Subthalamic Deep Brain Stimulation and Levodopa in Parkinson's disease: A Meta-Analysis of Combined Effects

Joaquin A. Vizcarra, MD[1]; Miguel Situ-Kcomt, MD[2]; Carlo Alberto Artusi, MD[3]; Andrew P Duker, MD[1]; Leonardo Lopiano MD, PhD[3]; Michael S. Okun, MD[4]; Alberto J. Espay, MD, MSc[1]; Aristide Merola, MD, PhD[1]

[1] Gardner Family Center for Parkinson's Disease and Movement Disorders, Department of Neurology, University of Cincinnati, Cincinnati, OH, USA

[2] Department of Neurology, University of Cincinnati, Cincinnati, OH, USA

[3] Department of Neuroscience "Rita Levi Montalcini", University of Turin, via Cherasco 15, 10126, Torino, Italy

[4] Department of Neurology, Center for Movement Disorders and Neurorestoration, McKnight Brain Institute, Gainesville FL, USA

Corresponding Author: Aristide Merola MD, PhD

260 Stetson St, Suite 2300, Cincinnati, OH 45267-0525, USA.

e-mail: merolaae@ucmail.uc.edu

Word count: 1777

Key words: Subthalamic Nucleus; Deep Brain Stimulation; Levodopa; Synergism; Parkinson

ABSTRACT

Introduction: While subthalamic nucleus deep brain stimulation (STN-DBS) and levodopa improve motor symptoms in Parkinson disease (PD) to a similar magnitude, their combined effect remains unclear. We sought to evaluate whether STN-DBS and levodopa yield differential effects on motor outcomes, dyskinesia, and activities of daily living (ADL) when combined compared to when administered alone.

Methods: We conducted a meta-analysis of all studies reporting motor, dyskinesia, and ADL outcomes after bilateral STN-DBS in PD with pre-surgical Unified Parkinson's Disease Rating Scale (UPDRS-III) in Medication-OFF and Medication-ON states and post-surgical assessments in four conditions: Stimulation-ON/Medication-ON, Stimulation-ON/Medication-OFF, Stimulation-OFF/Medication-ON, and Stimulation-OFF/Medication-OFF. Dyskinesia duration (UPDRS item 32) and ADL (UPDRS-II) were compared between high vs. low post-surgical levodopa equivalent daily dose (LEDD) reduction. Random-effects meta-analyses using generic-inverse variance were conducted. Confidence in outcomes effect sizes was assessed.

Results: Twelve studies were included (n= 401 patients). Stimulation-ON/Medication-ON was associated with an UPDRS-III improvement of -35.7 points [95% confidence interval, -40.4, -31.0] compared with Stimulation-OFF/Medication-OFF, -11.2 points [-14.0, -8.4] compared with Stimulation-OFF/Medication-ON and -9.5 points [-11.0, -8.0] compared to Stimulation-ON/Medication-OFF within 5 years. The difference was maintained beyond 5 years by -28.6 [-32.8, -24.4], -8.1 [-10.2, -5.9], and -8.0 [-10.3, -5.6], respectively. No difference was observed between Stimulation-ON/Medication-OFF and Stimulation-

OFF/Medication-ON within and beyond 5 years. Dyskinesia duration and ADL outcomes were similar in high vs. low post-surgical LEDD reduction.

Conclusion: STN-DBS and levodopa independently lessened motor severity in PD to a similar magnitude, but their combined effect was greater than either treatment alone, suggesting therapeutic synergism.

INTRODUCTION

With over 140,000 patients treated worldwide, subthalamic nucleus deep brain stimulation (STN-DBS) is an established treatment for motor complications in Parkinson disease (PD) [1]. The post-STN-DBS management, however, poses the challenge of identifying the optimal combination of dopaminergic therapies and stimulation settings.

Reduction in levodopa equivalent daily dose (LEDD) and other dopaminergic medications is widely endorsed after STN-DBS and has become an "anticipated benefit" of this surgical modality. This paradigm stems from the rationale that STN-DBS might reduce PD cardinal symptoms to a similar extent than levodopa (L-dopa) and that decreasing medications reduces postoperative dyskinesia [1,2]. On the other hand, medication reduction can elicit other problems, such as depression and apathy [3], which creates uncertainty as to the wisdom of aggressively lowering dopaminergic therapies in patients treated with STN-DBS [4]. While STN-DBS and L-dopa have been recognized as providing similar motor benefits, no systematic assessment of these two treatments combined has been performed in long-term studies [5]. In addition, the difference in motor complications and activities of daily living (ADL) between patients with high vs. low post-surgical LEDD reduction remains to be clarified.

In this meta-analysis, we sought to estimate the magnitude of difference between ON and OFF medication states and ON and OFF stimulation states to determine if STN-DBS and L-dopa may yield differential motor, dyskinesia and ADL outcomes when combined compared to when stimulation and medication are administered alone.

MATERIALS AND METHODS

This meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) and the Meta-analysis of Observational Studies in Epidemiology (MOOSE) guidelines [6,7]. Observational studies, randomized clinical trials (RCT), and non-randomized clinical trials (n-RCT) were included if meeting the following criteria: a) Surgical selection for bilateral STN-DBS, as per the Core Assessment Program for Surgical Interventional Therapies in PD (CAPSIT-PD) [8]; b) Pre-surgical assessment of motor symptoms in Medication-OFF (Med-OFF) and Medication-ON (Med-ON) conditions, as per the motor subscale of the Unified Parkinson's Disease Rating Scale (UPDRS-III); and c) Post-surgical assessment of motor symptoms in the following conditions: Stimulation-ON/Medication-ON (Stim-ON/Med-ON), Stimulation-OFF/Medication-OFF (Stim-OFF/Med-OFF), Stimulation-OFF/Medication-ON (Stim-OFF/Med-ON), and Stimulation-ON/Medication-OFF (Stim-ON/Med-OFF), using a supra-maximal L-dopa challenge dose to assess Medication-ON conditions (Supplementary Table 1), as per the CAPSIT-PD protocol [8].

Exclusion criteria were incomplete data reporting (i.e., lacking one or more of the four post-surgical CAPSIT-PD conditions) or sample sizes fewer than 5 patients. No restrictions were applied to gender, disease duration, disease severity, or DBS manufacturer.

Search Methods

We searched for eligible studies in Pubmed, Embase, Cochrane Movement Disorders Group Trials Register, Cochrane Central Register of Controlled Trials (CENTRAL), ClinicalTrials.gov, and the System for Information on Grey Literature in Europe (OpenGrey) up to December 31, 2017 using the following search terms: Parkinson disease, Parkinson,

deep brain stimulation, DBS, and follow-up (Supplementary Table 2). No language restrictions were applied.

Meta-analysis Design

We divided the analyses into short- (< 5 years) and long-term (≥ 5 years) after surgery, comparing the change in UPDRS-III in the four possible Stimulation/Medication conditions. In addition to the motor outcome (UPDRS-III), we examined ADL (UPDRS-II) and the change in the proportion of the waking day spent with dyskinesia (UPDRS item 32).

Selection of studies and Data Extraction

Abstracts were reviewed for eligibility criteria by three investigators (J.A.V., M.S., A.M.). Pertinent full-text articles were assessed and variables of interest extracted. Particular attention was paid to studies that shared the same population or published data from the same cohort at different time-points. In this scenario, the longest follow-up within each time interval (< 5 years and ≥ 5 years) was used for the analyses. Disagreements were anticipated to be settled by consensus.

Assessment of risk of bias and heterogeneity

Evidence quality was independently assessed by two investigators (J.A.V.; M.S.). For included studies, we used the Cochrane-validated "Quality Assessment Tool for Before-After Studies with No Control Group" [9]. Visual inspection of funnel plots was conducted to assess for publication bias [10]. Subsequently, the overall confidence in the effect for each outcome of interest was assessed following the Grading of Recommendations Assessment, Development and Evaluation (GRADE) methodology [11]. The degree of heterogeneity was deemed considerable if I^2 statistic was ≥ 75% and significance test (p-value) below 0.1 [12].

Measures of treatment effect

Random-effects meta-analyses using generic-inverse variance were used to pool the mean differences and standard errors of the following outcomes: a) motor score; b) dyskinesia duration; and c) ADL, at the pre-specified follow-up intervals, with 95% confidence intervals (C.I.) for these pooled estimates. Subgroup analyses were conducted to compare studies with high (≥ median) vs. low (< median) post-surgical LEDD reduction from baseline, within and beyond 5 years. All the analyses were performed in Review Manager® (RevMan, version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014).

RESULTS

Out of the 1,632 records derived from the initial search strategy (Supplementary Figure 1), 12 observational and non-randomized studies met full criteria [2,5,13–22], underwent data extraction (Table 1) and assessment for individual risk of bias (Supplementary Table 3). Pooled studies were assessed to determine the overall quality of evidence (Supplementary Table 4). The agreement was met between evaluators in all cases, and no signs of publication biases were observed in funnel plots (Supplementary Figure 2).

The total study population consisted of 401 PD patients treated with STN-DBS (n= 366 with < 5 years and n= 196 with ≥ 5 years of follow-up). Six patients had undergone previous neurosurgical procedures (n= 4 in Ostergaard and Sunde [13]; and n= 2 in Schupbach et al. [21]).

Follow up < 5 years

The Stim-ON/Med-ON condition reduced (improved) UPDRS-III by -35.7 points [95% C.I., -40.4, -31.0] compared with Stim-OFF/Med-OFF, but also by -11.2 [-14.0, -8.4] compared

with Stim-OFF/Med-ON and -9.5 [-11.0, -8.0] compared with Stim-ON/Med-OFF. No difference was observed between Stim-ON/Med-OFF and Stim-OFF/Med-ON conditions (Figure 1).

High vs. low post-surgical LEDD reduction (Figure 2) resulted in a similar improvement in dyskinesia duration (-1.4 [95% C.I. -1.5, -1.2] vs. -1.0 [95% C.I. -1.7, -0.4]; p = 0.33; $I^2 = 0$) and no significant differences in the ADL outcomes (0.6 [95% C.I. -0.5, 1.6] vs. -0.01 [95% C.I. -4.1, 4.1]; p = 0.79; $I^2 = 0$).

Follow up ≥ 5 years

The Stim-ON/Med-ON condition reduced UPDRS-III by -28.6 points [95% C.I., -32.8, -24.4] compared to Stim-OFF/Med-OFF, but also by -8.1 [-10.2, -5.9] compared with Stim-OFF/Med-ON and -8.0 [-10.3, -5.6] compared with Stim-ON/Med-OFF. No difference was observed between Stim-ON/Med-OFF and Stim-OFF/Med-ON conditions (Figure 3).

High vs. low post-surgical LEDD reduction (Figure 2) resulted in a similar improvement in dyskinesia duration (-1.1 [95% C.I. -1.3, -0.9] vs. 1.1 [95% C.I. -1.5, -0.7]; p = 0.99; $I^2 = 0$) and no significant differences in the ADL outcomes (5.6 [95% C.I. 1.0, 10.3] vs. 6.8 [95% C.I. 3.0, 10.6]; p = 0.71; $I^2 = 0$).

DISCUSSION

The results of this meta-analysis demonstrated that while there was similar individual efficacy of STN-DBS and L-dopa, their combined effect on motor severity was additive within and beyond 5 years of follow-up, with a UPDRS-III differential between 9.5 and 11.2 points in the short-term and between 8.0 and 8.1 points in the long-term. These values are

above the 3.25 point threshold considered the minimal clinically important difference (MID) for the UPDRS-III [23]. In addition, no difference was observed in the extent of dyskinesia duration improvement or in the ADL outcome between studies with high vs. low post-surgical LEDD reduction.

Taken together, these data argue against the paradigm of invariably aiming at reducing the dopaminergic tone as part of the post-surgical management of STN-DBS patients. In fact, there is no evidence that greater reduction in dopaminergic therapies might lead to better control of dyskinesia, while harnessing an additive effect between STN-DBS and L-dopa may be particularly relevant at advanced disease stages, in which the main sources of disability are relatively resistant to conventional medical and surgical therapies alone, such as gait, balance, speech, swallowing, and cognitive impairments [5,19,24,25]. Further, lower reduction in dopaminergic medications, as reported after unilateral STN-DBS [4], might result in lower incidence of apathy and depression [26].

The underlying mechanism behind the additive effect of stimulation and medication might reflect the complementary effects of both intervention, modulating both dopaminergic and non-dopaminergic pathways including, but not limited to, cholinergic and adrenergic circuits [27]. Further, there may be a differential modulation of nigro-striatal dopaminergic pathways between STN-DBS and L-dopa in advanced PD, when L-dopa response may be limited by aberrant synaptic plasticity, reduced density in D3 striatal dopamine receptors [28], and progressive loss of dopaminergic neurons in the caudal putamen [29]. Whether STN-DBS effectiveness might be hampered by advanced degeneration of dopaminergic and non-dopaminergic pathways remains unclear [30].

Some limitations attenuate the strength of our conclusions. First, we included uncontrolled, non-randomized clinical studies of small sample sizes, which lowers the confidence in the overall effect. Although heterogeneity was present, a meta-regression was not performed due to the limited number of studies included in the analyses [12]. To minimize these shortcomings, we carefully assessed the individual and overall quality of included studies as per the Cochrane and GRADE handbook recommendations [11]. Second, although standard in the pre-surgical evaluation of patients, the comparison between Med-OFF and a supra-therapeutic Med-ON condition may not represent an accurate estimate of the patient daily response to L-dopa therapy. Relatedly, not all measurements of UPDRS-III post-operative response to management may have accurately represented ecologically valid settings, such as their functioning at home. Also, the data on dyskinesia based on UPDRS item 32 lacks characterization of semiology, severity, and functional impairment and may not be sensitive enough to treatment. While the assessment of dyskinesia duration and ADL reflected information gathered from daily-living clinical experience, the possibility exists that subgroup analyses might be underpowered to detect small differences between high vs. low post-surgical LEDD reduction subgroups. Unfortunately, data from the full UPDRS-IV and from quality of life scales were not consistently available and would have prevented the construction of a pooled estimate with meta-analysis.

