

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Deep Brain Stimulation for Parkinson's Disease: A Review of Clinical Effectiveness, Cost- Effectiveness, and Guidelines

Service Line: Rapid Response Service
Version: 1.0
Publication Date: December 14, 2018
Report Length: 45 Pages

Authors: Chantelle Lachance, Carolyn Spry, Danielle MacDougall

Cite As: Deep brain stimulation for Parkinson's disease: A review of clinical effectiveness, cost-effectiveness, and guidelines. Ottawa: CADTH; 2018 Dec. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

Disclaimer: The information in this document is intended to help Canadian health care decision-makers, health care professionals, health systems leaders, and policy-makers make well-informed decisions and thereby improve the quality of health care services. While patients and others may access this document, the document is made available for informational purposes only and no representations or warranties are made with respect to its fitness for any particular purpose. The information in this document should not be used as a substitute for professional medical advice or as a substitute for the application of clinical judgment in respect of the care of a particular patient or other professional judgment in any decision-making process. The Canadian Agency for Drugs and Technologies in Health (CADTH) does not endorse any information, drugs, therapies, treatments, products, processes, or services.

While care has been taken to ensure that the information prepared by CADTH in this document is accurate, complete, and up-to-date as at the applicable date the material was first published by CADTH, CADTH does not make any guarantees to that effect. CADTH does not guarantee and is not responsible for the quality, currency, propriety, accuracy, or reasonableness of any statements, information, or conclusions contained in any third-party materials used in preparing this document. The views and opinions of third parties published in this document do not necessarily state or reflect those of CADTH.

CADTH is not responsible for any errors, omissions, injury, loss, or damage arising from or relating to the use (or misuse) of any information, statements, or conclusions contained in or implied by the contents of this document or any of the source materials.

This document may contain links to third-party websites. CADTH does not have control over the content of such sites. Use of third-party sites is governed by the third-party website owners' own terms and conditions set out for such sites. CADTH does not make any guarantee with respect to any information contained on such third-party sites and CADTH is not responsible for any injury, loss, or damage suffered as a result of using such third-party sites. CADTH has no responsibility for the collection, use, and disclosure of personal information by third-party sites.

Subject to the aforementioned limitations, the views expressed herein are those of CADTH and do not necessarily represent the views of Canada's federal, provincial, or territorial governments or any third party supplier of information.

This document is prepared and intended for use in the context of the Canadian health care system. The use of this document outside of Canada is done so at the user's own risk.

This disclaimer and any questions or matters of any nature arising from or relating to the content or use (or misuse) of this document will be governed by and interpreted in accordance with the laws of the Province of Ontario and the laws of Canada applicable therein, and all proceedings shall be subject to the exclusive jurisdiction of the courts of the Province of Ontario, Canada.

The copyright and other intellectual property rights in this document are owned by CADTH and its licensors. These rights are protected by the Canadian *Copyright Act* and other national and international laws and agreements. Users are permitted to make copies of this document for non-commercial purposes only, provided it is not modified when reproduced and appropriate credit is given to CADTH and its licensors.

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to Requests@CADTH.ca

Abbreviations

AGREE	Appraisal of Guidelines, Research and Evaluation
AMSTAR	A MeaSurement Tool to Assess systematic Reviews
BMI	body mass index
BMT	best medical therapy/best medical treatment
CADTH	Canadian Agency for Drugs and Technologies in Health
CRD	Centre for Reviews and Dissemination
DBS	deep brain stimulation
GPI	globus pallidus internus
GRADE	Grading Recommendations Assessment, Development and Evaluation
ICER	incremental cost-effectiveness ratio
ICUR	incremental cost-utility ratio
LEDD	levodopa-equivalent daily doses
NICE	National Institute for Health and Care Excellence
MDRS	Mattis Dementia Rating Scale
MMSE	Mini Mental State Examination
ODT	optimal drug therapy
PD	Parkinson's disease
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY	quality-adjusted life years
RCT	randomized controlled trial
UPDRS	Unified Parkinson disease rating scale
WAIS-III	Wechsler Adult Intelligence Scale – Third Edition

Context and Policy Issues

Parkinson's disease (PD) is one of the most common neurodegenerative disorders primarily affecting older adults.^{1,2} Over 67,000 Canadians have been diagnosed with PD, comprising 0.2% of the household population and 4.9% of the long-term care population.^{1,3} With the growing aging population in Canada, the number of Canadians diagnosed with PD is expected to double between 2011 and 2031.³

PD can have substantial impacts on those affected. Individuals with PD have a twofold increased risk of all-cause mortality.⁴ Common motor manifestations for individuals with PD include resting tremor, stiffness (rigidity), slowness of movements (bradykinesia), shuffling steps, soft voice, small handwriting (micrographia), and postural instability.⁵ Common non-motor manifestations include disturbances of mood, cognition, sleep, and autonomic dysfunction.^{6,7}

Currently, there is no cure for PD. To assist in mitigating the symptoms of the disease, individuals with PD are prescribed medication, such as levodopa, a dopamine agonist.⁶ After five years of pharmacological therapy, the majority of patients suffer medication-related complications, including dyskinesia and "on-off" fluctuations (i.e., sudden loss of benefit from medication). Other symptoms of the disease, problems with gait, balance, speech, swallowing, and cognition for example, may also become progressively resistant to pharmacologic therapies.^{6,8} Alternative treatments are warranted, especially once PD medications have become less effective for the patient.

Over the past few decades, deep brain stimulation (DBS) has been explored for the management of PD. DBS is a surgical treatment that modifies the irregular neuronal activity of the target region of the brain via electrical stimulation.⁹⁻¹¹ The procedure involves the placement of electrical leads into one (unilateral) or both (bilateral) sides of the basal

ganglia in the brain. The primary targets for DBS are usually the subthalamic nucleus (STN) or globus pallidus internus (GPI), but the thalamus can also be a target location. Symptoms such as tremor or dyskinesias determine which part of the brain should be targeted.⁹⁻¹¹ The DBS procedure is generally performed in two separate steps: implantation of the leads (usually using stereotactic methods) followed by implantation of the electrical pulse generator to which the leads are connected. The implantable pulse generator is implanted below the clavicle and it delivers the electrical pulses to the brain nuclei much like a pacemaker provides electrical stimulation to the heart to control heart rate.^{6,9,11}

In 2011, Canadian Agency for Drugs and Technologies in Health (CADTH) conducted a report on DBS for PD and provided a reference list of studies published between January 1, 2009 and August 19, 2011.¹² The current report aims to summarize evidence regarding the clinical and cost-effectiveness, as well as guidelines for the use of DBS in the treatment of PD. This report is, therefore, an update and upgrade to the previous CADTH report on this topic and may help knowledge users with purchasing decisions of DBS devices as a therapeutic intervention for patients with PD.

Research Questions

1. What is the comparative clinical effectiveness of deep brain stimulation versus standard care for treatment of Parkinson's disease?
2. What is the cost-effectiveness of deep brain stimulation versus standard care for treatment of Parkinson's disease?
3. What are evidence-based guidelines informing the use of deep brain stimulation for treatment of Parkinson's disease?

Key Findings

The clinical evidence on the effectiveness of deep brain stimulation for Parkinson's disease compared to standard of care (i.e., best medical therapy) was mixed. Several unique outcomes were used to quantify clinical effectiveness, making it challenging to compile the clinical evidence. Furthermore, the methodological quality of the evidence was moderate resulting in some uncertainty in the findings. Evidence from three systematic reviews and five clinical studies (one randomized controlled trial and four secondary analyses of randomized controlled trials suggests deep brain stimulation may be clinically effective at managing Parkinson's disease symptoms compared to standard of care; more research is warranted for definitive conclusions.

Four economic evaluations were identified that addressed the cost-effectiveness research question. Results suggest deep brain stimulation may be cost-effective treatment option compared to standard of care. The methodological quality of the evidence was moderate which should be considered when interpreting these results.

One evidence-based guideline satisfied the inclusion criteria for this report. This guideline recommends deep brain stimulation for people with advanced Parkinson's disease whose symptoms are not adequately controlled by best medical therapy.

Methods

Literature Search Methods

A limited literature search with main concepts appearing in the title or as major subject headings was conducted on key resources including PubMed, The Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused Internet search. Methodological filters were applied to limit retrieval to health technology assessments, systematic reviews, meta-analyses, randomized controlled trials (RCTs), non-randomized studies containing safety data, economic studies and guidelines. Where possible, retrieval was limited to the human population. The search was also limited to English language documents published between January 1, 2013 and November 15, 2018.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Patients of any age with Parkinson's disease
Intervention	Deep brain stimulation (i.e., surgically implanted device)
Comparator	Standard care (e.g., anti-Parkinson's medications like levodopa, dopamine agonists)
Outcomes	Q1. Clinical effectiveness (e.g., reduced tremor, improvement on standard measures of Parkinson's disease symptoms [e.g., UPDRS III off-medication score, PDQ-39 ADL score], reduced medication dependence); harms and/or adverse events Q2. Cost-Effectiveness Q3. Guidelines
Study Designs	Q1. Health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies pertaining to safety outcomes Q2. Economic evaluations, Q3. Evidence-based guidelines

PDQ-39 ADL = Parkinson's disease questionnaire: activities of daily living; UPDRS III = Unified Parkinson disease rating scale (motor section).

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, were duplicate publications (including systematic reviews with 100% overlap of their included studies), or were published prior to 2015 due to the volume of literature and availability of existing systematic reviews. Systematic reviews and clinical studies were excluded if all patients had DBS surgery and the comparison was "on" versus "off" stimulation (i.e., comparator was not standard of care as all patients had DBS surgery). Guidelines with unclear methodology were also excluded.

Critical Appraisal of Individual Studies

The included systematic reviews were critically appraised by one reviewer using A Measurement Tool to Assess systematic Reviews (AMSTAR) 2,¹³ randomized studies were

critically appraised using Downs and Black checklist,¹⁴ economic studies were assessed using the Drummond Checklist,¹⁵ and guidelines were assessed with the Appraisal of Guidelines, Research and Evaluation (AGREE) II instrument.¹⁶ Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations of each included study were described narratively.

Summary of Evidence

Quantity of Research Available

A total of 355 citations were identified in the literature search. Following screening of titles and abstracts, 311 citations were excluded and 44 potentially relevant reports from the electronic search were retrieved for full-text review. Four potentially relevant publications were retrieved from the grey literature search for full text review. Of these potentially relevant articles, 35 publications were excluded for various reasons, and 13 publications met the inclusion criteria and were included in this report. These comprised three systematic reviews, five clinical studies (1 RCT, 4 secondary analyses of RCTs), four economic evaluations, and one evidence-based guideline. Appendix 1 presents the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)¹⁷ flowchart of the study selection.

Additional references of potential interest are provided in Appendix 6.

Summary of Study Characteristics

A detailed summary of the characteristics of included publications are provided in Appendix 2.

Study Design

Three systematic reviews of clinical effectiveness outcomes were identified. All systematic reviews were published in 2016. One systematic review searched for relevant literature between the years 2000 and 2014,¹⁸ and two systematic reviews searched for relevant literature from inception to June 2015.^{19,20} All systematic reviews included comparative study designs ranging from randomized controlled trials only¹⁹ to including non-randomized and randomized controlled designs.^{18,20} Studies included in the systematic reviews were published between the years 2004 and 2015. The research question in one systematic review research question was broader than the research question of this report; thus, eight of the 16 included studies relevant for this report are described. Table 14 in Appendix 5 provides a detailed description of the overlap in the primary studies (n = 4) between the systematic reviews.

