brabenec L, Mekyska J, Galaz Z, Rektorova I. Speech disorders in Parkinson's disease: early diagnostics and effects of medication and brain stimulation. *J Neural Transm* 2017;124(3):303–334. <https://doi.org/10.1007/s00702-017-1676-0>
34. Hariz MI, Johansson F, Shamsgovara P, Johansson E, Hariz GM, Fagerlund M. Bilateral subthalamic nucleus stimulation in a parkinsonian patient with preoperative deficits in speech and cognition: persistent improvement in mobility but increased dependency: a case study. *Mov Disord* 2000;15(1):136–139. [https://doi.org/10.1002/1531-8257\(200001\)15:1<136::AID-MDS1021>3.0.CO;2-5](https://doi.org/10.1002/1531-8257(200001)15:1<136::AID-MDS1021>3.0.CO;2-5)
35. Ryu E, Agresti A. Modeling and inference for an ordinal effect size measure. *Statist Med* 2008;27(10):1703–1717. <https://doi.org/10.1002/sim.3079>
36. Grover T, Georgiev D, Kalliola R, et al. Effect of low versus high frequency subthalamic deep brain stimulation on speech intelligibility and verbal fluency in Parkinson's disease: a double-blind study. *J Parkinsons Dis* 2019;9(1):141–151. <https://doi.org/10.3233/JPD-181368>
37. Aldridge D, Theodoros D, Angwin A, Vogel AP. Speech outcomes in Parkinson's disease after subthalamic nucleus deep brain stimulation: a systematic review. *Parkinsonism Relat Disord* 2016;33:3–11. <https://doi.org/10.1016/j.parkreldis.2016.09.022>
38. Tanaka Y, Tsuboi T, Watanabe H, et al. Longitudinal speech change after subthalamic nucleus deep brain stimulation in Parkinson's disease patients: a 2-year prospective study. *J Parkinsons Dis* 2020;10(1):131–140. <https://doi.org/10.3233/JPD-191798>
39. Witt K, Granert O, Daniels C, et al. Relation of lead trajectory and electrode position to neuropsychological outcomes of subthalamic neurostimulation in Parkinson's disease: results from a randomized trial. *Brain* 2013;136(7):2109–2119. <https://doi.org/10.1093/brain/awt151>
40. Åström M, Tripoliti E, Hariz MI, et al. Patient-specific model-based investigation of speech intelligibility and movement during deep brain stimulation. *Stereotact Funct Neurosurg* 2010;88(4):224–233. <https://doi.org/10.1159/000314357>
41. Knudsen K, Krack P, Tonder L, et al. Programming parameters of subthalamic deep brain stimulators in Parkinson's disease from a controlled trial. *Parkinsonism Relat Disord* 2019;65:217–223. <https://doi.org/10.1016/j.parkreldis.2019.05.023>
42. 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(9):896–908. <https://doi.org/10.1056/NEJMoa060281>

43. Follett KA, Weaver FM, Stern M, et al. Pallidal versus subthalamic deep-brain stimulation for Parkinson's disease. *N Engl J Med* 2010; 362(22):2077–2091. <https://doi.org/10.1056/NEJMoa0907083>
44. Rusz J, Tykalova T, Ramig LO, Tripoliti E. Guidelines for speech recording and acoustic analyses in Dysarthrias of movement disorders. *Mov Disord* 2021;36(4):803–814. <https://doi.org/10.1002/mds.28465>

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.