In conclusion, our data confirm the comparable efficacy of STN-DBS and L-dopa but also suggest an additional benefit to be attained by their combined application, which is greater than each treatment alone. While a post-surgical reduction in dopaminergic therapies may be necessary to ameliorate dopaminergic side effects such as sedation, hallucinations, impulsivity, and orthostatic hypotension, our findings suggest that for any other reasons, including “simplification” of the daily therapeutic schedule, a significant reduction in

dopaminergic tone might preclude the potentially additive effect of STN-DBS and levodopa on PD motor symptoms. Further controlled, prospective studies will be needed to clarify optimal therapeutic strategies, but the available evidence supports the notion that clinicians may be missing an important source of outcome optimization in PD by aggressively reducing medications after bilateral STN-DBS.

ETHICAL STANDARDS

The manuscript does not contain clinical studies or patient data.

FINANCIAL DISCLOSURE AND CONFLICTS OF INTEREST

Dr. Vizcarra reports no conflict of interest.

Dr. Situ-Kcomt reports no conflict of interest.

Dr. Artusi reports no conflict of interest.

Dr. Duker has previously received honoraria but has not received industry support in the last 36 months.

Dr. Lopiano has received honoraria for lecturing and travel grants from Medtronic, UCB Pharma and AbbVie.

Dr. Okun serves as a consultant for the National Parkinson Foundation, and has received research grants from NIH, NPF, the Michael J. Fox Foundation, the Parkinson Alliance, Smallwood Foundation, the Bachmann-Strauss Foundation, the Tourette Syndrome Association, and the UF Foundation. His DBS research is supported by: R01 NR014852 and R01NS096008. He has previously received honoraria, but in the past >60 months has received no support from industry. He has received royalties for publications with Demos, Manson, Amazon, Smashwords, Books4Patients, and Cambridge (movement disorders books). He is an associate editor for New England Journal of Medicine Journal Watch Neurology. He has participated in CME and educational activities on movement disorders (in the last 36) months sponsored by PeerView, Prime, QuantiaMD, WebMD, Medicus, MedNet, Henry Stewart, and by Vanderbilt University. The institution and not Dr. Okun receives grants from Medtronic, Abbvie, Allergan, and ANS/St. Jude, and the PI has no financial interest in these grants. He has participated as a site PI and/or co-I for several NIH, foundation, and industry sponsored trials over the years but has not received honoraria.

Dr. Espay has received grant support from the NIH, Great Lakes Neurotechnologies and the Michael J Fox Foundation; personal compensation as a consultant/scientific advisory board member for Abbvie, TEVA, Impax, Acadia, Acorda, Cynapsus/Sunovion, Lundbeck, and USWorldMeds; publishing royalties from Lippincott Williams & Wilkins, Cambridge University Press, and Springer; and honoraria from Abbvie, UCB, USWorldMeds, Lundbeck, Acadia, the American Academy of Neurology, and the Movement Disorders Society. He serves on the editorial boards of the Journal of Parkinson's Disease and Parkinsonism and Related Disorders.

Dr. Merola is supported by NIH (KL2 TR001426) and has received speaker honoraria from CSL Behring, and Cynapsus Therapeutics. He has received grant support from Lundbeck.

FUNDING STATEMENT

Nothing to declare.

CONTRIBUTORSHIP STATEMENT

Dr. Vizcarra: conception, organization and execution of research project; design and execution of statistical analysis; writing of the first draft of manuscript.

Dr. Situ-Kcomt: execution of research project; review and critique of the manuscript.

Dr. Artusi: execution of research project; review and critique of the manuscript.

Dr. Duker: review and critique of statistical analysis; review and critique of the manuscript.

Dr. Lopiano: review and critique of statistical analysis; review and critique of the manuscript.

Dr. Okun: review and critique of statistical analysis; review and critique of the manuscript.

Dr. Espay: conception of research project; review and critique of statistical analysis; review and critique of the manuscript.

Dr. Merola: conception and organization of research project; review and critique of statistical analysis; writing of the first draft and review and critique of the manuscript of manuscript.

All the co-authors listed above gave their final approval of this manuscript version.

DATA ACCESS AND RESPONSIBILITY STATEMENT

Drs. Merola and Vizcarra had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

REFERENCES

1. Rowland NC, Sammartino F, Lozano AM (2017) Advances in surgery for movement disorders. Mov Disord 32:5–10.
2. Rodriguez-Oroz MC, Obeso JA, Lang AE, Houeto JL, Pollak P, Rehncrona S et al (2005) Bilateral deep brain stimulation in Parkinson's disease: a multicentre study with 4 years follow-up. Brain 128:2240–2249.
3. Schupbach M, Gargiulo M, Welter ML, Mallet L, Behar C, Houeto JL et al (2006) Neurosurgery in Parkinson disease: A distressed mind in a repaired body? Neurology 66:1811–1816.
4. Okun MS, Wu SS, Fayad S, Ward H, Bowers D, Rosado C et al (2014) Acute and chronic mood and apathy outcomes from a randomized study of unilateral STN and GPi DBS. PLoS One 9:1–16.
5. Zibetti M, Merola A, Rizzi L, Ricchi V, Angrisano S, Azzaro C et al (2011) Beyond nine years of continuous subthalamic nucleus deep brain stimulation in Parkinson's disease. Mov Disord 26:2327–2334.
6. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 151:264–9, W64.
7. Stroup DF (2000) Meta-analysis of Observational Studies in Epidemiology: A Proposal for Reporting. JAMA 283:2008–2012.
8. Defer G-L, Widner H, Marié R-M, Rémy P, Levivier M (1999) Core assessment program for surgical interventional therapies in Parkinson's disease (CAPSIT-PD). Mov Disord 14:572–584.
9. NHLBI (2014) Quality Assessment Tool for Before-After (Pre-Post) Studies With No Control Group. https://www.nhlbi.nih.gov/health-pro/guidelines/in-

develop/cardiovascular-risk-reduction/tools/before-after. Accessed 1 July 2017.

10. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J et al (2011) Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 342:d4002.
11. The GRADE Working Group (2013) Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach.
12. Higgins JPT, Green S, eds (2011) Cochrane Handbook for Systematic Reviews of Interventions, 5.1.0, The Cochrane Collaboration. Available from http://www.cochrane-handbook.org.
13. Østergaard K, Sunde NA (2006) Evolution of Parkinson's disease during 4 years of bilateral deep brain stimulation of the subthalamic nucleus. Mov Disord 21:624–631.
14. Ząbek M, Sobstyl M, Koziara H, Kądziołka B, Mossakowski Z, Dierzęcki S (2010) Bilateral subthalamic nucleus stimulation in the treatment of advanced Parkinson's disease. Five years' personal experience. Neurol Neurochir Pol 44:3–12.
15. Visser-Vandewalle V, Van Der Linden C, Temel Y, Celik H, Ackermans L, Spincemaille G et al (2005) Long-term effects of bilateral subthalamic nucleus stimulation in advanced Parkinson disease: A four year follow-up study. Parkinsonism Relat Disord 11:157–165.
16. Merola A, Romagnolo A, Bernardini A, Rizzi L, Artusi CA, Lanotte M et al (2015) Earlier versus later subthalamic deep brain stimulation in Parkinson's disease. Parkinsonism Relat Disord 21:972–975.
17. Merola A, Zibetti M, Angrisano S, Rizzi L, Ricchi V, Artusi CA et al (2011) Parkinson's disease progression at 30 years: A study of subthalamic deep brain-stimulated patients. Brain 134:2074–2084.
18. Moro E, Lozano AM, Pollak P, Agid Y, Rehncrona S, Volkmann J et al (2010) Long-

term results of a multicenter study on subthalamic and pallidal stimulation in Parkinson's disease. Mov Disord 25:578–586.

19. Castrioto A, Lozano A, Poon Y-Y, Lang AE, Fallis M, Moro E (2011) Ten-Year Outcome of Subthalamic Stimulation in Parkinson Disease. Arch Neurol 68:1550–1556.

20. Piboolnurak P, Lang AE, Lozano AM, Miyasaki JM, Saint-Cyr JA, Poon YYW et al (2007) Levodopa response in long-term bilateral subthalamic stimulation for Parkinson's disease. Mov Disord 22:990–997.

21. Schupbach WMM, Chastan N, Welter ML, Houeto JL, Mesnage V, Bonnet AM et al (2005) Stimulation of the subthalamic nucleus in Parkinson's disease: a 5 year follow up. J Neurol Neurosurg Psychiatry 76:1640–1644.

22. Simonin C, Tir M, Devos D, Kreisler A, Dujardin K, Salleron J et al (2009) Reduced levodopa-induced complications after 5 years of subthalamic stimulation in Parkinson's disease: A second honeymoon. J Neurol 256:1736–1741.

23. Horváth K, Aschermann Z, Ács P, Deli G, Janszky J, Komoly S et al (2015) Minimal clinically important difference on the Motor Examination part of MDS-UPDRS. Parkinsonism Relat. Disord. 21 1421–1426.

24. Fasano A, Romito LM, Daniele A, Piano C, Zinno M, Bentivoglio AR et al (2010) Motor and cognitive outcome in patients with Parkinson's disease 8 years after subthalamic implants. Brain 133:2664–2676.

25. Rizzone MG, Fasano A, Daniele A, Zibetti M, Merola A, Rizzi L et al (2014) Long-term outcome of subthalamic nucleus DBS in Parkinson's disease: From the advanced phase towards the late stage of the disease? Parkinsonism Relat Disord 20:376–381.

26. Thobois S, Lhommée E, Klinger H, Ardouin C, Schmitt E, Bichon A et al (2013) Parkinsonian apathy responds to dopaminergic stimulation of D2/D3 receptors with

piribedil. Brain 136:1568–1577.

27. Stemper B, Beric A, Welsch G, Haendl T, Sterio D, Hilz MJ (2006) Deep brain stimulation improves orthostatic regulation of patients with Parkinson disease. Neurology 67:1781–1785.

28. Joyce J, Ryoo H, Beach T, Caviness J, Stacy M, Gurevich E et al (2002) Loss of response to levodopa in Parkinson's disease and co-occurrence with dementia: role of D3 and not D2 receptors. Brain Res 955:138–152.

29. Porritt MJ, Kingsbury AE, Hughes AJ, Howells DW (2006) Striatal dopaminergic neurons are lost with Parkinson's disease progression. Mov Disord 21:2208–2211.

30. Fabbri M, Coelho M, Guedes LC, Chendo I, Sousa C, Rosa MM et al (2017) Response of non-motor symptoms to levodopa in late-stage Parkinson's disease: Results of a levodopa challenge test. Parkinsonism Relat Disord 39:37–43.