One single-centre RCT²¹ was identified that was not already captured in the included systematic reviews. A secondary analysis²² from a previous multi-centre RCT (captured in a systematic review)¹⁹ was identified that contained unique data for relevant outcome variables than those published in the original RCT; these data were not included in the primary report because the scale was validated after the RCT was published.²² Three post-hoc analyses²³⁻²⁵ from the same single-centre RCT (excluded from our report based on its year of publication; included in Appendix 6) were identified. One post-hoc study included all participants with baseline and 24-month UPDRS scores (n = 28) and a focused cohort of participants within the RCT with medication duration of one to four years at enrollment (n = 20).²³ The other two post-hoc studies included 29 participants.^{24,25} The five studies

examining clinical effectiveness are classified as clinical studies in this report.²¹⁻²⁵ All clinical studies were published between 2015 and 2018.²¹⁻²⁵

Four economic evaluations were identified, which included two cost-utility analyses, one cost-effectiveness, and one cost analysis. One cost-utility study used a cycle length of 1-year (2013) and a lifelong analytic time horizon from a German healthcare payer perspective.²⁶ The second cost-utility study conducted their work using a 1-, 5- and 10-year time horizons from a health and social care perspective.²⁷ The cost-effectiveness analysis used a 15-year time horizon from a United Kingdom payer (i.e., National Health Service) perspective.²⁸ Finally, the cost analysis analyzed medication cost and utilization from a 24-month pilot RCT and projected 10-year medication costs from the perspective of intervention (DBS plus optimal drug therapy) versus control group (optimal drug therapy).²⁹

One evidence-based guideline was identified, which was produced by the National Institute for Health and Care Excellence (NICE).³⁰ This guideline focuses on the diagnosis and management of PD in adults and contains a subsection about DBS informed by a systematic review of the evidence (RCTs for clinical evidence, cost-utility analyses for economic evidence) and input from content experts. The NICE Guideline Development Group recommendations are based on the trade-off between the benefits and harms of an intervention, taking into consideration the quality of the underpinning evidence.³⁰ The quality of evidence was evaluated using the Grading Recommendations Assessment, Development and Evaluation (GRADE) framework. The included recommendations are based on varying qualities of evidence (range: very low to high quality).

Country of Origin

The body of evidence originated from six countries: five studies from the United States (1 systematic review,¹⁸ 3 secondary analyses from 1 RCT,²³⁻²⁵ 1 economic evaluation²⁹), two studies from the United Kingdom (2 economic evaluations),^{27,28} two studies from China (2 systematic reviews),^{19,20} one study from Sweden (1 RCT),²¹ one study from Germany (economic evaluation),²⁶ and one study was conducted in both Germany and France (1 secondary analysis from a RCT);²² the guideline was published in the United Kingdom originally in 2006 and was revised in 2017.³⁰

Patient Population

All systematic reviews and clinical studies examined adult patients diagnosed with PD that underwent DBS surgery or best medical treatment (BMT; sometimes called best medical therapy or optimal drug therapy [ODT]).¹⁸⁻²⁵ There were no restrictions on sex or gender reported. Included systematic reviews and clinical studies did not report on whether patients were community-dwelling, residing in assisted living, or in long-term care.

The economic evaluations compared populations who received DBS compared to populations who received BMT to determine cost outcomes.²⁶⁻²⁹

The guideline aims to inform healthcare professionals, commissioners and providers, adults with PD and their families and caregivers about diagnosing and managing PD in people aged 18 years and over.³⁰

Interventions and Comparators

Two systematic reviews compared bilateral STN-DBS (intervention) to medication only (comparator).^{18,20} The third systematic review compared bilateral GPi-DBS to BMT (3 studies) and STN-DBS to BMT (5 studies; 4 bilateral electrode placement, 1 non-specified

electrode placement).¹⁹ One systematic review provided details on the comparator, which was “participants diagnosed with idiopathic PD and treated with dopaminergic drugs or dopamine-agonist, such as Madopar, Sinemet, pergolide, ropinirole, and pramipexole instead of STN-DBS.”²⁰ Definitions of BMT were not specified in the other two reviews.^{18,20} The follow-up period of the included studies ranged from 6 months to 5 years.

All five clinical studies investigated bilateral DBS to BMT. One RCT investigated bilateral caudal zona incerta (cZi) electrode placement²¹ and one clinical study conducted a secondary analysis comparing bilateral STN-DBS plus BMT to BMT alone.²² In addition, the three post-hoc analyses of the Charles et al. (2014)³¹ trial compared bilateral STN-DBS plus BMT to BMT alone.²³⁻²⁵

Two economic evaluations compared STN-DBS to BMT^{26,28} and two economic evaluations compared STN-DBS plus BMT to BMT.^{27,29} When clinical data was used to inform the evaluation, all patients received bilateral DBS except for two patients²⁷ (i.e., 176 of 178 procedures were bilateral).³²

The DBS subsection of the guideline compared patients who received DBS to patients who received BMT or levodopa-carbidopa intestinal gel plus BMT.³⁰

Outcomes

One systematic review investigated global cognition using Mini Mental State Examination (MMSE) and Mattis Dementia Rating Scale (MDRS), memory using Digital Span Backward, paired associate learning, Rey Auditory Verbal Learning Test-total, and Rey Auditory Verbal Learning Test-delayed recall, verbal fluency via phonemic and semantic fluency measures, and executive function using Raven’s Coloured Matrices, Stroop Color Word Test, Trail Making and Trail Making.²⁰ One systematic review reported Unified Parkinson’s Disease Rating Scale (UPDRS), Parkinson’s Disease Questionnaire 39 items (PDQ-39), levodopa-equivalent dose, and adverse events.¹⁹ Depending on the included primary study, whole score and sub scores of UPDRS were reported with stimulation on and sometimes reporting with stimulation off. One systematic review focused on verbal fluency by ascertaining letter fluency and category fluency scores.¹⁸

Several different outcomes were ascertained for the clinical studies. One secondary analysis reported motor symptoms using UPDRS-III (whole score and subscores), health-related quality of life using the PDQ-39, changes in dyskinesia and motor fluctuation scores using UPDRS-IV, and changes in LEDD (levodopa-equivalent daily doses; baseline versus 6-months for all outcomes).²¹ Another secondary analysis compared changes in behaviour using the validated Ardouin Scale of Behavior in PD, apathy using the Starkstein Apathy Scale, and depression using the Beck Depression Inventory, comparing baseline to 24-months. For the three post-hoc analyses based from the Charles et al. RCT,³¹ one study compared UPDRS-III and IV scores and used a composite score termed ‘clinically important worsening’ defined by the authors as both a ≥ 3 -point increase in UPDRS Part III and a ≥ 1 -point increase in Part IV from baseline to 24 months;²³ one study compared body-mass index (BMI) from baseline to 24 months;²⁴ and one study conducted a ‘neuropsychological test battery consisting of 12 tests, yielding a total of 22 specific neuropsychological variables and 10 personality test variables, administered in the off medication state’;²⁵ all included outcomes from this study did not overlap with the other two post-hoc studies.

Costs and quality-adjusted life-years (QALYs) were reported in three of the economic evaluations.²⁶⁻²⁸ Of these economic evaluations, a willingness-to-pay threshold was

reported in two studies,^{27,28} incremental cost-utility ratio (ICUR) was reported in one study²⁶ and incremental cost-effectiveness ratio (ICER) was reported in three studies.²⁶⁻²⁸ One economic evaluation reported medication costs for a two-year trial and provided a 10-year projection of medication costs.²⁹

Outcomes considered for the NICE guidelines³⁰ include: adverse events (perioperative, long-term complications including falls), symptom severity (UPDRS, dyskinesia, 'on' and 'off' time), disease progression (Hoehn & Yahr score), neuropsychiatric non-motor features (cognitive impairment, sleep disorder, suicidal ideation), health-related quality of life (patient, caregiver), information to inform decision making, resource use and cost (including medication load), and time to full time institutional care.

Summary of Critical Appraisal

Additional details regarding the strengths and limitations of included publications are provided in Appendix 3.

Systematic Reviews

The three systematic reviews¹⁸⁻²⁰ generally met the required criteria of the AMSTAR checklist.¹³ All reviews described their research question and inclusion criteria, searched for literature using multiple databases, searched reference lists of included studies to identify additional literature, and described the included studies in adequate detail.¹⁸⁻²⁰ These strengths increase the reproducibility of the findings. Two reviews mentioned performing data selection in duplicate^{18,19} and one study described performing data extraction in duplicate.²⁰ All studies described their statistical methods for synthesizing the data,¹⁸⁻²⁰ and two reviews described their assessment of risk of bias using validated tools.^{20,33} One study did not conduct a risk of bias assessment with included studies; therefore, review authors did not assess the potential risk of bias in individual studies on the results of the meta-analysis.¹⁸ The two reviews that performed a risk of bias assessment did not describe the potential impact of risk of bias on individual studies (i.e., excluding studies due to high risk of bias not described).^{19,20} Heterogeneity was discussed for results with $I^2 \geq 50\%$ or $P < 0.05$ in one review;²⁰ the remaining two reviews mentioned heterogeneity but provided limited details.^{18,19} Therefore, we are unaware of how this heterogeneity may affect the results of the review.^{18,19} Finally, no reviews mentioned having a predefined protocol about the process and objective of their review nor did they explain their selection of the study designs for the inclusion in the review. A protocol defines important aspects of a study a priori (e.g., outcome measures), helping to minimize bias and reduce ambiguity. By the included reviews not having a protocol available before conducting the review, we must interpret the findings with caution.

Clinical Studies

Three of the five clinical studies clearly reported their objectives, interventions, comparators, and main outcomes;^{21,22,23,25} two studies did not describe the intervention and comparator in adequate detail (e.g., surgery process not described) but did provide the clinicaltrial.gov registration numbers to locate the original publication.^{24,25} Four of the five clinical studies clearly described the characteristics of the study population;^{21,22,24,25} one study did not describe the population characteristics sufficiently, including a lack of details of their focused cohort that is different from the primary RCT.²³ All studies used appropriate statistical tests to assess the main outcomes, assessed the outcomes using the same time period between intervention and control groups, clearly described their main findings,

provided estimates of random variability, and stated their funding sources.²¹⁻²⁵ When examining the external validity of the findings (i.e., representativeness of the findings of the study), it is unclear if subjects who were asked to participate in the study were representative of the entire population recruited and if those subjects who were prepared to participate representative of the recruited population. Regarding internal validity (i.e., examining bias and confounders), it is unclear how long the recruitment period was, if participants in different groups were recruited over the same period of time, or if the intervention assignment was concealed until recruitment was complete for all clinical studies.²¹⁻²⁵ Moreover, it was not possible to blind patients to the intervention due to the nature of the intervention being investigated (i.e., brain surgery); all studies were single-blinded whereby the evaluator was blinded to the treatment allocations.²¹⁻²⁵ Four of the five clinical studies were secondary analysis of RCTs²²⁻²⁵ and not powered based on the outcomes included in this report, rather any power calculations conducted would be for the primary RCT.