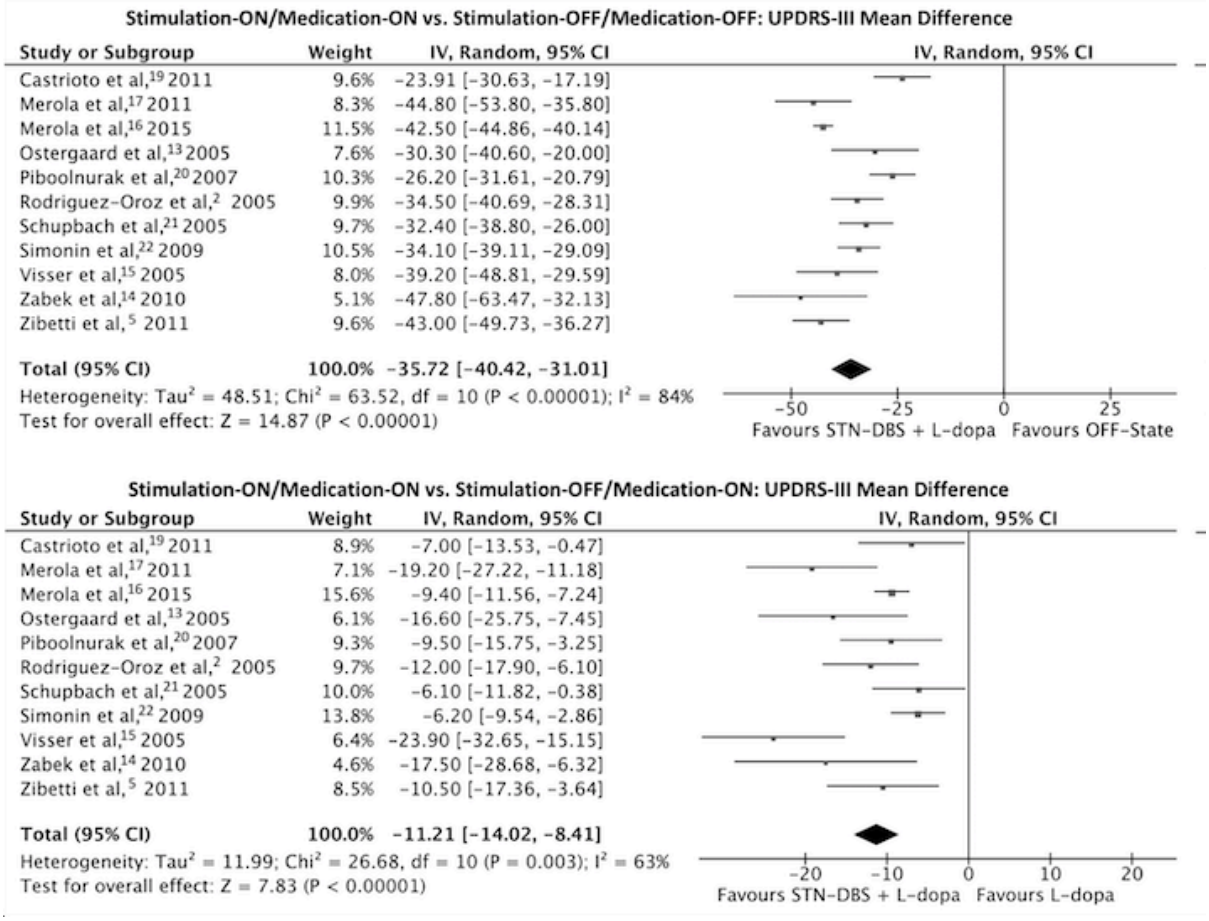
FIGURES CAPTIONS

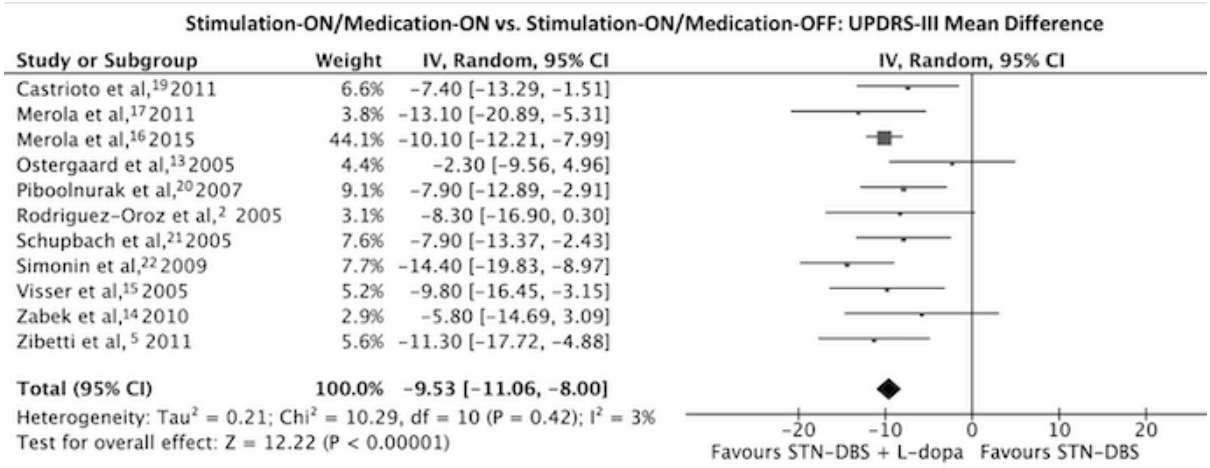
Fig. 1 STN-DBS and L-dopa effect on UPDRS-III at < 5 years. Inverse variance method was used to calculate mean differences and data were pooled using a random-effects model. Results are shown as point estimates and 95% confidence interval.
STN-DBS + L-dopa: Stimulation-ON/Medication-ON; STN-DBS: Stimulation-ON/Medication-OFF; L-dopa: Stimulation-OFF/Medication-ON; OFF state: Stimulation-OFF/Medication-OFF. UPDRS-III: Unified Parkinson's Disease Rating Scale part III. IV: inverse variance; C.I.: confidence interval

Fig. 2 Dyskinesia duration and ADL in higher vs. lower LEDD reduction subgroups. Inverse variance method was used to calculate mean differences and data were pooled using a random-effects model. Results are shown as point estimates and 95% confidence interval.
STN-DBS + L-dopa: Stimulation-ON/Medication-ON; STN-DBS: Stimulation-ON/Medication-OFF; L-dopa: Stimulation-OFF/Medication-ON; OFF state: Stimulation-OFF/Medication-OFF. UPDRS-II: Unified Parkinson's Disease Rating Scale part II; UPDRS-IV #32: Unified Parkinson's Disease Rating Scale –Item 32. I.V.: inverse variance; C.I.: confidence interval

Fig. 3 STN-DBS and L-dopa effect on UPDRS-III at ≥ 5 years**.** Inverse variance method was used to calculate mean differences and data were pooled using a random-effects model. Results are shown as point estimates and 95% confidence interval.
STN-DBS + L-dopa: Stimulation-ON/Medication-ON; STN-DBS: Stimulation-ON/Medication-OFF; L-dopa: Stimulation-OFF/Medication-ON; OFF state: Stimulation-OFF/Medication-OFF. UPDRS-III: Unified Parkinson's Disease Rating Scale part III. IV: inverse variance; C.I.: confidence interval

Figure 1.


Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-OFF: UPDRS-III Mean Difference
Study or Subgroup
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Castrioto et al,[19] 2011 9.6% -23.91 [-30.63, -17.19]
Merola et al,[17] 2011 8.3% -44.80 [-53.80, -35.80]
Merola et al,[16] 2015 11.5% -42.50 [-44.86, -40.14]
Ostergaard et al,[13] 2005 7.6% -30.30 [-40.60, -20.00]
Piboolnurak et al,[20] 2007 10.3% -26.20 [-31.61, -20.79]
Rodriguez-Oroz et al,[2] 2005 9.9% -34.50 [-40.69, -28.31]
Schupbach et al,[21] 2005 9.7% -32.40 [-38.80, -26.00]
Simonin et al,[22] 2009 10.5% -34.10 [-39.11, -29.09]
Visser et al,[15] 2005 8.0% -39.20 [-48.81, -29.59]
Zabek et al,[14] 2010 5.1% -47.80 [-63.47, -32.13]
Zibetti et al,[5] 2011 9.6% -43.00 [-49.73, -36.27]
Total (95% CI) 100.0% -35.72 [-40.42, -31.01]
Heterogeneity: Tau² = 48.51; Chi² = 63.52, df = 10 (P < 0.00001); I² = 84%
Test for overall effect: Z = 14.87 (P < 0.00001)
-50
-25
0
25
Favours STN-DBS + L-dopa
Favours OFF-State
Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference
Study or Subgroup
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Castrioto et al,[19] 2011 8.9% -7.00 [-13.53, -0.47]
Merola et al,[17] 2011 7.1% -19.20 [-27.22, -11.18]
Merola et al,[16] 2015 15.6% -9.40 [-11.56, -7.24]
Ostergaard et al,[13] 2005 6.1% -16.60 [-25.75, -7.45]
Piboolnurak et al,[20] 2007 9.3% -9.50 [-15.75, -3.25]
Rodriguez-Oroz et al,[2] 2005 9.7% -12.00 [-17.90, -6.10]
Schupbach et al,[21] 2005 10.0% -6.10 [-11.82, -0.38]
Simonin et al,[22] 2009 13.8% -6.20 [-9.54, -2.86]
Visser et al,[15] 2005 6.4% -23.90 [-32.65, -15.15]
Zabek et al,[14] 2010 4.6% -17.50 [-28.68, -6.32]
Zibetti et al,[5] 2011 8.5% -10.50 [-17.36, -3.64]
Total (95% CI) 100.0% -11.21 [-14.02, -8.41]
Heterogeneity: Tau² = 11.99; Chi² = 26.68, df = 10 (P = 0.003); I² = 63%
Test for overall effect: Z = 7.83 (P < 0.00001)
-20
-10
0
10
20
Favours STN-DBS + L-dopa
Favours L-dopa


Stimulation-ON/Medication-ON vs. Stimulation-ON/Medication-OFF: UPDRS-III Mean Difference
Study or Subgroup
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Castrioto et al,[19] 2011 6.6% -7.40 [-13.29, -1.51]
Merola et al,[17] 2011 3.8% -13.10 [-20.89, -5.31]
Merola et al,[16] 2015 44.1% -10.10 [-12.21, -7.99]
Ostergaard et al,[13] 2005 4.4% -2.30 [-9.56, 4.96]
Piboolnurak et al,[20] 2007 9.1% -7.90 [-12.89, -2.91]
Rodriguez-Oroz et al,[2] 2005 3.1% -8.30 [-16.90, 0.30]
Schupbach et al,[21] 2005 7.6% -7.90 [-13.37, -2.43]
Simonin et al,[22] 2009 7.7% -14.40 [-19.83, -8.97]
Visser et al,[15] 2005 5.2% -9.80 [-16.45, -3.15]
Zabek et al,[14] 2010 2.9% -5.80 [-14.69, 3.09]
Zibetti et al,[5] 2011 5.6% -11.30 [-17.72, -4.88]
Total (95% CI) 100.0% -9.53 [-11.06, -8.00]
Heterogeneity: Tau² = 0.21; Chi² = 10.29, df = 10 (P = 0.42); I² = 3%
Test for overall effect: Z = 12.22 (P < 0.00001)
-20 -10 0 10 20
Favours STN-DBS + L-dopa
Favours STN-DBS

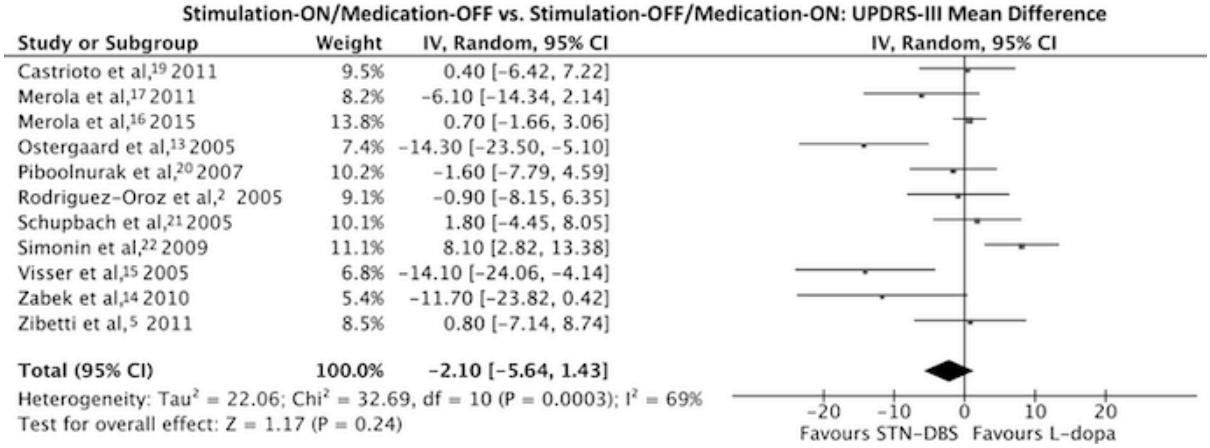
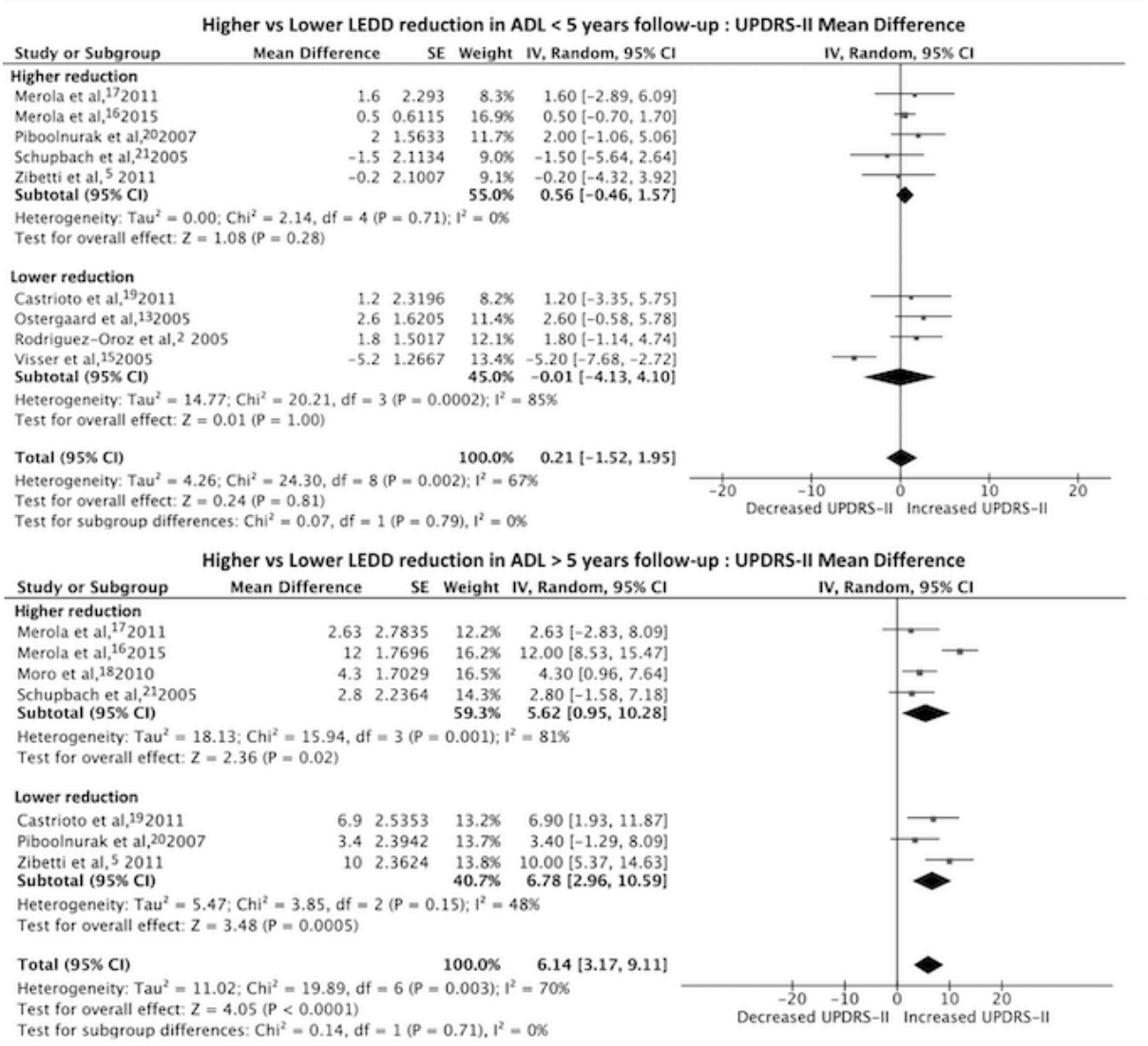
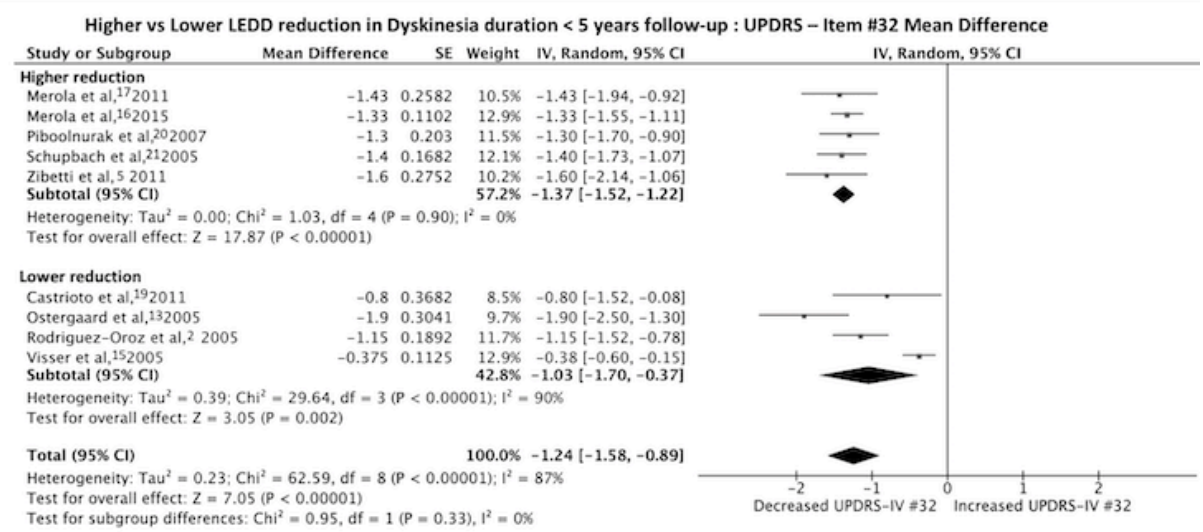

Stimulation-ON/Medication-OFF vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference
Study or Subgroup
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Castrioto et al,[19] 2011 9.5% 0.40 [-6.42, 7.22]
Merola et al,[17] 2011 8.2% -6.10 [-14.34, 2.14]
Merola et al,[16] 2015 13.8% 0.70 [-1.66, 3.06]
Ostergaard et al,[13] 2005 7.4% -14.30 [-23.50, -5.10]
Piboolnurak et al,[20] 2007 10.2% -1.60 [-7.79, 4.59]
Rodriguez-Oroz et al,[2] 2005 9.1% -0.90 [-8.15, 6.35]
Schupbach et al,[21] 2005 10.1% 1.80 [-4.45, 8.05]
Simonin et al,[22] 2009 11.1% 8.10 [2.82, 13.38]
Visser et al,[15] 2005 6.8% -14.10 [-24.06, -4.14]
Zabek et al,[14] 2010 5.4% -11.70 [-23.82, 0.42]
Zibetti et al,[5] 2011 8.5% 0.80 [-7.14, 8.74]
Total (95% CI) 100.0% -2.10 [-5.64, 1.43]
Heterogeneity: Tau² = 22.06; Chi² = 32.69, df = 10 (P = 0.0003); I² = 69%
Test for overall effect: Z = 1.17 (P = 0.24)
-20 -10 0 10 20
Favours STN-DBS
Favours L-dopa

Figure 2


Higher vs Lower LEDD reduction in ADL < 5 years follow-up : UPDRS-II Mean Difference
Study or Subgroup | Mean Difference | SE | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Higher reduction
Merola et al,[17]2011 1.6 2.293 8.3% 1.60 [-2.89, 6.09]
Merola et al,[16]2015 0.5 0.6115 16.9% 0.50 [-0.70, 1.70]
Piboolnurak et al,[20]2007 2 1.5633 11.7% 2.00 [-1.06, 5.06]
Schupbach et al,[21]2005 -1.5 2.1134 9.0% -1.50 [-5.64, 2.64]
Zibetti et al,[5] 2011 -0.2 2.1007 9.1% -0.20 [-4.32, 3.92]
Subtotal (95% CI) 55.0% 0.56 [-0.46, 1.57]
Heterogeneity: Tau² = 0.00; Chi² = 2.14, df = 4 (P = 0.71); I² = 0%
Test for overall effect: Z = 1.08 (P = 0.28)
Lower reduction
Castrioto et al,[19]2011 1.2 2.3196 8.2% 1.20 [-3.35, 5.75]
Ostergaard et al,[13]2005 2.6 1.6205 11.4% 2.60 [-0.58, 5.78]
Rodriguez-Oroz et al,[2] 2005 1.8 1.5017 12.1% 1.80 [-1.14, 4.74]
Visser et al,[15]2005 -5.2 1.2667 13.4% -5.20 [-7.68, -2.72]
Subtotal (95% CI) 45.0% -0.01 [-4.13, 4.10]
Heterogeneity: Tau² = 14.77; Chi² = 20.21, df = 3 (P = 0.0002); I² = 85%
Test for overall effect: Z = 0.01 (P = 1.00)
Total (95% CI) 100.0% 0.21 [-1.52, 1.95]
Heterogeneity: Tau² = 4.26; Chi² = 24.30, df = 8 (P = 0.002); I² = 67%
Test for overall effect: Z = 0.24 (P = 0.81)
Test for subgroup differences: Chi² = 0.07, df = 1 (P = 0.79), I² = 0%
-20 -10 0 10 20
Decreased UPDRS-II Increased UPDRS-II
Higher vs Lower LEDD reduction in ADL > 5 years follow-up : UPDRS-II Mean Difference
Study or Subgroup | Mean Difference | SE | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Higher reduction
Merola et al,[17]2011 2.63 2.7835 12.2% 2.63 [-2.83, 8.09]
Merola et al,[16]2015 12 1.7696 16.2% 12.00 [8.53, 15.47]
Moro et al,[18]2010 4.3 1.7029 16.5% 4.30 [0.96, 7.64]
Schupbach et al,[21]2005 2.8 2.2364 14.3% 2.80 [-1.58, 7.18]
Subtotal (95% CI) 59.3% 5.62 [0.95, 10.28]
Heterogeneity: Tau² = 18.13; Chi² = 15.94, df = 3 (P = 0.001); I² = 81%
Test for overall effect: Z = 2.36 (P = 0.02)
Lower reduction
Castrioto et al,[19]2011 6.9 2.5353 13.2% 6.90 [1.93, 11.87]
Piboolnurak et al,[20]2007 3.4 2.3942 13.7% 3.40 [-1.29, 8.09]
Zibetti et al,[5] 2011 10 2.3624 13.8% 10.00 [5.37, 14.63]
Subtotal (95% CI) 40.7% 6.78 [2.96, 10.59]
Heterogeneity: Tau² = 5.47; Chi² = 3.85, df = 2 (P = 0.15); I² = 48%
Test for overall effect: Z = 3.48 (P = 0.0005)
Total (95% CI) 100.0% 6.14 [3.17, 9.11]
Heterogeneity: Tau² = 11.02; Chi² = 19.89, df = 6 (P = 0.003); I² = 70%
Test for overall effect: Z = 4.05 (P < 0.0001)
Test for subgroup differences: Chi² = 0.14, df = 1 (P = 0.71), I² = 0%
-20 -10 0 10 20
Decreased UPDRS-II Increased UPDRS-II


Higher vs Lower LEDD reduction in Dyskinesia duration < 5 years follow-up : UPDRS – Item #32 Mean Difference
Study or Subgroup
Mean Difference
SE
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Higher reduction
Merola et al,[17]2011 −1.43 0.2582 10.5% −1.43 [−1.94, −0.92]
Merola et al,[16]2015 −1.33 0.1102 12.9% −1.33 [−1.55, −1.11]
Piboolnurak et al,[20]2007 −1.3 0.203 11.5% −1.30 [−1.70, −0.90]
Schupbach et al,[21]2005 −1.4 0.1682 12.1% −1.40 [−1.73, −1.07]
Zibetti et al,[5] 2011 −1.6 0.2752 10.2% −1.60 [−2.14, −1.06]
Subtotal (95% CI) 57.2% −1.37 [−1.52, −1.22]
Heterogeneity: Tau² = 0.00; Chi² = 1.03, df = 4 (P = 0.90); I² = 0%
Test for overall effect: Z = 17.87 (P < 0.00001)
Lower reduction
Castrioto et al,[19]2011 −0.8 0.3682 8.5% −0.80 [−1.52, −0.08]
Ostergaard et al,[13]2005 −1.9 0.3041 9.7% −1.90 [−2.50, −1.30]
Rodriguez-Oroz et al,[2] 2005 −1.15 0.1892 11.7% −1.15 [−1.52, −0.78]
Visser et al,[15]2005 −0.375 0.1125 12.9% −0.38 [−0.60, −0.15]
Subtotal (95% CI) 42.8% −1.03 [−1.70, −0.37]
Heterogeneity: Tau² = 0.39; Chi² = 29.64, df = 3 (P < 0.00001); I² = 90%
Test for overall effect: Z = 3.05 (P = 0.002)
Total (95% CI) 100.0% −1.24 [−1.58, −0.89]
Heterogeneity: Tau² = 0.23; Chi² = 62.59, df = 8 (P < 0.00001); I² = 87%
Test for overall effect: Z = 7.05 (P < 0.00001)
Test for subgroup differences: Chi² = 0.95, df = 1 (P = 0.33), I² = 0%
−2
−1
0
1
2
Decreased UPDRS-IV #32
Increased UPDRS-IV #32

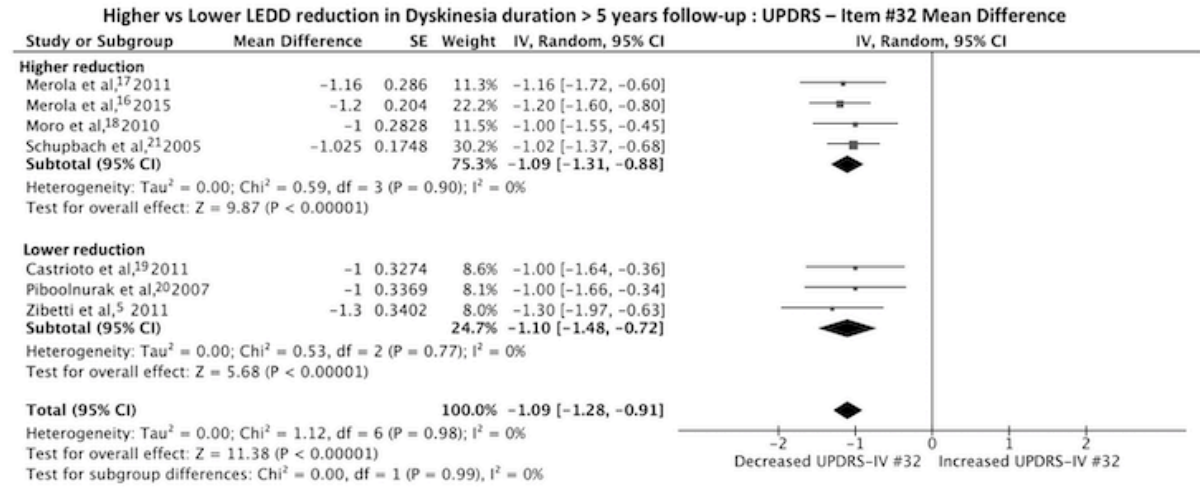
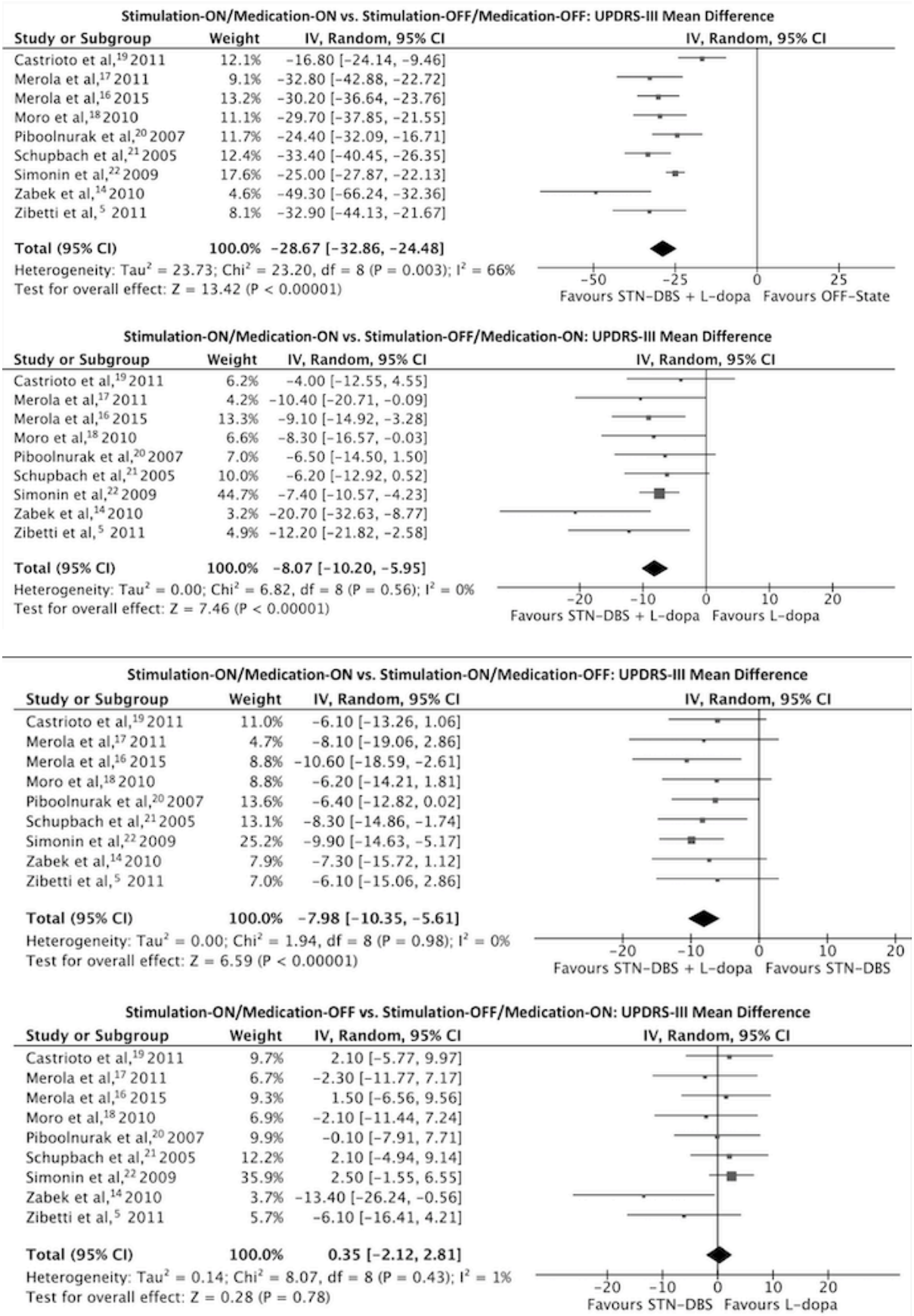