Economic Evaluations

The four included economic evaluations²⁶⁻²⁹ satisfied the majority of the criteria outlined in the Drummond checklist.¹⁵ All economic evaluations described the viewpoint of the analysis the rationale for choosing the interventions, time horizon, form of the economic evaluation and justified their approach in relation to the question addressed.²⁶⁻²⁹ All evaluations also described the sources of effectiveness and the primary outcome measure of the analysis. However, none of the economic evaluations adequately described the alternatives that were being compared (i.e., BMT);²⁶⁻²⁹ knowing the type, brand, dose, and frequency of medication is important for reproducibility of findings. In addition, it is unclear whether any of the findings from the economic evaluations could be applied to the Canadian population.²⁶⁻²⁹ Moreover, one report did not describe important details including, the currency and year that the costs were calculated, the viewpoint of the analysis, and basic details on the models (e.g., discount rate), parameters and statistical methods used.²⁹ This report did reference four previous studies which may have described information, but it was not included in the economic evaluation. Finally, another evaluation described the limitations of previous cost-effectiveness studies conducted, but did not elaborate on what limitations were present in their current report (e.g., what are the limitations of using a deterministic model?).²⁷

Guidelines

The included guideline³⁰ met the majority of required criteria of the AGREE II¹⁶ tool. Strengths of the guidelines are the overall objectives and populations to whom the guidelines apply are specifically described; guideline development groups included individuals from all relevant professional groups; the target users of the guidelines are clearly defined; systematic methods were used to search for evidence; the criteria for selecting the evidence, the strengths and limitations of the body of evidence, and the methods for formulating the recommendations are clearly described; the health benefits, side effects, and risks have been considered in formulating the recommendations; there is an explicit link between the recommendations and the supporting evidence; the guideline was externally reviewed by experts prior to its publication; a procedure for updating the guideline is provided; the recommendations are specific and unambiguous; and the different options for management of the condition or health issue are clearly presented. These features may increase the reliability of the recommendations as they demonstrate sound methodology and make these publications less prone to biases.

There were a few aspects of the guideline that were unclear. Under the applicability subsection of the tool, it was unclear whether the potential resource implications of applying the recommendations have been considered and whether the guideline presents monitoring or auditing criteria. Under the editorial independence subsection, it was unclear whether the views of the funding body have not influenced the content of the guideline and whether competing interests of guideline development group members have been recorded and or addressed.

Summary of Findings

A detailed summary of findings and authors' conclusions from the included studies is available in Appendix 4.

Clinical Effectiveness of Deep Brain Stimulation to Standard of Care

Unified Parkinson's Disease Rating Scale (UPDRS; Total/Comprehensive Scores)

One secondary analysis of an RCT found that the STN-DBS plus BMT group improved at each time point on Total UPDRS versus the BMT group after 24 months.²³ This study also used components of the UPDRS III and IV to comprise a composite score called 'clinically important worsening' (i.e., ≥ 3 point increase in UPDRS Part III and ≥ 1 point increase in Part IV). Authors of this report found 54% of the BMT group and 27% of the DBS plus BMT group experienced clinically important worsening. Thus, the authors reported a 50% relative risk reduction for participants in the DBS plus BMT group compared to participants in the BMT only group. When applied to the focused cohort of participants in the DBS in early PD pilot trial on medication for 1 to 4 years at enrollment, there was an 80% reduction in the risk of clinically important worsening experienced by the DBS plus BMT group versus BMT group.²³

Mentation, Behavior and Mood

Part I of UPDRS.

One systematic review found no differences between the GPi-DBS group versus the BMT group (2 studies included) or STN-DBS group versus the BMT group on Part I of the UPDRS (i.e., mentation, behavior, and mood)¹⁹.

One systematic review found a significant difference in Mattis Dementia Rating Scale (MDRS) scores in favour of the control group versus STN-DBS, but found no differences between treatment groups for MMSE scores.²⁰ This review also found that the control group performed better at two of the four memory tests assessed (Rey Auditory Verbal Learning Test-total and Rey Auditory Verbal Learning Test-delayed recall); no differences were found between treatment groups for the Digital Span Backward and Paired Associate Learning tests.²⁰ No clinically significant differences were found in executive function between the STN-DBS and control group using 4 different executive function tests.²⁰ The Stroop Color Word test, a test for executive function reported a significant difference in favour of the control group, but had significant heterogeneity).²⁰ Moreover, a secondary analysis of an RCT found those who were treated with BMT performed better on three tests examining attention, working memory, and processing speed tests (WAIS-III Digit Span test, Paced Auditory Serial Addition Test for the slowest and fastest paced rates, Stroop Color and Word Test) at 12 months, but there was no difference between groups at 24 months. This study also found that BMT performed better than DBS plus BMT for components of the Wisconsin Card Sorting test at 12 months (perserverative errors) and 24 months (perserverative

errors, categories achieved). No other differences between groups were identified in the other memory, personality, and visual-spatial orientation tests.²⁵

In one secondary analyses, patients in STN-DBS plus BMT were compared to BMT only after 24 months on several behavioural factors.²² They identified that neuropsychiatric fluctuations subscale of the Ardouin scale differed significantly between treatment groups, favouring bilateral STN-DBS plus BMT group.²² The hyperdopaminergic behaviours score of the Ardouin scale decreased significantly with bilateral STN-DBS plus BMT but increased significantly with BMT only. Moreover, nocturnal hyperactivity, diurnal somnolence, creativity, and hobbyism increased with BMT alone and decreased under STN-DBS plus BMT. The same study found no differences between groups for the Starkstein Apathy Scale and Beck Depression Inventory.²²

Activities of Daily Living

Part II of UPDRS

One systematic review found that the GPi-DBS group outperformed the BMT group in Part II of the UPDRS when in the GPi-DBS group was on the off-medication phase (1 study), but no differences were found on the on-medication phase (2 studies).¹⁹ The STN-DBS group outperformed the BMT group in Part III of the UPDRS in both the on-medication (4 studies) and off-medication phases (3 studies).¹⁹

Verbal fluency

In one systematic review with 10 included studies, patients in the STN-DBS group had greater deficits in letter and category fluency than patients treated with BMT alone.¹⁸ A second systematic review also found significant differences in phonemic and semantic fluency in favour of the BMT only group when compared to the STN-DBS group.²⁰ In addition, a secondary analysis of a RCT also identified that the BMT only group performed better than the DBS plus BMT group in phonemic fluency at 12 months; however, no differences between groups were found with phonemic fluency after 24 months or with semantic fluency at either 12 months or 24 months.²⁵

Motor Examination

Part III of UPDRS

One systematic review found that the GPi-DBS group outperformed the BMT group in Part III of the UPDRS when in the GPi-DBS group was on the on-medication phase (3 studies) and off-medication phase (1 study).¹⁹ The STN-DBS group outperformed the BMT group in Part III of the UPDRS when in the STN-DBS group was on the on-medication phase (5 studies) only. There were no differences between STN-DBS and BMT groups when in the off-medication phase (3 studies).¹⁹

Two clinical studies collected UPDRS III scores for their patients to observe differences at 6 months (RCT)²¹ and at 24 months (secondary analysis of an RCT).²³ The scores were significantly better for DBS patients on-stimulation off-medication compared with medical patients off-medication.²¹ In the medical group on-medication versus DBS group on-medication on-stimulation, there were no significant differences in UPDRS scores.²¹ Looking at sub scores within the UPDRS III, tremor and akinesia improved in the DBS group on-stimulation off-medication versus the medical group off medication. No differences were identified between groups in speech, axial symptoms or dyskinesia.²¹ The secondary analyses found that the STN-DBS plus BMT group improved at each time point on Part III of the UPDRS versus the BMT group after 24 months ($P = 0.02$).²³

One secondary analysis examined manual dexterity using the Purdue Pegboard test. At 12 months and 24 months, no differences were found between the DBS plus BMT versus BMT only group.²⁵

Complications of Therapy

Part IV of UPDRS

One systematic review found that both GPi-DBS (on-medication phase; 2 studies) and STN-DBS (on-medication phase, 1 study) groups performed better on Part IV of UPDRS than the BMT groups.¹⁹ However, one RCT did not find changes between groups (DBS versus BMT) in any of the UPDRS IV sub items after 6 months.²¹

Health-related quality of life

PDQ-39

Findings from one systematic review suggest that DBS (STN or GPi) improved PDQ-39 scores compared with BMT (6 studies).¹⁹ One RCT found no difference between groups (DBS versus BMT) in the PDQ-39 summary index after 6 months.²¹

Medication Use

One systematic review found a decrease in medication dose in favour of DBS (STN or GPi) compared to BMT.¹⁹ Conversely, one RCT did not find differences between DBS versus BMT LEDD from baseline to 6 months.²¹ Results from a secondary analysis of an RCT found that the STN-DBS group had a significant decrease in LEDD by 39% compared to the BMT group (baseline to 24 months).²² Moreover, the authors provided descriptive details that antidepressants were stopped in 12 patients in the STN-DBS plus BMT group (versus 4 patients in BMT group) and 1 patient started neuroleptic drugs in the STN-DBS group (versus 9 patients in BMT group) over a 24-month period.²²

Adverse events

In a secondary analysis of an RCT, one report descriptively reported psychiatric serious events that happened after 24 months. In the STN-DBS group there were 19 psychiatric serious events, including 2 suicides, experienced by 17 patients (12%). In the BMT group there were 31 psychiatric serious events, including 1 suicide, experienced by 23 patients (17%).²²

In another secondary analysis of an RCT, BMI was compared at 24 months between STN-DBS plus BMT and BMT groups and no differences were found.²⁴

Cost-Effectiveness of Deep Brain Stimulation to Standard of Care

One economic evaluation calculated the cost for patients participating in a 24-month trial of DBS plus BMT versus BMT.²⁹ The study found that there was a cumulative cost savings for the DBS plus BMT group of \$7,150 (per patient, US currency presumed) over the 24-month study period.²⁹ This study used the trial data to project the medication costs over 10 years and estimated a \$64,590 cost savings in favour of the DBS plus BMT. Thus, a low dose medication regimen in the DBS plus BMT group may suggest a long-term reduction in medication costs.²⁹

One cost-utility analysis estimated the cost-effectiveness of bilateral STN-DBS over a 15-year time horizon from a UK payer perspective.²⁸ Compared to BMT, DBS was estimated to provide an ICER of £19,887 per QALY gained for PD patients. After conducting various

sensitivity analyses, the ICER was acceptable based on the willingness-to-pay thresholds of £30,000 per QALY gained.²⁸

From a life-long time horizon, one cost-utility analysis calculated the ICUR and ICER of bilateral STN-DBS versus BMT.²⁶ The authors estimated an ICUR of €22,710 per QALY gained and an ICER of €89 per PDQ-39 summary index point gained in favour of the STN-DBS group. Sensitivity analyses revealed that the outcome was most sensitive to battery exchange but never exceeded the threshold of €50,000.²⁶

One economic evaluation estimated the cost-effectiveness of DBS plus BMT over three time periods to examine the short- and long-term implications.²⁷ Over a 1 year period, they estimated an ICER of £468,528 per QALY, which largely exceeded the willingness-to-pay threshold for health gains in the UK of £20,000 to £30,000 per QALY gained. Additional sensitivity analyses revealed DBS had a low probability of being cost-effective after 1-year. For the 5-year projection, an ICER of £45,180 was found. One-way sensitivity analyses revealed that DBS may be cost-effective at year 5, but required any of the following: “a 10-year life span for the original implantable pulse generator and electrodes; surgery QALY gains increase by around 30%; and annual follow-up costs in the BMT group to increase by 30+%.²⁷ Reducing annual health and personal social services follow-up costs in the DBS group by 50% brought the ICER into the higher range of being cost-effective. Two way sensitivity analyses revealed that DBS may be effective at year 5, but required a 10-year life span for the original implantable pulse generator and electrodes and QALY gains in the surgery arm to increase by 10%. Moreover, the 10-year projection estimated an ICER of £70,537 per QALY gained, which is an increase compared to year 5. This calculation was most sensitive to the increased probability of battery replacements. Sensitivity analyses revealed that DBS can be cost-effective at 10-years if the implantable pulse generator and electrodes have a 10-year lifespan and increasing drug costs for the BMT group by 30%.²⁷

Guidelines

The NICE guidelines recommend (i) offering patients with advanced PD BMT which may include intermittent apomorphine injection and/or continuous subcutaneous apomorphine infusion; ii) not offering DBS to patients with PD whose symptoms are adequately controlled by BMT; and (iii) consider DBS for people with advanced PD whose symptoms are not adequately controlled by BMT.³⁰ These recommendations were developed using evidence of varying quality, ranging from very low quality (e.g., some adverse events and neuropsychological outcomes data) to high quality of evidence (e.g., some health-related quality of life patient data).

Limitations

There are certain limitations to consider when reviewing the report.

A major limitation of the body of clinical evidence is that it is often unclear whether the intervention of studies that compared DBS to BMT always had a drug treatment plan for DBS patients. More methodological details are needed to get an improved understanding of the treatment groups. In addition, the definition of standard care was not well reported in the included studies; information on the drug name, dose, frequency, compliance, and adherence data would be helpful for interpreting the findings. Furthermore, there were a large number of clinical outcomes used to assess clinical effectiveness and the results of these outcomes were mixed. Even though many studies included the validated UPDRS, data from the scale were reported in a variety of ways, including a total score, sub scores,

and composite scores. Reporting the total UPDRS score is easier to make comparisons across studies.

The report included four secondary analyses of two RCTs. Due to the nature of secondary analyses, included studies were not powered for the findings described in the report as the RCT would have been powered for the primary outcome(s) of the trial. None of the included studies were conducted in Canada. Therefore, it is unclear how generalizable the results of these studies, especially the economic evaluations, are to the Canadian settings as costs associated with surgery and medication may vary between countries. Finally, the recommendations presented in the evidence-based guideline were based on varying quality of evidence, so more primary studies are needed to fill the knowledge gaps.

Conclusions and Implications for Decision or Policy Making

A total of 13 relevant publications were identified, which comprised three systematic reviews,¹⁸⁻²⁰ five clinical studies (1 RCT and 4 secondary analyses),²¹⁻²⁵ four economic studies,²⁶⁻²⁹ and one evidence-based guideline.³⁰ The methodological quality of the studies was moderate resulting in some uncertainty in the findings. As such, caution should be applied in the interpretation of the studies.

Overall, the clinical effectiveness findings were mixed depending on the outcome examined. Most of the clinical outcome data identified in this report fell into one of the four categories of the UPDRS: (I) mentation, behaviour and mood;^{19,20,22} (II) activities of daily living;^{18-20,25} (III) motor examination;^{19,21,23,25} and (IV) complications of therapy.^{19,21} When considering total UPDRS scores, findings from one secondary analysis found that the STN-DBS BMT group performed better than the BMT group after 24-months.²³ For mentation, behaviour, and mood outcomes, one systematic review did not find any differences in UPDRS I scores when comparing DBS to standard care.¹⁸ Mixed findings from systematic review and clinical study data were identified for various cognition and behaviour outcomes,^{19,25} and no differences between groups were found for depression and apathy tests.²² For the activities of daily living outcomes, one systematic review found varying results for the UPDRS II.¹⁹ Two systematic reviews^{18,20} and one secondary analysis²⁵ identified higher verbal fluency scores in the standard of care (BMT) group compared to the DBS group. For the motor examination outcomes, three studies examined UPDRS scores (1 systematic review, 2 clinical studies) and the conclusions were mixed and seemed to depend on whether the DBS was inserted at the STN or GPi and whether the DBS group was being tested on the on-medication or off-medication.^{19,21,23} For complications of therapy outcomes, evidence from one systematic review suggests DBS performed better on the UPDRS IV compared to standard of care,¹⁹ but an RCT did not find differences in the same outcomes at a 6-month follow-up.²¹

Clinical effectiveness studies also examined health-related quality of life, medication dependence, and adverse events. After a 6-month follow-up, one RCT found no differences in health-related quality of life scores in the DBS group versus standard of care.²¹ However, a systematic review pooled results from six studies to reveal that DBS patients had better health-related quality of life scores than standard of care.¹⁹ A systematic review found reduced medication dependence in DBS patients versus patient on BMT alone (follow-up range: 6 months to 5 years).¹⁹ This difference was not observed in a RCT with a 6-month follow-up period²¹ meaning that differences may be more apparent between DBS and standard of care groups when observed over a longer follow-up period.^{19,22} Moreover, a secondary analysis reported fewer serious psychiatric events and suicides in the DBS

group compared to the medical group; however, more data is required to calculate statistical differences.²²

The economic evaluations suggest DBS may be cost-effective compared to standard of care (i.e. BMT), with the longer time horizons estimating larger QALY gains for the DBS group.²⁶⁻²⁹ Since the included economic evaluations were conducted from the perspective of countries other than Canada, it is unknown whether DBS would be cost-effective from a Canadian healthcare perspective. Economic evaluations from a Canadian lens are needed.

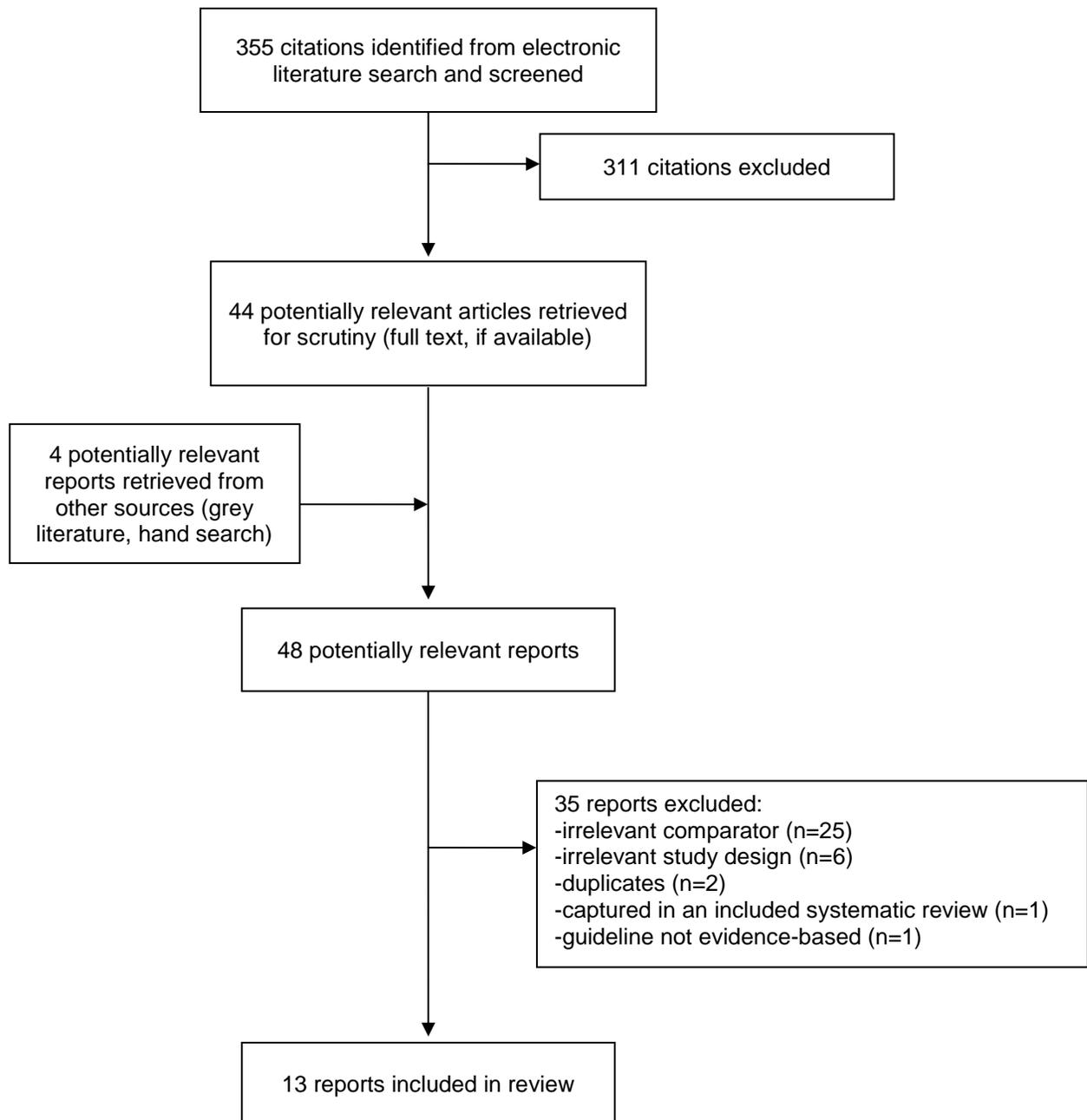
Finally, the evidence-based guideline from the United Kingdom put forth the recommendation of considering DBS for people with advanced PD whose symptoms are not adequately controlled by BMT based on variable quality of evidence.³⁰ Similar to the other included literature in this report, this guideline is intended for patients with PD in Britain and may not be generalizable to patients with PD in Canada.

The current report suggests that DBS versus standard of care may be a clinically effective means to treat patients with Parkinson's disease. However, findings from the available clinical evidence were inconsistent so no clear conclusions can be made. This is consistent with the previous CADTH report conducted in 2011.¹² In the previous CADTH report, there was limited evidence regarding the cost-effectiveness of DBS.¹² This report provides four additional economic evaluations that suggest that DBS may be a cost-effective intervention. Furthermore, a recent evidence-based guideline provides the recommendation that DBS may be considered for patients with advanced PD if their symptoms are not adequately controlled by BMT.