Higher vs Lower LEDD reduction in Dyskinesia duration > 5 years follow-up : UPDRS – Item #32 Mean Difference
Study or Subgroup
Mean Difference
SE
Weight
IV, Random, 95% CI
IV, Random, 95% CI
Higher reduction
Merola et al,[17]2011 −1.16 0.286 11.3% −1.16 [−1.72, −0.60]
Merola et al,[16]2015 −1.2 0.204 22.2% −1.20 [−1.60, −0.80]
Moro et al,[18]2010 −1 0.2828 11.5% −1.00 [−1.55, −0.45]
Schupbach et al,[21]2005 −1.025 0.1748 30.2% −1.02 [−1.37, −0.68]
Subtotal (95% CI) 75.3% −1.09 [−1.31, −0.88]
Heterogeneity: Tau² = 0.00; Chi² = 0.59, df = 3 (P = 0.90); I² = 0%
Test for overall effect: Z = 9.87 (P < 0.00001)
Lower reduction
Castrioto et al,[19]2011 −1 0.3274 8.6% −1.00 [−1.64, −0.36]
Piboolnurak et al,[20]2007 −1 0.3369 8.1% −1.00 [−1.66, −0.34]
Zibetti et al,[5] 2011 −1.3 0.3402 8.0% −1.30 [−1.97, −0.63]
Subtotal (95% CI) 24.7% −1.10 [−1.48, −0.72]
Heterogeneity: Tau² = 0.00; Chi² = 0.53, df = 2 (P = 0.77); I² = 0%
Test for overall effect: Z = 5.68 (P < 0.00001)
Total (95% CI) 100.0% −1.09 [−1.28, −0.91]
Heterogeneity: Tau² = 0.00; Chi² = 1.12, df = 6 (P = 0.98); I² = 0%
Test for overall effect: Z = 11.38 (P < 0.00001)
Test for subgroup differences: Chi² = 0.00, df = 1 (P = 0.99), I² = 0%
−2
−1
0
1
2
Decreased UPDRS-IV #32
Increased UPDRS-IV #32

Figure 3


Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-OFF: UPDRS-III Mean Difference
Study or Subgroup | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Castrioto et al,[19] 2011 12.1% -16.80 [-24.14, -9.46]
Merola et al,[17] 2011 9.1% -32.80 [-42.88, -22.72]
Merola et al,[16] 2015 13.2% -30.20 [-36.64, -23.76]
Moro et al,[18] 2010 11.1% -29.70 [-37.85, -21.55]
Piboolnurak et al,[20] 2007 11.7% -24.40 [-32.09, -16.71]
Schupbach et al,[21] 2005 12.4% -33.40 [-40.45, -26.35]
Simonin et al,[22] 2009 17.6% -25.00 [-27.87, -22.13]
Zabek et al,[14] 2010 4.6% -49.30 [-66.24, -32.36]
Zibetti et al,[5] 2011 8.1% -32.90 [-44.13, -21.67]
Total (95% CI) 100.0% -28.67 [-32.86, -24.48]
Heterogeneity: Tau² = 23.73; Chi² = 23.20, df = 8 (P = 0.003); I² = 66%
Test for overall effect: Z = 13.42 (P < 0.00001)
-50 -25 0 25
Favours STN-DBS + L-dopa Favours OFF-State
Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference
Study or Subgroup | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Castrioto et al,[19] 2011 6.2% -4.00 [-12.55, 4.55]
Merola et al,[17] 2011 4.2% -10.40 [-20.71, -0.09]
Merola et al,[16] 2015 13.3% -9.10 [-14.92, -3.28]
Moro et al,[18] 2010 6.6% -8.30 [-16.57, -0.03]
Piboolnurak et al,[20] 2007 7.0% -6.50 [-14.50, 1.50]
Schupbach et al,[21] 2005 10.0% -6.20 [-12.92, 0.52]
Simonin et al,[22] 2009 44.7% -7.40 [-10.57, -4.23]
Zabek et al,[14] 2010 3.2% -20.70 [-32.63, -8.77]
Zibetti et al,[5] 2011 4.9% -12.20 [-21.82, -2.58]
Total (95% CI) 100.0% -8.07 [-10.20, -5.95]
Heterogeneity: Tau² = 0.00; Chi² = 6.82, df = 8 (P = 0.56); I² = 0%
Test for overall effect: Z = 7.46 (P < 0.00001)
-20 -10 0 10 20
Favours STN-DBS + L-dopa Favours L-dopa
Stimulation-ON/Medication-ON vs. Stimulation-ON/Medication-OFF: UPDRS-III Mean Difference
Study or Subgroup | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Castrioto et al,[19] 2011 11.0% -6.10 [-13.26, 1.06]
Merola et al,[17] 2011 4.7% -8.10 [-19.06, 2.86]
Merola et al,[16] 2015 8.8% -10.60 [-18.59, -2.61]
Moro et al,[18] 2010 8.8% -6.20 [-14.21, 1.81]
Piboolnurak et al,[20] 2007 13.6% -6.40 [-12.82, 0.02]
Schupbach et al,[21] 2005 13.1% -8.30 [-14.86, -1.74]
Simonin et al,[22] 2009 25.2% -9.90 [-14.63, -5.17]
Zabek et al,[14] 2010 7.9% -7.30 [-15.72, 1.12]
Zibetti et al,[5] 2011 7.0% -6.10 [-15.06, 2.86]
Total (95% CI) 100.0% -7.98 [-10.35, -5.61]
Heterogeneity: Tau² = 0.00; Chi² = 1.94, df = 8 (P = 0.98); I² = 0%
Test for overall effect: Z = 6.59 (P < 0.00001)
-20 -10 0 10 20
Favours STN-DBS + L-dopa Favours STN-DBS
Stimulation-ON/Medication-OFF vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference
Study or Subgroup | Weight | IV, Random, 95% CI | IV, Random, 95% CI
Castrioto et al,[19] 2011 9.7% 2.10 [-5.77, 9.97]
Merola et al,[17] 2011 6.7% -2.30 [-11.77, 7.17]
Merola et al,[16] 2015 9.3% 1.50 [-6.56, 9.56]
Moro et al,[18] 2010 6.9% -2.10 [-11.44, 7.24]
Piboolnurak et al,[20] 2007 9.9% -0.10 [-7.91, 7.71]
Schupbach et al,[21] 2005 12.2% 2.10 [-4.94, 9.14]
Simonin et al,[22] 2009 35.9% 2.50 [-1.55, 6.55]
Zabek et al,[14] 2010 3.7% -13.40 [-26.24, -0.56]
Zibetti et al,[5] 2011 5.7% -6.10 [-16.41, 4.21]
Total (95% CI) 100.0% 0.35 [-2.12, 2.81]
Heterogeneity: Tau² = 0.14; Chi² = 8.07, df = 8 (P = 0.43); I² = 1%
Test for overall effect: Z = 0.28 (P = 0.78)
-20 -10 0 10 20
Favours STN-DBS Favours L-dopa

Supplementary Material

Dopaminergic Dose Reduction after Subthalamic Deep Brain Stimulation: Time to Revisit a Common Practice?

Joaquin A. Vizcarra, MD[1]; Miguel Situ-Kcomt, MD[2]; Carlo Alberto Artusi, MD[3]; Andrew P Duker, MD[1]; Leonardo Lopiano MD, PhD[3]; Michael S. Okun, MD[4]; Alberto J. Espay, MD, MSc[1]; Aristide Merola, MD, PhD[1]

[1] Gardner Family Center for Parkinson's Disease and Movement Disorders, Department of Neurology, University of Cincinnati, Cincinnati, OH, USA

[2] Department of Neurology, University of Cincinnati, Cincinnati, OH, USA

[3] Department of Neuroscience "Rita Levi Montalcini", University of Turin, via Cherasco 15, 10126, Torino, Italy

[4] Department of Neurology, Center for Movement Disorders and Neurorestoration, McKnight Brain Institute, Gainesville FL, USA

Corresponding Author: Aristide Merola MD, PhD

260 Stetson St, Suite 2300, Cincinnati, OH 45267-0525, USA.

e-mail: merolaae@ucmail.uc.edu

Supplementary Files

Supplementary Figure 1: Study selection algorithm

Supplementary Table 1: ON/OFF evaluation of included studies

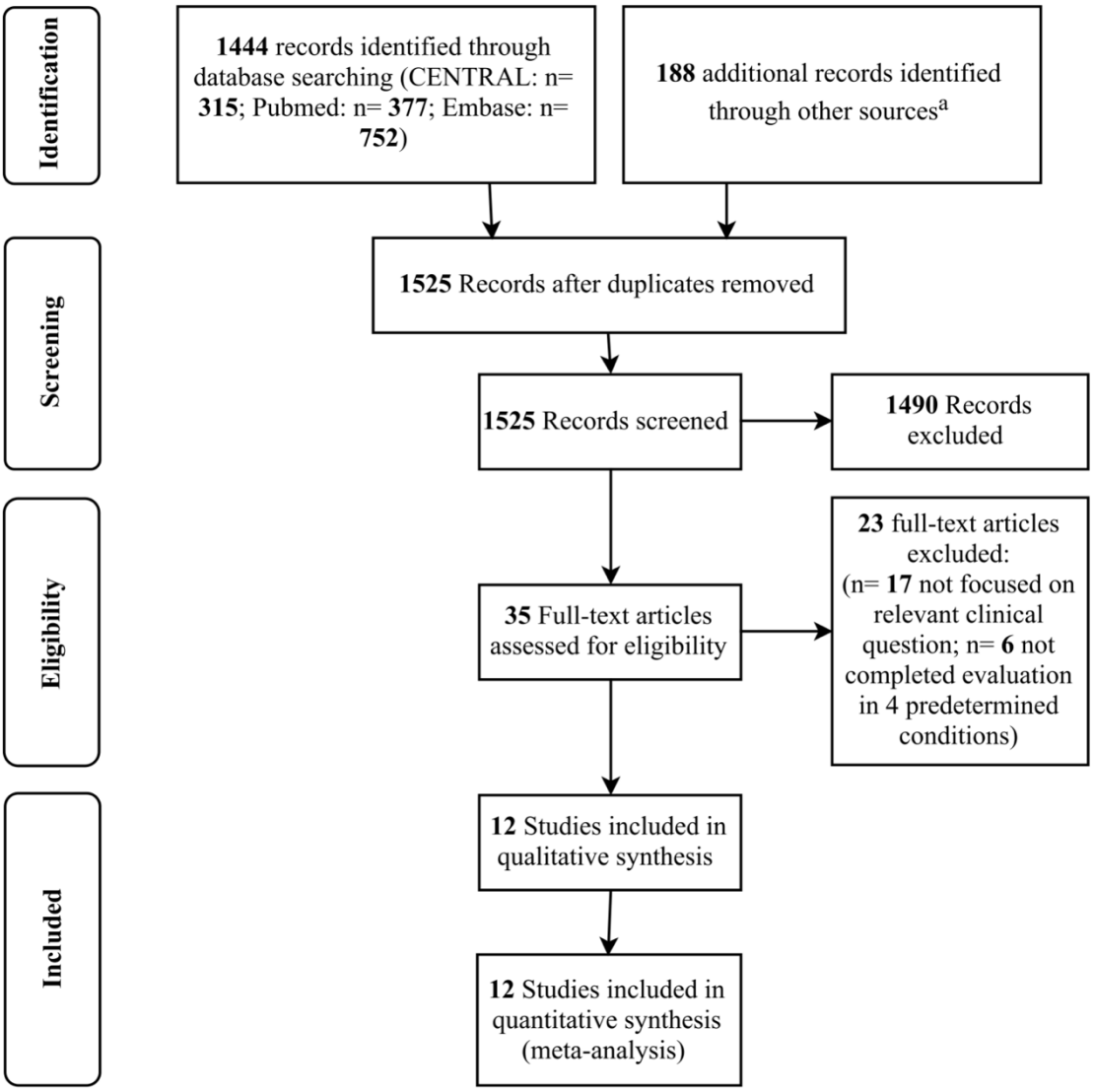
Supplementary Table 2: Database search methods

Supplementary Table 3: Individual quality appraisal

Supplementary Table 4: Grading of Recommendations Assessment, Development and Evaluation (GRADE) profile

Supplementary Figure 2: Funnel plots

Supplementary Figure 1: Study selection algorithm



CENTRAL: Central Register of Controlled Trials; [a] clinicaltrials.gov and System for Information on Grey Literature in Europe.