References

1. Wong SL, Gilmour H, Ramage-Morin PL. Parkinson's disease: prevalence, diagnosis and impact. *Health Rep.* 2014;25(11):10-14.
2. Hirtz D, Thurman D, Gwinn-Hardy K, Mohamed M, Chaudhuri A, Zalutsky R. How common are the "common" neurologic disorders? *Neurology.* 2007;68(5):326-337.
3. Mapping connections: an understanding of neurological conditions in Canada. Ottawa (ON): Public Health Agency of Canada; 2014: www.canada.ca/en/public-health/services/reports-publications/mapping-connections-understanding-neurological-conditions.html
4. Driver J, Kurth T, Buring J, Gaziano J, Logroscino G. Parkinson disease and risk of mortality: a prospective comorbidity-matched cohort study. *Neurology.* 2008;70(16 Part 2):1423-1430.
5. Lang AE, Lozano AM. Parkinson's disease. *N Engl J Med.* 1998;339(16):1130-1143.
6. Okun MS. Deep-brain stimulation for Parkinson's disease. *N Engl J Med.* 2012;367(16):1529-1538.
7. Chaudhuri KR, Martinez - Martin P, Schapira AH, et al. International multicenter pilot study of the first comprehensive self - completed nonmotor symptoms questionnaire for Parkinson's disease: the NMSQuest study. *Mov Disord.* 2006;21(7):916-923.
8. Olanow CW, Watts RL, Koller WC. An algorithm (decision tree) for the management of Parkinson's disease (2001): treatment guidelines. *Neurology.* 2001;56(suppl 5):S1-S88.
9. Benabid AL, Chabardes S, Torres N, et al. Functional neurosurgery for movement disorders: a historical perspective. *Prog Brain Res.* 2009;175:379-391.
10. Bötzel K, Kraft E. Strategies for treatment of gait and posture associated deficits in movement disorders: the impact of deep brain stimulation. *Restor Neurol Neurosci.* 2010;28(1):115-122.
11. Voges J, Koulousakis A, Sturm V. Deep brain stimulation for Parkinson's disease. *Acta Neurochir Suppl.* 2007;97(Pt 2):171-184.
12. Deep brain stimulation for Parkinson's disease and neurological movement disorders: clinical effectiveness, cost-effectiveness, and guidelines. (CADTH Rapid response report: reference list). Ottawa (ON): CADTH; 2011: <https://www.cadth.ca/sites/default/files/pdf/htis/aug-2011/RA0543%20Deep%20Brain%20Stimulation%20final%20Abs.pdf>. Accessed 2018 Nov 28.
13. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ.* 2017;358:j4008. <http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf>. Accessed 2018 Dec 13.
14. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf>. Accessed 2018 Dec 13.
15. Higgins JPT, Green S, editors. Figure 15.5.a: Drummond checklist (Drummond 1996). *Cochrane handbook for systematic reviews of interventions.* London (GB): The Cochrane Collaboration; 2011: http://handbook-5-1.cochrane.org/chapter_15/figure_15_5_a_drummond_checklist_drummond_1996.htm. Accessed 2018 Dec 13.
16. Consortium ANS. The AGREE II Instrument. Hamilton (ON): AGREE Enterprise; 2017: <https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf>. Accessed 2018 Dec 13.
17. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol.* 2009;62(10):e1-e34.
18. Wyman-Chick KA. Verbal fluency in Parkinson's patients with and without bilateral deep brain stimulation of the subthalamic nucleus: a meta-analysis. *J Int Neuropsychol Soc.* 2016;22(4):478-485.
19. Xie CL, Shao B, Chen J, Zhou Y, Lin SY, Wang WW. Effects of neurostimulation for advanced Parkinson's disease patients on motor symptoms: a multiple-treatments meta-analysis of randomized controlled trials. *Sci Rep.* 2016;6:25285.
20. Xie Y, Meng X, Xiao J, Zhang J, Zhang J. Cognitive changes following bilateral deep brain stimulation of subthalamic nucleus in Parkinson's disease: a meta-analysis. *Biomed Res Int.* 2016;2016:3596415.
21. Blomstedt P, Stenmark Persson R, Hariz GM, et al. Deep brain stimulation in the caudal zona incerta versus best medical treatment in patients with Parkinson's disease: a randomised blinded evaluation. *J Neurol Neurosurg Psychiatry.* 2018;89(7):710-716.
22. Lhomme 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.
23. Hacker ML, Tonascia J, Turchan M, et al. Deep brain stimulation may reduce the relative risk of clinically important worsening in early stage Parkinson's disease. *Parkinsonism Relat Disord.* 2015;21(10):1177-1183.
24. Millan SH, Hacker ML, Turchan M, Molinari AL, Currie AD, Charles D. Subthalamic nucleus deep brain stimulation in early stage Parkinson's disease is not associated with increased body mass index. *Parkinsons Dis.* 2017;2017:7163801.
25. Tramontana MG, Molinari AL, Konrad PE, et al. Neuropsychological effects of deep brain stimulation in subjects with early stage Parkinson's disease in a randomized clinical trial. *J Parkinsons Dis.* 2015;5(1):151-163.
26. Dams J, Balzer-Geldsetzer M, Siebert U, et al. Cost-effectiveness of neurostimulation in Parkinson's disease with early motor complications. *Mov Disord.* 2016;31(8):1183-1191.
27. McIntosh E, Gray A, Daniels J, et al. Cost-utility analysis of deep brain stimulation surgery plus best medical therapy versus best medical therapy in patients with Parkinson's: economic evaluation alongside the PD SURG trial. *Mov Disord.* 2016;31(8):1173-1182.
28. Fundament T, Eldridge PR, Green AL, et al. Deep brain stimulation for Parkinson's disease with early motor complications: a UK cost-effectiveness analysis. *PLoS One.* 2016;11(7):e0159340.
29. Hacker ML, Currie AD, Molinari AL, et al. Subthalamic nucleus deep brain stimulation may reduce medication costs in early stage Parkinson's disease. *J Parkinsons Dis.* 2016;6(1):125-131.
30. National Institute for Health Care and Excellence. Parkinson's disease in adults: diagnosis and management. (NICE guideline NG71) 2017; <https://www.nice.org.uk/guidance/ng71>. Accessed 2018 Nov 30.
31. Charles D, Konrad PE, Neimat JS, et al. Subthalamic nucleus deep brain stimulation in early stage Parkinson's disease. *Parkinsonism Relat Disord.* 2014;20(7):731-737.
32. Williams A, Gill S, Varma T, et al. Deep brain stimulation plus best medical therapy versus best medical therapy alone for advanced Parkinson's disease (PD SURG trial): a randomised, open-label trial. *Lancet Neurol.* 2010;9(6):581-591.
33. Xie T, Padmanaban M, Bloom L, et al. Effect of low versus high frequency stimulation on freezing of gait and other axial symptoms in Parkinson patients with bilateral STN DBS: a mini-review. *Transl Neurodegener.* 2017;6:13.
34. Charles D, Tolleson C, Davis TL, et al. Pilot study assessing the feasibility of applying bilateral subthalamic nucleus deep brain stimulation in very early stage Parkinson's disease: study design and rationale. *J Parkinsons Dis.* 2012;2(3):215-223.
35. Schuepbach WM, Rau J, Knudsen K, et al. Neurostimulation for Parkinson's disease with early motor complications. *N Engl J Med.* 2013;368(7):610-622.
36. Charles P, Dolhun R, Gill C, et al. Deep brain stimulation in early Parkinson's disease: enrollment experience from a pilot trial. *Parkinsonism Relat Disord.* 2012;18(3):268-273.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of Included Systematic Reviews and Meta-Analyses

First Author, Publication Year, Country	Study Designs and Numbers of Primary Studies Included	Population Characteristics	Intervention, Comparator	Clinical Outcomes, Length of Follow-Up
Wyman-Chick, 2016, ¹⁸ USA	10 studies, including randomized (i.e., RCTs) and non-randomized comparative studies (i.e., convenience samples)	831 patients with PD (439 who underwent bilateral STN-DBS, 392 non-surgical patients)	Bilateral STN-DBS, medication only	Verbal fluency via ascertaining letter fluency and category fluency scores, follow-up range: 3 to 36 months
Xie, 2016a, ²⁰ China	10 studies, including 3 RCTs and 7 non-randomized controlled trials	797 patients with PD (414 who underwent bilateral STN-DBS, 383 non-surgical patients)	Bilateral DBS, medication only (i.e., “dopaminergic drugs or dopamine agonist, such as Madopar, Sinemet, pergolide, ropinirole, and pramipexole instead of STN DBS”)	Global cognition using MMSE and MDRS; memory using Digital Span Backward, paired associate learning, RAVLT-total, and RAVLT-delayed recall; verbal fluency via phonemic and semantic fluency; executive function using Raven’s Coloured Matrices, Stroop Color Word Test, Boston naming, Trail Making a and Trail Making b, follow-up range: 6 to 24 months
Xie, 2016b, ¹⁹ China	16 RCTs; 8 relevant studies for the report	2,186 patients with PD; 1,347 patients relevant to the report	Among the relevant studies: Bilateral GPi-DBS, BMT (3 studies) STN-DBS, BMT (5 studies; 4 bilateral; 1 unspecified)	UPDRS (depending on primary study, whole score and subscores, stimulation on, and sometimes off), PDQ-39, LED, follow-up range: 6 months to 5 years

BMT = best medical therapy; DBS = deep brain stimulation; GPi = globus pallidus interna; LED= levodopa-equivalent dose; MDRS = Mattis Dementia Rating Scale; MMSE = Mini Mental State Examination; PDQ-39 = Parkinson’s Disease Questionnaire 39 items; RCT = randomized controlled trial; RAVLT-total = Rey Auditory Verbal Learning Test-total; STN = subthalamic nucleus; UPDRS =Unified Parkinson’s Disease Rating Scale; USA = United States of America

Table 3: Characteristics of Included Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention	Comparator(s)	Clinical Outcomes, Length of Follow-Up
Blomstedt, 2018, ²¹ Sweden	observer-blind RCT	n = 19 patients diagnosed with PD who would normally be considered for bilateral STN DBS (15 male, 4 female) Mean age for intervention group: 57±11.4 years Mean age for control group: 60.9±9.2 years	Bilateral cZi DBS	BMT	Motor symptoms using the UPDRS-III (whole score and subscores) Health-related quality of life using the PDQ-39 Changes in dyskinesia and motor fluctuation scores using UPDRS-IV Changes in LEDD 6-month follow-up for all outcomes
L'Hommée 2018, ²² France and Germany NCT00354133	Secondary analysis of an open-label RCT	n = 251 patients with PD (<61 years) and disabling motor fluctuations lasting for up to 3 years, with minimum disease duration of 4 years and at least 50% levodopa responsiveness (179 men, 72 women) Mean age for intervention group: 52.9±6.6 years Mean age for control group: 52.2±6.1 years	Bilateral STN-DBS plus BMT	BMT alone	Changes in behaviour using the Ardouin Scale of Behavior in PD Apathy using the Starkstein Apathy Scale, Depression using the Beck Depression Inventory. 24-month follow-up for all outcomes
Millan, 2017, ²⁴ USA NCT00282152	prospective, single-blind RCT	n = 29 patients with a clinical diagnosis of idiopathic early-stage PD (26 men, 3 women) Mean age for intervention group: 60±6.8 years Mean age for control group: 60±7.0 years	Bilateral STN-DBS + ODT	ODT	BMI 24-month follow-up

Table 3: Characteristics of Included Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention	Comparator(s)	Clinical Outcomes, Length of Follow-Up
Hacker 2015, ²³ USA NCT00282152	Post-hoc analysis of participants from a prospective, single-blind, parallel-group, pilot trial RCT	n = 28 patients with a clinical diagnosis of idiopathic early-stage PD part of pilot trial and subset (n = 20) of patients taking PD medications 1 to 4 years at enrollment of RCT	Bilateral STN-DBS + ODT	ODT	UPDRS-III and IV 'Clinically important worsening' (≥ 3 point increase in UPDRS Part III and ≥ 1 point increase in Part IV) 24-month follow-up
Tramontana, 2015, ²⁵ USA NCT00282152	prospective, single-blind, parallel-group RCT	n = 30 patients with a clinical diagnosis of probable idiopathic early-stage PD (27 men, 3 women) Mean age for intervention group: 60±6.8 years Mean age for control group: 60±7.0 years	Bilateral STN-DBS + ODT	ODT	Manual dexterity using Purdue Pegboard; Visual-spatial orientation (non-motor) using Benton Judgment of Line Orientation Test; Confrontation naming using Boston Naming Test; Rapid word production using Verbal Fluency; Attention/working memory using WAIS-III Digit Span; Attention/working memory/processing speed using Paced Auditory Serial Addition Test; immediate and delayed verbal memory using WMS-III Word List Learning I & II; Immediate and delayed visual memory using WMS-III Memory for Faces I & II; Executive functioning/flexible problem-solving using Wisconsin Card Sorting Test; Processing speed/control of competing responses using Stroop Color and Word Test; Emotional status and personality using Minnesota Multiphasic Personality Inventory-2

Table 3: Characteristics of Included Clinical Studies

First Author, Publication Year, Country	Study Design	Population Characteristics	Intervention	Comparator(s)	Clinical Outcomes, Length of Follow-Up
					12- and 24-month follow-up for all outcomes

BMI = body-mass index; BMT = best medical therapy; cZi = caudal zona incerta; DBS = deep brain stimulation; GPi = globus pallidus interna ; LEDD= levodopa-equivalent daily doses; ODT = optimal drug therapy; PD = Parkinson's disease; PDQ-39 = Parkinson's Disease Questionnaire 39 items; RCT = randomized controlled trial; STN = subthalamic nucleus; UPDRS =Unified Parkinson's Disease Rating Scale; USA = United States of America; WAIS-III = Wechsler Adult Intelligence Scale – Third Edition; WMS-III = Wechsler Memory Scale – Third Edition

Table 4: Characteristics of Included Economic Evaluations

First Author, Publication Year, Country	Type of Analysis, Time Horizon, Perspective	Decision Problem	Population Characteristics	Intervention and Comparator(s)	Approach	Clinical and Cost Data Used in Analysis	Main Assumptions
Dams, 2016, ²⁶ Germany	Cost utility, A cycle length of 1 year (2013) and a lifelong analytic time horizon, German health care payer perspective	To assess the cost utility of STN-DBS compared with the BMT for German patients below the age of 61 with early motor complications of PD.	Patients with PD (disease duration: 4+ years; disease severity rating below H & Y III on medication). All patients had suffered from motor complications for up to maximally 3 years.	Intervention: Bilateral STN DBS Comparator: BMT	Adapted Markov state-transition model	<u>Clinical data</u> RCT with 24-month follow-up ("EARLYSTIM" NCT00354133; Effectiveness was measured using the PDQ-39. <u>Cost data</u> direct costs were obtained from a cost of illness study; adaptations to STN DBS costs calculated using data from "EARLYSTIM" study; Utilities measured by the EuroQoL index with German tariffs and used to calculate QALYs.	A cost-effective therapy < 50,000 EUR/QALY
Fundament, 2016, ²⁸ UK	Cost-effectiveness, 15-year time horizon, UK payer (National Health Service) perspective	To evaluate the cost-effectiveness of DBS, versus BMT, among Parkinson's patients with early onset of motor complications	Patients with PD with early onset motor complications	Intervention: Bilateral STN DBS Comparator: BMT	Markov model	Data derived from a systematic review of clinical evidence; data from the EARLYSTIM study; and a UK Clinical Practice Research Datalink dataset including DBS patients. Effectiveness was measured using the UPDRS.	UK maximum willingness-to-pay threshold of £30,000 per QALY gained

First Author, Publication Year, Country	Type of Analysis, Time Horizon, Perspective	Decision Problem	Population Characteristics	Intervention and Comparator(s)	Approach	Clinical and Cost Data Used in Analysis	Main Assumptions
Hacker, 2016, ²⁹ USA	Cost analysis, 24 month (RCT duration) and 10-year (projection) time horizon, intervention versus control group perspective	To analyze medication cost and utilization from the pilot trial of DBS in early stage PD and to project 10 year medication costs.	Patients with Patients with advanced stage PD	Intervention: Bilateral STN DBS+ODT Comparator: ODT	Trial-based model	<u>Clinical data</u> RCT (NCT#00282152) <u>Cost data</u> Medication data collected at each visit and converted to levodopa equivalents. Daily medication costs were calculated using prices listed in the Drug Topics Red Book (2010 Edition).	lowest cost available was selected for medications with multiple pill quantities listed; 10% annual increase in medication cost was projected for both groups; a reduced annual increase in medication cost (5%) was projected for the intervention group
McIntosh, 2016, ²⁷ UK	Cost-utility analysis, 1-, 5-, and 10-year time horizon, health and social care perspective	To the costs and outcomes of DBS surgery and BMT at 1 year and estimates their comparative cost-effectiveness at 1, 5, and 10 years.	Patients with advanced stage PD	Intervention: DBS+BMT (176 of 178 procedures bilateral) ³² Comparator: BMT	Deterministic model	<u>Clinical data</u> RCT (ISRCTN34111222) for first year with extrapolation to 5 and 10 years <u>Cost data</u> patient-specific micro costing of preoperative, operative, and postoperative resources	Within-trial drug utilization, annual health and social service costs would continue out to 5 and 10 years for all resources other than annual clinic appointments and nurse visits, which would reduce to 2 and 1, respectively, and that the majority of surgical

First Author, Publication Year, Country	Type of Analysis, Time Horizon, Perspective	Decision Problem	Population Characteristics	Intervention and Comparator(s)	Approach	Clinical and Cost Data Used in Analysis	Main Assumptions
							complications would occur during the first 6 months postoperatively. Within-trial utility values, preserving between-arm differences, for each arm at the 1-year follow-up point were employed within a trajectory to estimate longer-term QALYs.

BMT = best medical therapy; DBS = deep brain stimulation; EUR = euro (currency); H & Y = Hoehn and Yahr scale; ODT = optimal drug therapy; PD = Parkinson's disease; PDQ-39 = Parkinson's Disease Questionnaire 39 items; QALYs = quality-adjusted life years; RCT = randomized controlled trial; STN = subthalamic nucleus; UK = United Kingdom; UPDRS = Unified Parkinson's Disease Rating Scale; USA = United States of America

Table 5: Characteristics of Included Guideline

Intended Users, Target Population	Intervention and Practice Considered	Major Outcomes Considered	Evidence Collection, Selection, and Synthesis	Evidence Quality Assessment	Recommendations Development and Evaluation	Guideline Validation
National Institute for Health and Care Excellence (NICE) Guideline, 2017 ³⁰						
Healthcare professionals Commissioners and providers Adults with PD and their families and caregivers	Relevant for this report, treatment approaches and management strategies for PD, including pharmacological (e.g., Levodopa) and non-pharmacological (e.g., DBS)	adverse events (perioperative, long-term complications including falls), symptom severity (UPDRS, dyskinesia, 'on' and 'off' time), disease progression (Hoehn & Yahr score), neuropsychiatric non-motor features (cognitive impairment, sleep disorder, suicidal ideation), health-related quality of life (patient, caregiver), information to inform decision making, resource use and cost (including medication load), and time to full time institutional care	Systematic review was conducted to identify and synthesize relevant literature (dates of search unclear) Call for evidence from stakeholders of unpublished RCT data and cost-utility analysis data	GRADE	The NICE Guideline Development Group makes a recommendation based on the trade-off between the benefits and harms of an intervention, taking into consideration the quality of the underpinning evidence	Recommendations were developed by the Guideline Development Group based on assessment of clinical- and cost-effectiveness evidence, with input from stakeholders (i.e., 2 expert witnesses involved in an RCT of DBS who provided details on their RCT and insight into its strengths and weaknesses) <i>*Stakeholders were not present when the evidence was reviewed and recommendations were made.</i>

DBS = deep brain stimulation; GRADE = Grading of Recommendations Assessment, Development, and Evaluation; PD = Parkinson's disease; RCT = randomized controlled trial

Appendix 3: Critical Appraisal of Included Publications

Table 6: Strengths and Limitations of Systematic Reviews and Meta-Analyses using AMSTAR 2¹³

Strengths	Limitations
Wyman-Chick, 2016 ¹⁸	
<ul style="list-style-type: none"> • Research questions/inclusion criteria for the review included the components of PICO • Multiple databases searched, keywords provided for the literature search, specific journals searched, and reference lists of included studies searched • Data selection performed in duplicate • Included studies described in adequate detail • Methods for statistical combination of results described • Hedge’s correction for small sample size bias calculated before descriptive analyses conducted. • Review authors reported no conflicts of interest 	<ul style="list-style-type: none"> • It is unclear whether review methods were established prior to the conduct of the review • Did not explain their selection of the study designs for inclusion in the review • It is not clear if data extraction was performed in duplicate • Reasons for excluding studies not provided and there were no accompanying list of excluded studies • Did not conduct ROB assessment (i.e., Cochrane ROB tool for RCTs) with included studies • Sources of funding not reported for the included studies • Did not describe how different study designs (i.e., RCTs versus non-randomized designs) would be weighed or accounted for in meta-analysis • Heterogeneity discussed but limited details; did not state how the heterogeneity present may affect the results • Review authors did not assess the potential impact of ROB in individual studies on the results of the meta-analysis; excluding studies due to high risk of bias was not described
Xie, 2016a ²⁰	
<ul style="list-style-type: none"> • Research questions/inclusion criteria for the review included the components of PICO • Multiple databases searched, select keywords provided for the literature search, and reference lists of included studies searched • Data extraction performed in duplicate • Included studies described in adequate detail • Cochrane ROB (for RCTs) and MINORS (for non-randomized studies) used to assess ROB of included studies • Appropriate methods for statistical combination of results used • Sensitivity analysis conducted • Heterogeneity discussed for results with $I^2 \geq 50\%$ or $p < 0.05$ • Begg’s test and funnel plots used to evaluate publication bias; evidence of publication bias was discussed. • Review authors reported no conflicts of interest 	<ul style="list-style-type: none"> • It is unclear whether review methods were established prior to the conduct of the review • Did not explain their selection of the study designs for inclusion in the review • Details on publication restrictions not provided • The literature search did not include trial registries or grey literature • It is unclear if study selection was performed in duplicate • Reasons for excluding studies not provided and there were no accompanying list of excluded studies • Sources of funding not reported for the included studies • Review authors did not assess the potential impact of ROB in individual studies on the results of the meta-analysis; excluding studies due to high ROB was not described
Xie, 2016b ¹⁹	
<ul style="list-style-type: none"> • Research questions/inclusion criteria for the review included the components of PICO 	<ul style="list-style-type: none"> • No predefined protocol about the process and objective of this review

Table 6: Strengths and Limitations of Systematic Reviews and Meta-Analyses using AMSTAR 2¹³

Strengths	Limitations
<ul style="list-style-type: none"> • Multiple databases searched, select keywords provided for the literature search, and reference lists of included studies searched • Data selection performed in duplicate • Included studies described in detail • Cochrane ROB used to assess ROB of RCTs • Methods for statistical combination of results used adequately described • Review authors reported no conflicts of interest 	<ul style="list-style-type: none"> • Did not explain their selection of the study designs for inclusion in the review • Details on publication restrictions not provided • It is not clear if data extraction was performed in duplicate • The literature search did not include trial registries or grey literature • Reasons for excluding studies not provided and there were no accompanying list of excluded studies • Sources of funding not reported for the included studies • Heterogeneity discussed but limited details; did not state how the heterogeneity present may affect the results • Review authors did not assess the potential impact of ROB in individual studies on the results of the meta-analysis; excluding studies due to high ROB was not described • Small study bias mentioned but limited details

MINORS = methodological index for nonrandomized studies; PICO = population, intervention, comparator, outcome; RCT = randomized controlled trial; ROB = risk of bias

Table 7: Strengths and Limitations of Clinical Studies using Downs and Black checklist¹⁴