Supplementary Table 1: ON/OFF evaluation of included studies

Author	Pre-op ON/OFF Evaluation	Post-op ON/OFF Evaluation
Ostergaard et al.[13]	OFF: 10-12 hours off meds. ON: 1 hour after usual levodopa dose	Stim-OFF/Med-OFF: 10-12 hours both off. Stim-ON/Med-OFF: 30 min after turning on stim. Stim-OFF/Med-ON: 60 min later. Stim-ON/Med-ON:30 min after turning on stim.
Zabek et al.[14]	OFF: 12 hours off meds. ON: 1 hour after 200mg levodopa/50mg benzeraside	Stim-OFF/Med-OFF: 4 hours stim. off. Stim-ON/Med-OFF: NR. Stim-OFF/Med-ON: NR. Stim-ON/Med-ON: NR.
Visser-Vandewalle et al.[15]	best ON and practically defined OFF state as described in the CAPIT-PD	Stim-OFF/Med-OFF: practically defined. off. Stim-ON/Med-OFF: NR. Stim-OFF/Med-ON: best on. Stim-ON/Med-ON: practically defined.
Merola et al. 2015[16]	OFF: 12 hours off meds. ON: 40 min after 150% of usual levodopa dose	Stim-OFF/Med-OFF: CAPSIT-PD. Stim-ON/Med-OFF: CAPSIT-PD. Stim-OFF/Med-ON: CAPSIT-PD. Stim-ON/Med-ON: CAPSIT-PD.
Rodríguez-Oroz et al.[2]	OFF: 12 hours off meds. ON: 1 hour after 150% of usual levodopa dose	Stim-OFF/Med-OFF: 1-2 hours off stim. Stim-ON/Med-OFF: 30 min after turning stim on. Stim-OFF/Med-ON: 1-2 hours after turning stim off. Stim-ON/Med-ON: 30 min after turning stim on.
Moro et al.[18]	OFF: 12 hours off meds. ON: best response to morning dose	Stim-OFF/Med-OFF: 1-2 hours off stim. Stim-ON/Med-OFF: 30 min after turning stim on. Stim-OFF/Med-ON: 1-2 hours after turning stim off. Stim-ON/Med-ON: 30 min after turning stim on.
Castrioto et al.[19]	OFF: 12 hours off meds. ON: 1 hour after 150% of usual levodopa dose	Stim-OFF/Med-OFF: 1 hour off stim. Stim-ON/Med-OFF: 60 min after turning stim on. Stim-OFF/Med-ON: 60 min after medication on. Stim-ON/Med-ON: 60 min after turning on stim.
Piboolnurak et al.[20]	OFF: CAPIT-PD conditions. ON: CAPIT-PD conditions	Stim-OFF/Med-OFF: NR. Stim-ON/Med-OFF: 30 min after turning stim on. Stim-OFF/Med-ON: NR. Stim-ON/Med-ON: 30 min after turning on stim.
Schupbach et al.[21]	OFF: 12 hours off meds. ON: 1 hour after 150% of usual levodopa dose	Stim-OFF/Med-OFF: 10-12 hours off (27 pax) or for at least 1.5hours (10 pax). Stim-ON/Med-OFF: 30 min after turning stim on. Stim-OFF/Med-ON: 60 min after turning stim off. Stim-ON/Med-ON: 30 min after turning on stim.
Zibetti et al.[5]	OFF: 12 hours off meds. ON: 1 hour after 150% of usual levodopa dose	Stim-ON/Med-OFF: 10-12 hours off meds. Stim-OFF/Med-OFF: 60 min. after turning stim off. Stim-OFF/Med-ON: 40 min after med on. Stim-ON/Med-ON: NR.
Simonin et al.[22]	OFF: 12 hours off meds. ON: 1 hour after 150% of usual levodopa dose	Stim-ON/Med-OFF: 10-12 hours off meds. Stim-OFF/Med-OFF: 120 min. after turning stim off. Stim-OFF/Med-ON: after pre-surgical dose of med on. Stim-ON/Med-ON: 30 min after turning stim on.
Merola et al. 2011[17]	OFF: 12 hours off meds. ON: 40 min after 150% of usual levodopa dose	Stim-OFF/Med-OFF: CAPSIT-PD. Stim-ON/Med-OFF: CAPSIT-PD. Stim-OFF/Med-ON: CAPSIT-PD. Stim-ON/Med-ON: CAPSIT-PD.

NR: Non-applicable. Pax: patients. Stim: stimulation. Med: medication. Min: minutes

Supplementary Table 2: Database search methods

Pubmed and CENTRAL		Embase
Search number	**Keyword**	**Search term**
#1	Parkinson's disease	('parkinson disease' OR parkinson) AND ('brain depth stimulation' OR dbs) AND ('follow up'/exp OR 'follow up')
#2	Parkinson	
#3	Deep brain stimulation	
#4	DBS	
#5	follow-up	
#6	#1 OR #2	
#7	#3 OR #4	
#8	#6 AND #7 AND #5	

CENTRAL indicates Central Register of Controlled Trials

Supplementary Table 3: Individual quality appraisal

	Ostergaard et al.[13]	**Zabek et al.**[14]	**Visser-Vandewalle et al.**[15]	**Merola et al. 2015**[16]	**Rodríguez-Oroz et al.**[2]	**Merola et al. 2011**[17]	**Moro et al.**[18]	**Castrioto et al.**[19]	**Piboolnurak et al.**[20]	**Schupbach et al.**[21]	**Zibetti et al.**[5]	**Simonin et al.**[22]
1. Was the study question or objective clearly stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Were eligibility selection criteria for the study population prespecified and clearly described	No	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
3. Were the participants in the study representative of those who would be eligible for the test/service/intervention in the general or clinical population of interest?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
4. Were all eligible participants that met the prespecified entry criteria enrolled?	Yes	Yes	Yes	Yes	Yes	N/A	Yes	Yes	Yes	Yes	Yes	Yes
5. Was the sample size sufficiently large to provide confidence in the findings?	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
6. Was the test/service/intervention clearly described and delivered consistently across the study population?	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
7. Were the outcome measures prespecified, clearly defined, valid, reliable, and assessed consistently across all study participants?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

8. Were the people assessing the outcomes blinded to the participants' exposures/interventions?	No	No	No	No	No	No	No	No	No	No	No	No
9. Was the loss to follow-up after baseline 20% or less? Were those lost to follow-up accounted for in the analysis?	Yes	Yes	Yes	No	No	Yes	No	Yes	No	Yes	Yes	Yes
10. Did the statistical methods examine changes in outcome measures from before to after the intervention? Were statistical tests done that provided p values for the pre-to-post changes?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
11. Were outcome measures of interest taken multiple times before the intervention and multiple times after the intervention?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
12. If the intervention was conducted at a group level, did the statistical analysis take into account the use of individual-level data to determine effects at the group level?	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Quality Rating	**Good**	**Poor**	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**	**Good**

Supplementary table 4: Grading of Recommendations Assessment, Development and Evaluation (GRADE) profile

Number of Studies	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Overall quality of evidence
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than OFF-state in < 5 years of follow-up?						
11	Not serious	Not serious[a]	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than L-dopa in < 5 years of follow-up?						
11	Not serious	Not serious[d]	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than STN-DBS in < 5 years of follow-up?						
11	Not serious	Not serious	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS improve the UPDRS-III score to a greater extent than L-dopa in < 5 years of follow-up?						
11	Not serious	Serious[e]	Not serious[b]	Not serious	Undetected	+
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than OFF-state in ≥ 5 years of follow-up?						
9	Not serious	Not serious[f]	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than L-dopa in ≥ 5 years of follow-up?						
9	Not serious	Not serious	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS + L-dopa improve the UPDRS-III score to a greater extent than STN-DBS in ≥ 5 years of follow-up?						
9	Not serious	Not serious	Not serious[b]	Not serious[c]	Undetected	++
Does STN-DBS improve the UPDRS-III score to a greater extent than L-dopa in ≥ 5 years of follow-up?						
9	Not serious	Not serious	Not serious[b]	Not serious	Undetected	++
Is there a difference between higher versus lower LEDD reduction in the UPDRS-II score in < 5 years of follow-up?						
9	Not serious	Serious[g]	Not serious[b]	Not serious	Undetected	+
Is there a difference between higher versus lower LEDD reduction in the UPDRS-IV (item 32) score in < 5 years of follow-up?						
7	Not serious	Serious[h]	Not serious[b]	Not serious	Undetected	+
Is there a difference between higher versus lower LEDD reduction in the UPDRS-II score in ≥ 5 years of follow-up?						
9	Not serious	Serious[i]	Not serious[b]	Not serious	Undetected	+
Is there a difference between higher versus lower LEDD reduction in the UPDRS-IV (item 32) score in ≥ 5 years of follow-up?						
7	Not serious	Not serious	Not serious[b]	Not serious	Undetected	++

STN-DBS + L-dopa = Stimulation-ON/Medication-ON; OFF-state = Stimulation-OFF/Medication-OFF; L-dopa= Stimulation-OFF/Medication-ON; STN-DBS= Stimulation-ON/Medication-OFF. PD: Parkinson's disease; UPDRS: Unified Parkinson's Disease Rating Scale; STN-DBS: Subthalamic Nucleus-Deep Brain Stimulation; L-dopa: levodopa.

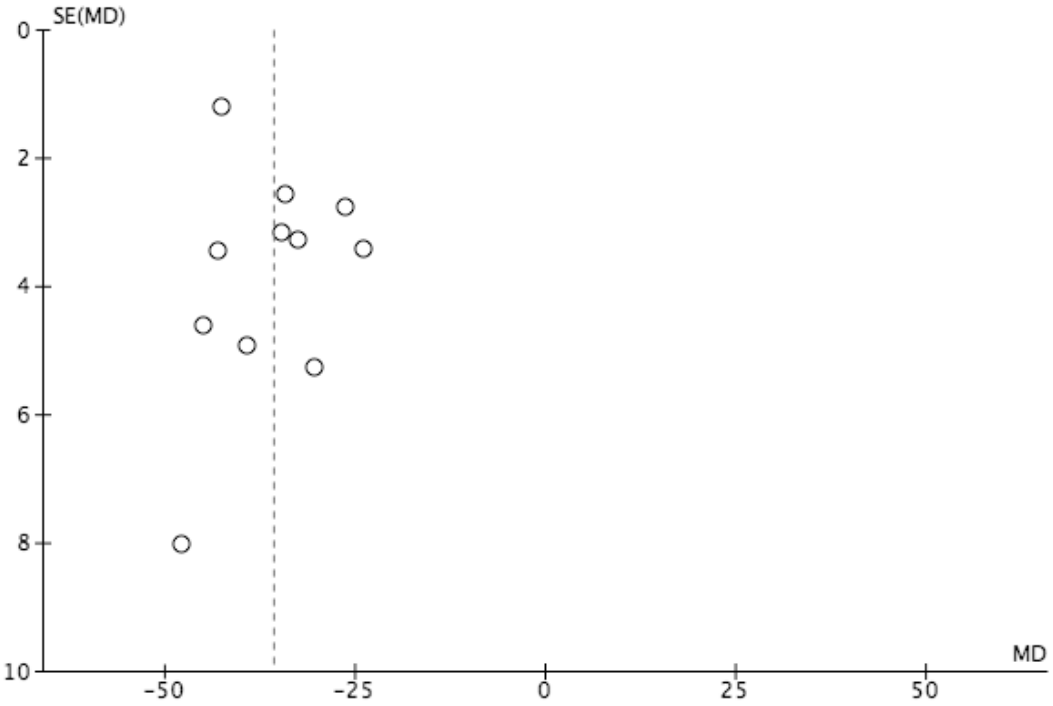
++ (Low) = our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect; + (Very low) = we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[a]: Heterogeneity of questionable importance due to unidirectional effect despite I^2 value of 84%; [b]: Studies have similar population characteristics, interventions, and controls. Same outcomes; [c]: Lower end of confidence interval exceeds minimal clinically important difference of -3.25 in scale. Optimal information size not obtained as only one group was involved with no control; [d]: Heterogeneity of questionable importance due to unidirectional effect despite I^2 value of 63%; [e]: I^2 value is 69%, point estimates vary widely across studies, almost no overlap in confidence intervals; [f]: Heterogeneity of questionable importance due to unidirectional effect despite I^2 value of 66%. [g]: Heterogeneity in subgroups of probable importance due to bidirectional effect and I^2 value of 0% and 85% in the higher and lower LEDD reduction subgroups, respectively. [h]: Heterogeneity in subgroups of probable importance due to bidirectional effect and I^2 value of 81% and 48% in the higher and lower LEDD reduction subgroups, respectively. [i]: Heterogeneity in subgroups of questionable importance due to unidirectional effect despite I^2 value of 0% and 90% in the higher and lower LEDD reduction subgroups, respectively. Subgroup analysis warrants caution providing reason for degrading our confidence in the estimate.