Strengths	Limitations
Blomstedt, 2018 ²¹	
<ul style="list-style-type: none"> Objectives, intervention, comparator, and main outcomes of the study were clearly described Characteristics of the study population were clearly described Estimates of the random variability were provided as standard deviation values The evaluators were blind to the treatment allocations The main findings of the study were clearly described Patients lost to follow-up were described The time period between intervention and outcome the same for intervention and control groups (i.e., 6 months) Adverse events were reported Funding for the study clearly stated Appropriate statistical tests were used to assess the main outcomes 	<ul style="list-style-type: none"> It is unclear whether the participants were representative of the source population Due to the type of intervention, blinding of participants not possible Attrition of participants after directly after randomization (n = 4; 1 from intervention group, 3 from control group). It is unclear if any of the dropouts were due to their allocation It is unclear how long the recruitment period was, if participants in different groups were recruited over the same period of time, or if the intervention assignment was concealed until recruitment was complete. Patients' compliance and/or adherence with PD medication not described Sample size for statistical power was not calculated
L'Hommée, 2018 ²²	
<ul style="list-style-type: none"> Objectives, intervention, comparator, and main outcomes of the study were clearly described with references to previous work and guidelines, as appropriate. Characteristics of the study population were clearly described Estimates of the random variability were provided as standard deviation and standard error values, as appropriate At preoperative assessment, both clinicians and patients were masked to treatment assignment The main findings of the study were clearly described The time period between intervention and outcome the same for intervention and control groups (i.e., 24 months) Appropriate statistical tests were used to assess the outcomes of interest Adverse events were reported in detail Funding for the study clearly stated 	<ul style="list-style-type: none"> It is unclear whether the participants were representative of the source population It is unclear in this report how long the recruitment period was, if participants in different groups were recruited over the same period of time, or if the intervention assignment was concealed until recruitment was complete. Patients' compliance and/or adherence with PD medication not described Sample size for statistical power was calculated for the primary outcome and not the outcomes of this analysis. Reasons for why patients lost to follow-up not described
Millan, 2017 ²⁴	
<ul style="list-style-type: none"> Objectives and main outcome of the study were clearly described Valid outcome measure used (i.e., BMI) Characteristics of the study population were clearly described The main findings of the study were clearly described Estimates of the random variability were provided as standard deviation values The time period between intervention and outcome the same for intervention and control groups Appropriate statistical tests were used to assess the main outcome Adverse event described for 1 participant, which was 	<ul style="list-style-type: none"> Intervention and comparator not described in detail (e.g., surgery process not described); patients' compliance and/or adherence with PD medication not described Single blind study; assumption: outcome assessor(s) blinded given the nature of the trial, but not described It is unclear whether the participants were representative of the source population It is unclear how long the recruitment period was, if participants in different groups were recruited over the same period of time, or if the intervention assignment was concealed until recruitment was complete Sample size for statistical power was not calculated

Table 7: Strengths and Limitations of Clinical Studies using Downs and Black checklist¹⁴

Strengths	Limitations
<p>deemed unrelated to the intervention; weight loss from this adverse event was accounted by performing a secondary analysis with the outlier excluded</p> <ul style="list-style-type: none"> Funding for the study clearly stated 	
<p>Hacker 2015²³</p>	
<ul style="list-style-type: none"> Objectives, intervention, comparator, and main outcomes of the study were clearly described UPDRS III outcome was conducted by evaluators who were blind to the treatment allocations Appropriate statistical tests were used to assess the main outcomes The main findings of the study were clearly described Estimates of the random variability were provided as 95% confidence intervals The time period between intervention and outcome the same for intervention and control groups (i.e., 24 months) Funding for the study clearly stated 	<ul style="list-style-type: none"> Lack of details on the study population Due to the type of intervention, blinding of participants not possible; open-label for all measures except for a single-blind assessment of UPDRS Part III No power calculation due to the nature of this post-hoc analysis With this post-hoc analysis of a pilot RCT, this report did not describe (but may be retrieved from original publication³¹) <ul style="list-style-type: none"> detailed characteristics of the study population how long the recruitment period was if participants in different groups were recruited over the same period of time if the intervention assignment was concealed until recruitment was complete patients lost to follow-up from original pilot RCT Patients' compliance and/or adherence with PD medication adverse events
<p>Tramontana, 2015²⁵</p>	
<ul style="list-style-type: none"> Objectives, intervention, comparator, and main outcomes of the study were clearly described Characteristics of the study population were clearly described Estimates of the random variability were provided as standard deviation values The main findings of the study were clearly described Patients lost to follow-up were described The time period between intervention and outcome assessment the same for intervention and control groups (i.e., 12- and 24-months) Adverse events were reported Funding for the study clearly stated Appropriate statistical tests were used to assess outcomes Report referenced primary study protocol³⁴ for further details about power calculations and randomization process 	<ul style="list-style-type: none"> Intervention and comparator not described in detail (e.g., surgery process not described in this report) Due to the type of intervention, blinding of participants not possible (single-blind study) It is unclear whether the participants were representative of the source population It is unclear how long the recruitment period was, if participants in different groups were recruited over the same period of time, or if the intervention assignment was concealed until recruitment was complete. Patients' compliance and/or adherence with PD medication not described

BMI = body mass index; PD = Parkinson's disease; RCT = randomized controlled trial; UPDRS = Unified Parkinson's Disease Rating Scale

Table 8: Strengths and Limitations of Economic Studies using the Drummond Checklist¹⁵

Strengths	Limitations
Dams, 2016 ²⁶	
<ul style="list-style-type: none"> • The research question, its economic importance, the viewpoint of the analysis, and the rationale for choosing the interventions stated • The form of economic evaluation is stated and justified in relation to the question addressed • The sources of effectiveness estimated are stated • The primary outcome measure for the economic evaluation is stated • Sources and methods used for estimating costs are stated • The currency used for all costs (i.e., 2013, euros) was stated • Details of the models are given and the key parameters are justified • The time horizon is stated • Both costs and health effects were half-cycle corrected and discounted with 3% per annum to adjust for future values • Details of statistical methods and approaches to sensitivity analyses are provided • The choice of variables for sensitivity analysis is justified; ranges for sensitivity analysis are provided • Incremental analysis reported • Provided a comparison of base-case findings from this report to other cost studies of DBS for Parkinson’s • The conclusions follow from the data reported and are clearly stated with appropriate limitations identified 	<ul style="list-style-type: none"> • The alternatives being compared (i.e., BMT) are not clearly described; the report references the RCT³⁵ where these details may be found. • It is uncertain if these findings can be applied to the local population (i.e., Canada)
Fundament, 2016 ²⁸	
<ul style="list-style-type: none"> • The research question, its economic importance, the viewpoint of the analysis, and the rationale for choosing the interventions stated • The form of economic evaluation is stated and justified in relation to the question addressed • The sources of effectiveness estimated are described • The primary outcome measure for the economic evaluation is clearly stated • Sources and methods used for estimating costs are stated • The currency used for all costs (Great British Pounds) was stated • Details of the models are given and the key parameters are justified • The time horizon is stated • Costs and QALYs were both discounted at 3.5% per year, according to NICE methods guidance • Details of statistical methods and approaches to sensitivity analyses are provided • The choice of variables for sensitivity analysis is justified; ranges for sensitivity analysis are provided • Incremental analysis is reported • The conclusions follow from the data reported and are 	<ul style="list-style-type: none"> • The alternatives being compared (i.e., BMT) are not clearly described; the report references the RCT³⁵ where these details may be found. • It is uncertain if these findings can be applied to the local population (i.e., Canada)

Table 8: Strengths and Limitations of Economic Studies using the Drummond Checklist¹⁵

Strengths	Limitations
clearly stated with appropriate limitations identified	
Hacker, 2016 ²⁹	
<ul style="list-style-type: none"> • The research question, its economic importance, and the rationale for choosing the interventions stated • The form of economic evaluation is stated and justified in relation to the question addressed • The sources of effectiveness estimated are described • The primary outcome measure for the analysis is stated • Sources and methods used for estimating costs are stated • The time horizon is stated • The conclusions follow from the data reported and limitations are identified 	<ul style="list-style-type: none"> • The viewpoint of the analysis not explicitly described • Report references 4 other studies that provides information about study design, enrollment experience, baseline characteristics, and outcomes;^{23,31,34,36} not described in detail in this report. Thus, alternatives being compared (i.e., optimal drug therapy) are not clearly described in report • The currency of costs not explicitly stated (implied US dollars, implied medications costs are based on those medications available between 2006 and 2012) • Lacking basic details of models, parameters, and statistical methods described • No mention of incremental analysis • It is uncertain if these findings can be applied to the local population (i.e., Canada)
McIntosh, 2016 ²⁷	
<ul style="list-style-type: none"> • The research question, the viewpoint of the analysis, and the rationale for choosing the interventions stated • The form of economic evaluation is stated and justified in relation to the question addressed • The sources of effectiveness estimated are described • The primary outcome measure for the economic evaluation is clearly stated • Sources and methods used for estimating costs are stated • The currency used for all costs (i.e., 2010 Great British Pounds) was stated • Details of the models are given and the key parameters are justified • The time horizon is stated • Discount rate for annuitization of capital items was 3.5% (HM Treasury) • Details of statistical methods and approaches to sensitivity analyses are provided • The choice of variables for sensitivity analysis is justified; ranges for sensitivity analysis are provided • Incremental analysis is reported • The conclusions follow from the data reported and are clearly stated 	<ul style="list-style-type: none"> • The economic importance not described in introduction • The alternatives being compared (i.e., BMT) are not clearly described • It is uncertain if these findings can be applied to the local population (i.e., Canada) • Limitations of the report not described

BMT = best medical treatment; DBS = deep brain stimulation; NICE= National Institute for Health and Care Excellence; RCT = randomized controlled trial; QALY = quality-adjusted life years

Table 9: Strengths and Limitations of Guideline using AGREE II¹⁶

Item	Guideline NICE 2017 ³⁰
Domain 1: Scope and Purpose	
1. The overall objective(s) of the guideline is (are) specifically described.	✓
2. The health question(s) covered by the guideline is (are) specifically described.	✓
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.	✓
Domain 2: Stakeholder Involvement	
4. The guideline development group includes individuals from all relevant professional groups.	✓
5. The views and preferences of the target population (patients, public, etc.) have been sought.	✓
6. The target users of the guideline are clearly defined.	✓
Domain 3: Rigour of Development	
7. Systematic methods were used to search for evidence.	✓
8. The criteria for selecting the evidence are clearly described.	✓
9. The strengths and limitations of the body of evidence are clearly described.	✓
10. The methods for formulating the recommendations are clearly described.	✓
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	✓
12. There is an explicit link between the recommendations and the supporting evidence.	✓
13. The guideline has been externally reviewed by experts prior to its publication.	✓
14. A procedure for updating the guideline is provided.	✓
Domain 4: Clarity of Presentation	
15. The recommendations are specific and unambiguous.	✓
16. The different options for management of the condition or health issue are clearly presented.	✓
17. Key recommendations are easily identifiable.	✓
Domain 5: Applicability	
18. The guideline describes facilitators and barriers to its application.	✓
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	✓
20. The potential resource implications of applying the recommendations have been considered.	unclear
21. The guideline presents monitoring and/or auditing criteria.	unclear
Domain 6: Editorial Independence	
22. The views of the funding body have not influenced the content of the guideline.	unclear
23. Competing interests of guideline development group members have been recorded and addressed.	unclear

Appendix 4: Main Study Findings and Authors' Conclusions

Table 10: Summary of Findings Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusion
Wyman-Chick, 2016 ¹⁸	
<p>Verbal fluency. <i>Distribution of effect sizes (n = 10 studies) analyzed to determine if patients in the STN-DBS group experience greater deficits in verbal fluency versus patients in the medication only group.</i></p> <ul style="list-style-type: none"> • Letter fluency: mean effect size = -0.47 (SE = 0.08, z = -5.67, P < 0.001, medium effect size). Patients in STN-DBS group had greater deficits in letter fluency than patients treated with medication only • Category fluency: mean effect size = -0.31 (SE = 0.08, z = -3.91, P < 0.001, small effect size). Patients in STN-DBS group had greater deficits in category fluency than patients treated with medication only. 	<p><i>“The results of the current meta-analysis indicate that patients with PD who underwent bilateral STN-DBS surgery have statistically significant deficits in verbal fluency compared to patients with PD who are treated with medication only. Furthermore, deficits in letter fluency were greater than deficits in category fluency. A medium effect size was observed for letter fluency and a small effect size was observed for category fluency.”</i> (p 483)¹⁸</p> <p><i>“While the results of this meta-analysis reflect statistically significant differences in verbal fluency between patients who have undergone STN-DBS and those managed with medication only, the clinical impact of these findings is unclear.”</i> (p 483)¹⁸</p>
Xie, 2016a ²⁰	
<p>All outcomes comparing STN-DBS to control group. Any significant differences found were in favour of the control group.</p> <p>Global cognition</p> <ul style="list-style-type: none"> • MMSE: non-significant difference (d = -0.06; 95% CI, -0.23 to 0.36) • MDRS: significant difference (d = -0.21; 95% CI, -0.42 to -0.01) <p>Memory</p> <ul style="list-style-type: none"> • Digital Span Backward: non-significant difference (d = -0.14; 95% CI, -0.33 to 0.05) • Paired associate learning: non-significant difference (d = -0.69; 95% CI, -2.01 to 0.63) • RAVLT-total: significant difference (d = -2.06; 95% CI, -4.06 to -0.06) • RAVLT-delayed recall: significant difference (d = -1.41; 95% CI, -2.23 to -0.58) <p>Verbal fluency</p> <ul style="list-style-type: none"> • Phonemic fluency: significant difference (d = -0.49; 95% CI, -0.66 to -0.31) • Semantic fluency: significant difference (d = -0.39; 95% CI, -0.63 to -0.15) <p>Executive function</p> <ul style="list-style-type: none"> • Raven's Coloured Matrices: non-significant difference 	<p><i>“The results indicate that chronic stimulation of STN could cause subtle decline in global cognition, memory, phonemic fluency, semantic fluency, and executive function.”</i> (p 5)²⁰</p> <p><i>“In conclusion, STN DBS is relatively safe for PD patients, although small decline in global cognitive function, memory, verbal fluency, and executive function is observed. Further studies should focus on the exact mechanisms and factors associated with memory, verbal fluency, and executive function decline in order to provide practical information to clinical doctors”</i> (p 5)²⁰</p>