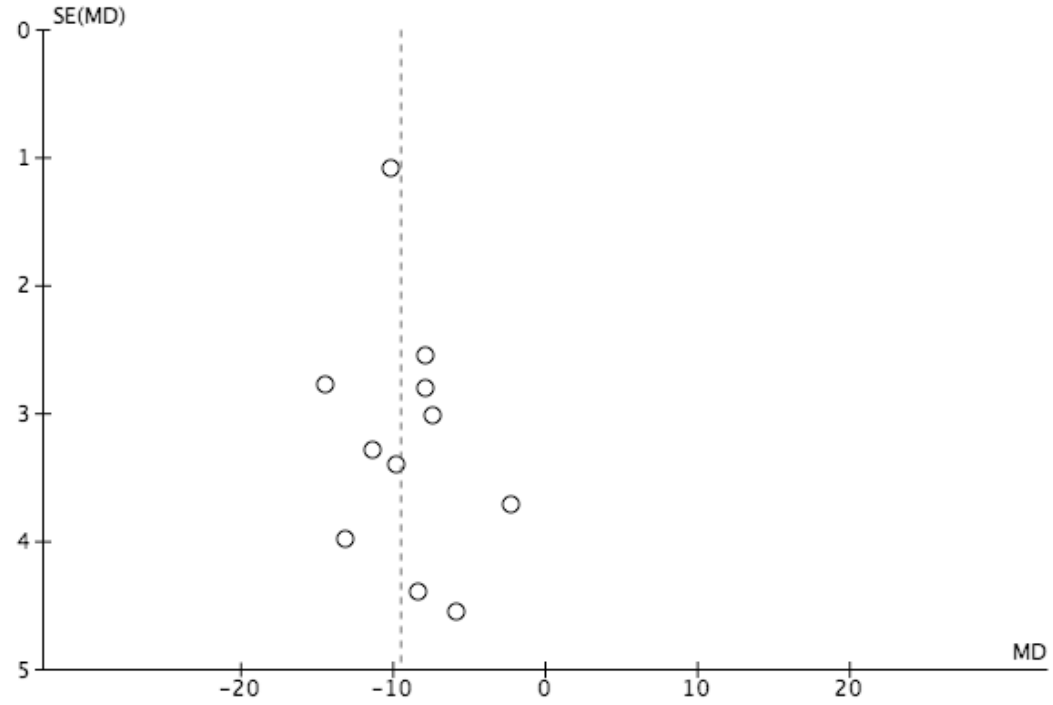
Supplementary Figure 2: Funnel plots

Less than 5 years

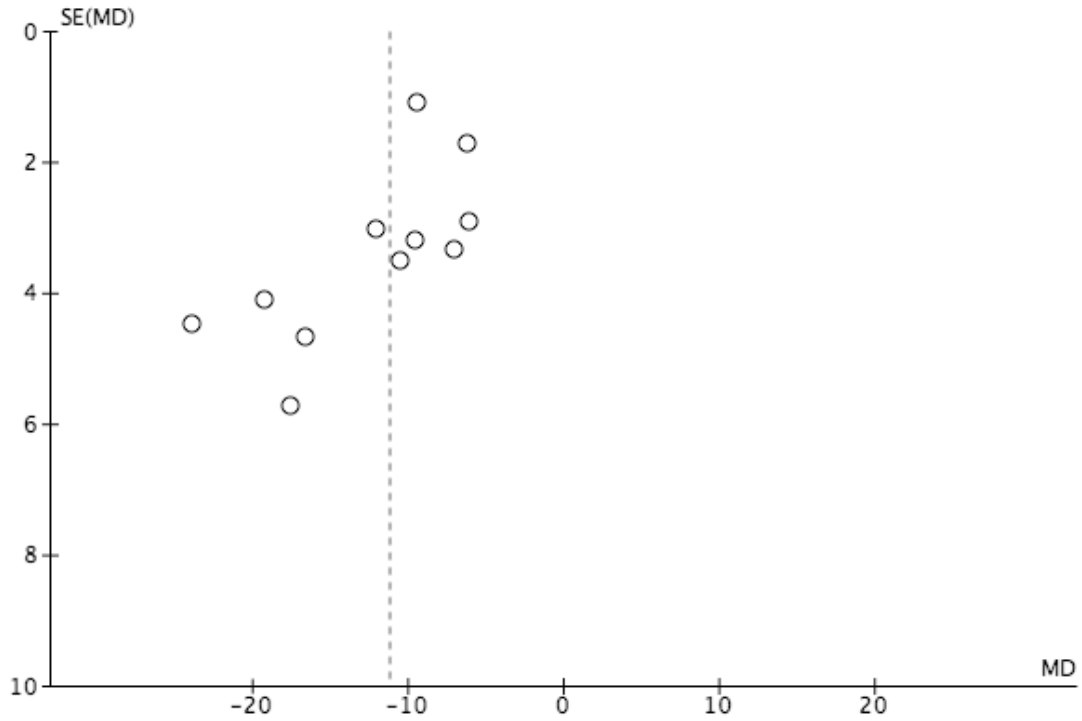
1) Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-OFF: UPDRS-III Mean Difference



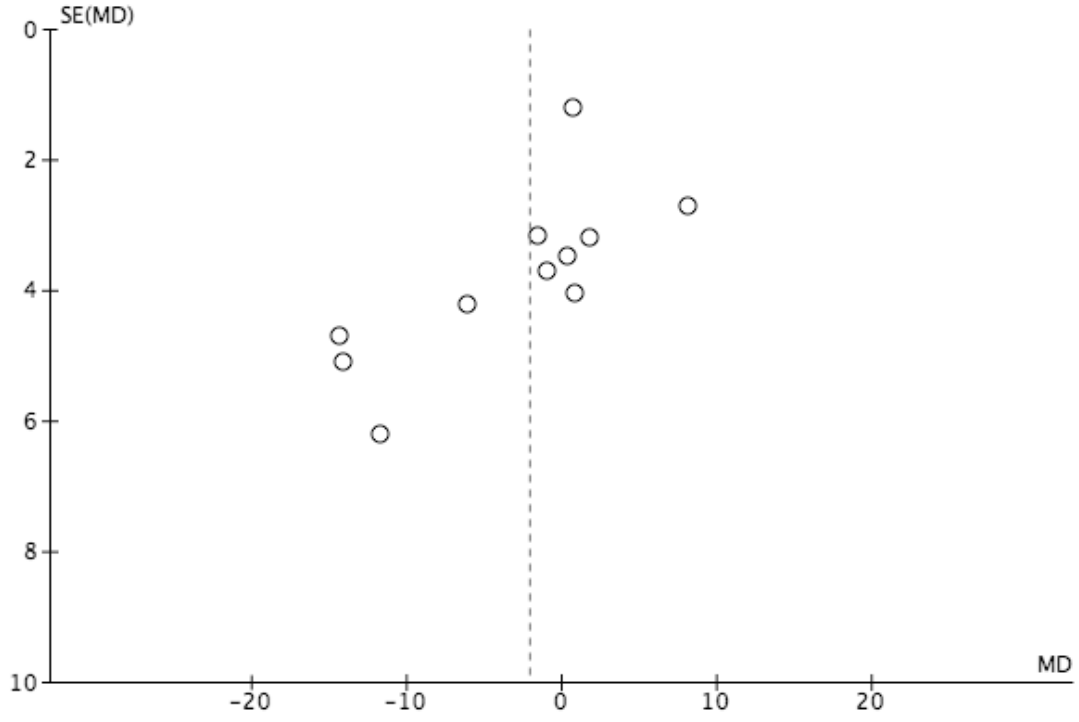
2) Stimulation-ON/Medication-ON vs. Stimulation-ON/Medication-OFF: UPDRS-III Mean Difference



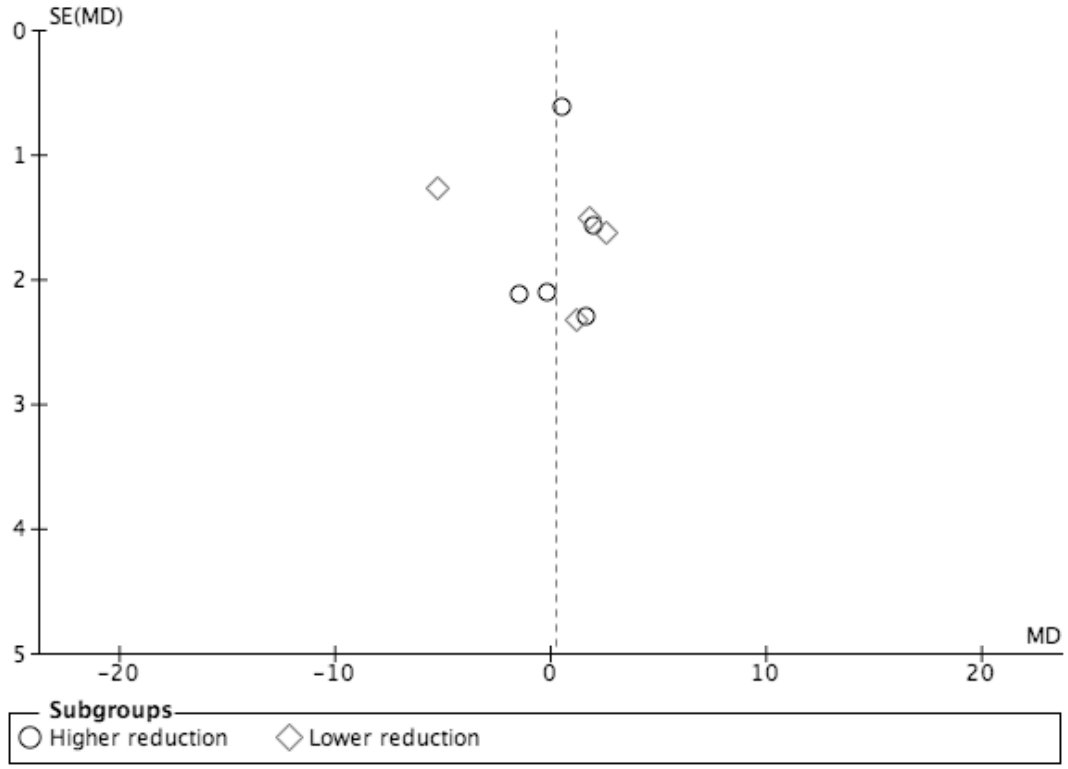
3) Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference



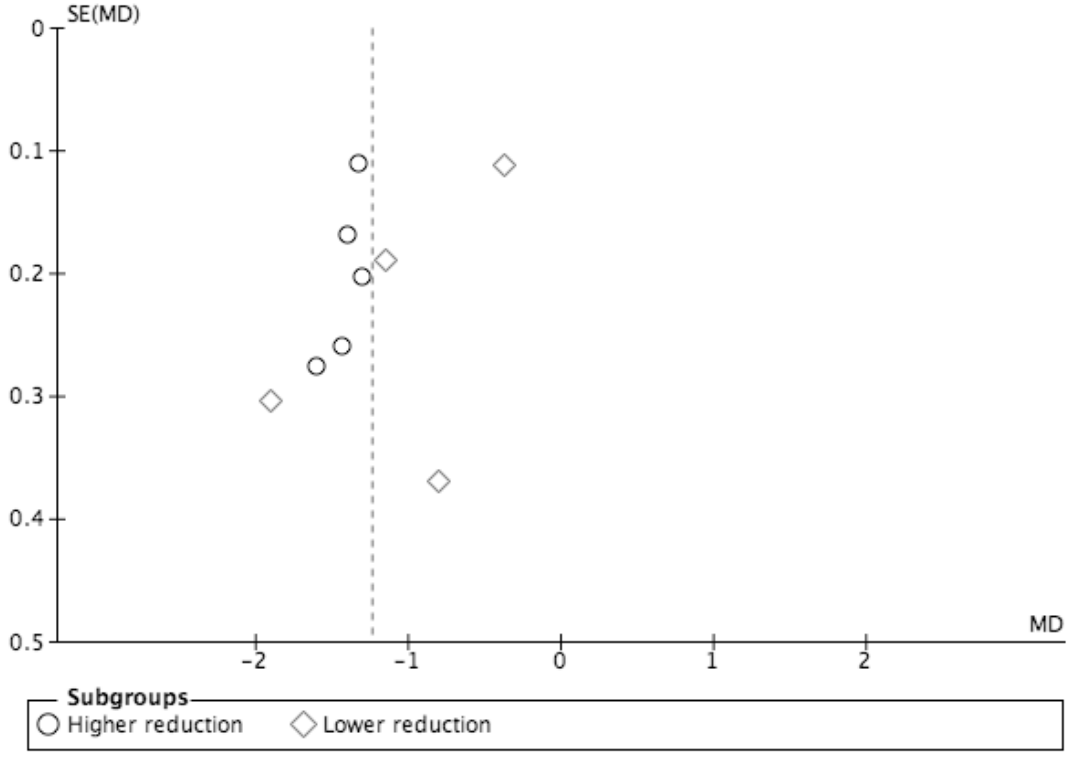
4) Stimulation-ON/Medication-OFF vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference



5) Higher vs Lower LEDD reduction in ADL: UPDRS-II Mean Difference

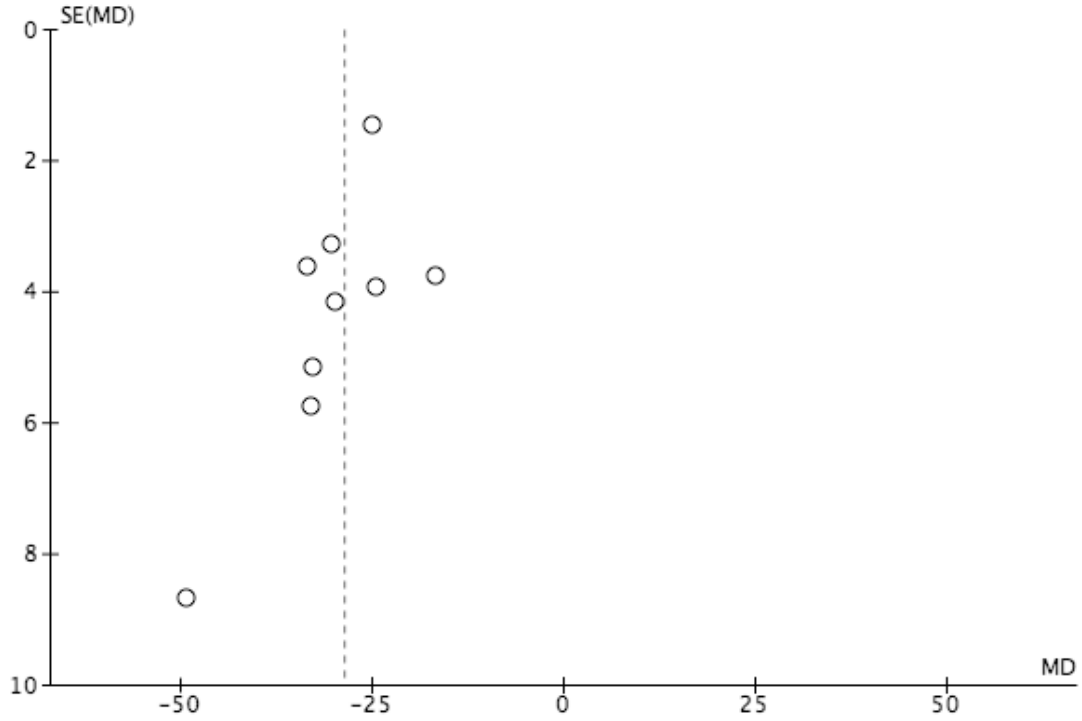


6) Higher vs Lower LEDD reduction in Dyskinesia duration: UPDRS – Item #32 Mean Difference

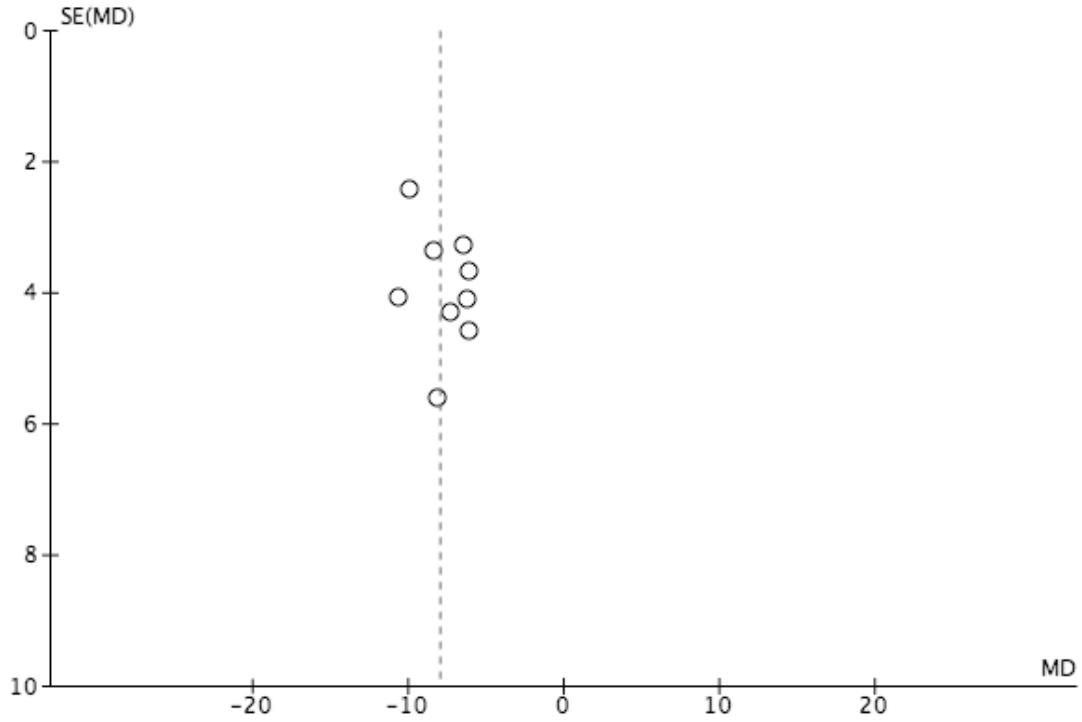


More than 5 years

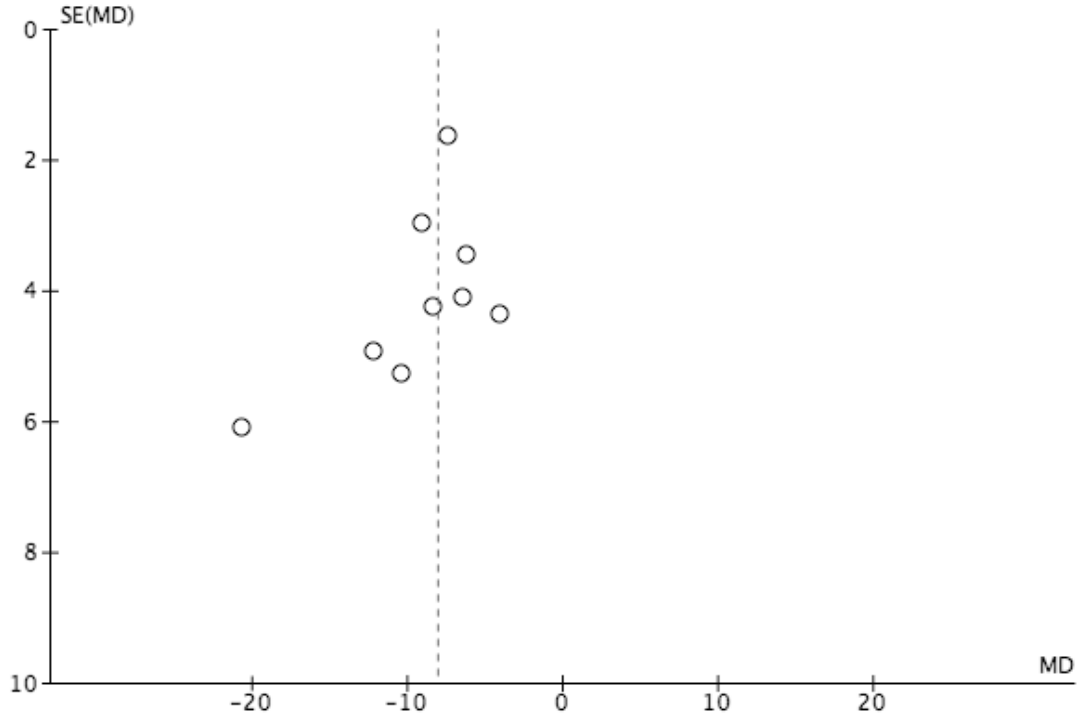
7) Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-OFF: UPDRS-III Mean Difference



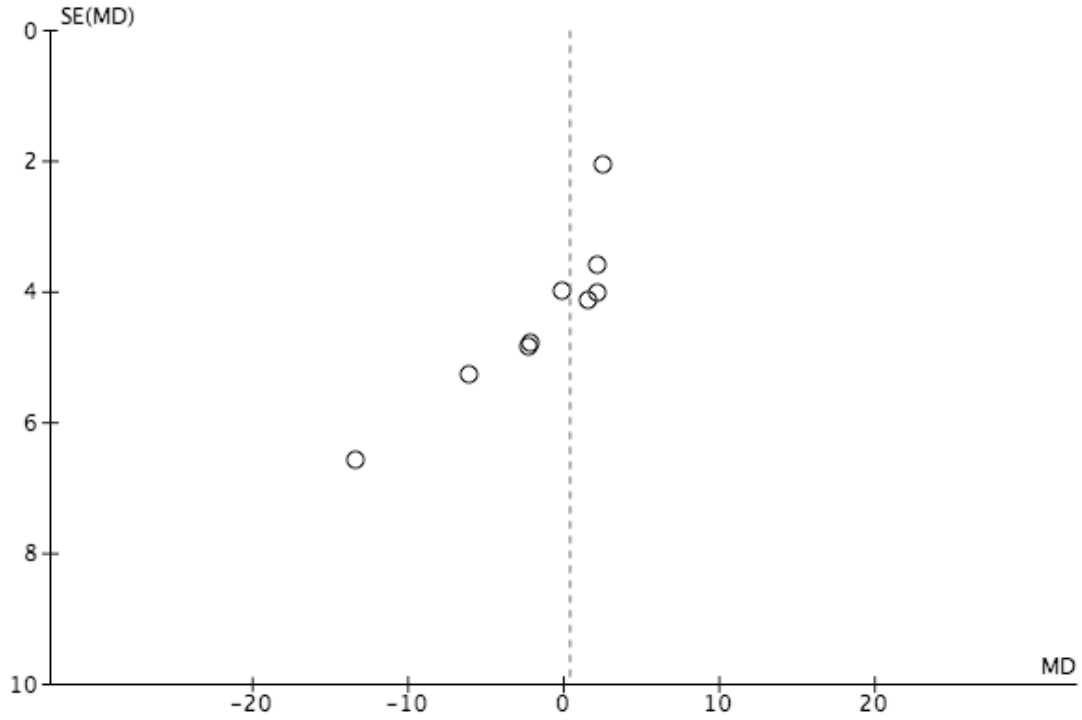
8) Stimulation-ON/Medication-ON vs. Stimulation-ON/Medication-OFF: UPDRS-III Mean Difference



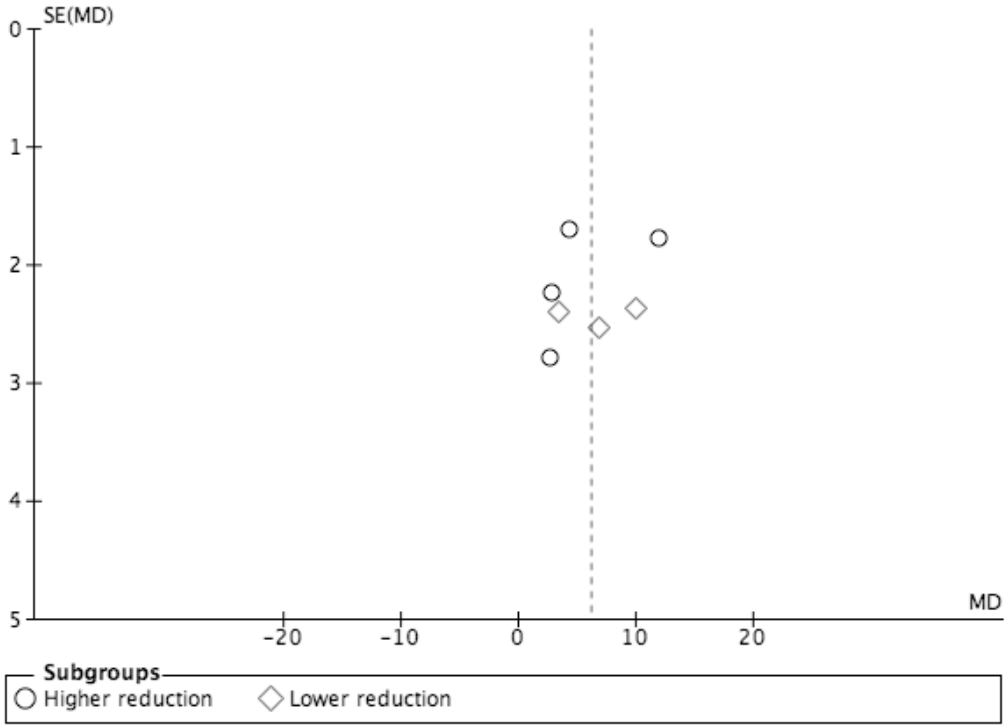
9) Stimulation-ON/Medication-ON vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference



10) Stimulation-ON/Medication-OFF vs. Stimulation-OFF/Medication-ON: UPDRS-III Mean Difference



11) Higher vs Lower LEDD reduction in ADL: UPDRS-II Mean Difference



12) Higher vs Lower LEDD reduction in Dyskinesia duration: UPDRS – Item #32 Mean Difference