Table 10: Summary of Findings Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusion
<p>($d = -0.15$; 95% CI, -0.56 to 0.25)</p> <ul style="list-style-type: none"> • <u>Stroop Color Word Test</u>: significant heterogeneity, $I^2 = 75\%$, reported significant difference ($d = -0.23$; 95% CI, -0.39 to -0.06) • <u>Boston naming</u>: non-significant difference ($d = -0.02$; 95% CI, -0.21 to 0.26) • <u>Trail Making a</u>: non-significant difference ($d = -0.03$; 95% CI, -0.17 to 0.22) • <u>Trail Making b</u>: significant heterogeneity, $I^2 = 92\%$, reported non-significant difference ($d = -0.08$; 95% CI, -0.10 to 0.27) 	
Xie, 2016b ¹⁹	
<p>UPDRS - GPi-DBS versus BMT groups</p> <ul style="list-style-type: none"> • <u>Part I</u> (2 studies): non-significant difference ($P = 0.10$) • <u>Part II</u> on-medication phase (2 studies): non-significant difference between GPi-DBS versus BMT groups ($P = 0.14$, WMD = -3.09; 95% CI, -7.20 to 1.03) off-medication phase (1 study): significant difference in favour of GPi-DBS versus BMT ($P < 0.0001$, WMD = 6.30; 95% CI, -8.18 to -4.42) • <u>Part III</u> on phase (3 studies): significant difference in favour of GPi-DBS versus BMT ($P < 0.0001$, WMD = -4.09; 95% CI, -4.45 to -3.72) off phase (1 study): significant difference in favour of GPi-DBS versus BMT ($P < 0.0001$, WMD = -16.70; 95% CI, -20.41 to -12.99) • <u>Part IV</u> on phase (from 2 studies): significant difference in favour of GPi-DBS versus BMT ($P < 0.0001$, WMD = -3.60; 95% CI, -4.68 to -2.53) <p>UPDRS - STN-DBS versus BMT groups</p> <ul style="list-style-type: none"> • <u>Part I</u> (2 studies): non-significant difference, $P = 0.56$ • <u>Part II</u> on phase (4 studies): significant difference in favour of STN-DBS versus BMT ($P = 0.005$, WMD = -1.50; 95% CI, -2.53 to -0.46); off phase (3 studies): significant difference in favour of STN-DBS versus BMT ($P < 0.0001$, WMD = -9.30; 95% CI, -10.92 to -7.68) • <u>Part III</u> on phase (5 studies): significant difference in favour of 	<p><i>“Based on current available information, either GPi-DBS or STN-DBS plus BMT was superior to BMT alone in terms of reducing UPDRS scores and improving quality of life and decreased medication requirements.” (p 9)¹⁹</i></p>

Table 10: Summary of Findings Included Systematic Reviews and Meta-Analyses

Main Study Findings	Authors' Conclusion
<p>STN-DBS versus BMT ($P = 0.0007$, WMD = - 3.23; 95% CI, - 5.09 to - 1.37); off phase (3 studies): non-significant difference, $P = 0.11$</p> <ul style="list-style-type: none"> Part IV on-phase (1 study): significant difference in favour of STN-DBS versus BMT ($P < 0.001$, WMD = - 3.50; 95% CI, - 5.02 to - 1.98) <p>PDQ-39</p> <ul style="list-style-type: none"> 6 studies compared DBS (STN or GPi) to BMT and found a significant difference in favor of DBS ($P < 0.0001$, WMD = - 7.93; 95% CI, - 8.51 to - 7.35) 3 studies compared GPi-DBS to BMT and found a significant difference in favor of DBS ($P = 0.0002$, WMD = - 5.16; 95% CI, - 7.91 to - 2.41) <p>LED</p> <ul style="list-style-type: none"> 6 studies compared DBS (STN or GPi) with BMT and found a decrease in medication dose in favor of DBS ($P < 0.0001$, WMD = - 417.00; 95% CI, - 565.80 to - 268.20) 	

CI = 95% confidence interval; d = effect size; DBS = deep brain stimulation; GPi = globus pallidus interna; LED = Levodopa equivalent dose; MDRS = Mattis Dementia Rating Scale; MMSE = Mini Mental State Examination; PD = Parkinson's disease; PDQ-39 = Parkinson's disease Questionnaire; RAVLT-total = Rey Auditory Verbal Learning Test-total; RAVLT-delayed recall = Rey Auditory Verbal Learning Test-delayed recall; SE = standard error; STN = subthalamic nucleus; UPDRS = Unified Parkinson's Disease Rating Scale; WMD = weighted mean difference

Table 11: Summary of Findings of Included Clinical Studies

Main Study Findings	Authors' Conclusion
Blomstedt, 2018 ²¹	
<p>Motor symptoms (baseline versus 6-months)</p> <ul style="list-style-type: none"> UPDRS-III The scores were significantly better for surgical patients on-stimulation off-medication (19.5 ± 7.8) compared with medical patients off-medication (37.2 ± 12.2, $P = 0.001$). In the medical group on-medication (21.8 ± 13.4) versus surgical group on-medication on-stimulation (18.5 ± 12.4), no significant differences in UPDRS scores. <p>Health-related quality of life (baseline versus 6-months)</p> <ul style="list-style-type: none"> PDQ-39 Summary index was not different between medical and surgical groups. <p>Changes in tremor, speech, dyskinesia, and axial sub scores (baseline versus 6-months)</p> <ul style="list-style-type: none"> UPDRS-III sub scores Tremor improved ($P = 0.001$) and akinesia ($P = 0.015$) in the surgical group on-stimulation off-medication compared with the medical group off-medication. No differences between the groups in speech, 	<p><i>"In this observer-blinded study on patients with advanced PD randomised to DBS in the cZi or to best medical management, it was shown that DBS in the cZi was efficient. The powerful effect of cZi DBS on tremor was confirmed. There was a modest, although significant, improvement on akinesia but not of the same magnitude as the improvement reported following STN DBS. Future reports will provide data on longer-term follow-up. Also, further studies are needed in order to determine the role of cZi DBS in the therapeutic armamentarium for PD, especially in relation to STN DBS. What can be stated so far is that the cZi as a target for DBS in PD will be an addition to the existing established targets, which may increase the possibility to tailor the surgery to the needs of the individual patients."</i> (p 715)²¹</p>

Table 11: Summary of Findings of Included Clinical Studies

Main Study Findings	Authors' Conclusion
<p>axial symptoms, dyskinesia or in LEDD.</p> <ul style="list-style-type: none"> • <u>UPDRS-IV</u>. No statistically significant changes were seen between the groups in any of the UPDRS IV sub items 36–39. <p>Medication use (baseline versus 6-months)</p> <ul style="list-style-type: none"> • No differences between the groups in LEDD. 	
L'Hommée, 2018 ²²	
<p>Behaviour (baseline versus 24 months)</p> <ul style="list-style-type: none"> • <u>Ardouin Scale of Behavior in PD</u>. Neuropsychiatric fluctuations subscale of the Ardouin scale differed significantly between treatment groups, favouring bilateral STN-DBS plus BMT (mean change = -0.65 [SE = 0.15]) versus BMT alone (-0.02 [0.15] points; $P = 0.0028$) • Hyperdopaminergic behaviours score of the Ardouin scale decreased significantly with bilateral STN-DBS plus BMT but increased significantly with BMT alone (mean change -1.26 [0.35] points versus 1.12 [0.35] points; $P < 0.0001$). • Nocturnal hyperactivity ($P < 0.0001$), diurnal somnolence ($P = 0.0294$), creativity ($P = 0.0212$), and hobbyism ($P = 0.0017$) increased with BMT alone and decreased under STN-DBS plus BMT. Excess in motivation grew over time with both treatments but to a greater extent with BMT alone ($P = 0.0313$) • <u>Psychiatric serious events</u> were descriptively described. STN-DBS = 19 events experienced in 17 patients (12%); BMT group = 31 events including 1 suicide in 23 patients (17%). Two suicides reported in STN-DBS group and one in the BMT group. <p>Apathy (baseline versus 24 months)</p> <ul style="list-style-type: none"> • <u>Starkstein Apathy Scale</u> no significant differences between groups <p>Depression (baseline versus 24 months)</p> <ul style="list-style-type: none"> • <u>Beck Depression Inventory</u> no significant differences between groups <p>Medication use (baseline versus 24 months)</p> <ul style="list-style-type: none"> • <u>LEDD</u> Significant decrease in LEDD by 39% (mean change -363.3 mg/day, SE = 41.8) in STN-DBS group versus an increase by 21% (mean change 245.8 mg/day, SE 40.4) in BMT group ($P < 0.0001$) • <u>Antidepressants</u> were stopped in 12 patients in STN-DBS group plus BMT versus 4 patients in BMT alone group • <u>Neuroleptics</u> were started in 9 patients assigned BMT alone versus 1 patient STN-DBS plus BMT. 	<p><i>“In individuals with PD and early motor complications of levodopa treatment, we found that subthalamic stimulation led to stabilisation of both motor and non-motor complications of dopaminergic treatment. Controlled data using the Ardouin scale show that applying subthalamic stimulation in this population permits relative safety with respect to postoperative apathy, anxiety, and depression; a clear cut improvement of disabling neuropsychiatric fluctuations; and easier long-term management of behavioural complications of dopamine replacement therapy compared with dopaminergic medications alone.” (p 230)²²</i></p>

Table 11: Summary of Findings of Included Clinical Studies

Main Study Findings	Authors' Conclusion
Millan, 2017 ²⁴	
<p>BMI</p> <ul style="list-style-type: none"> Between-group difference in change in BMI score at 24 months was not significant ($P = 0.62$). 	<p><i>“These results suggest that STN-DBS does not cause increased BMI in early stage PD. More study is needed to confirm these findings and the FDA has approved a phase III multicenter, randomized, double-blind, placebo-controlled, pivotal clinical trial evaluating DBS in early stage PD” (p 3)²⁴</i></p>
Hacker 2015 ²³	
<p>UPDRS-III and IV (comparing baseline to 24 months)</p> <ul style="list-style-type: none"> DBS + ODT group improved versus ODT group at each time point on Total UPDRS ($P = 0.04$) and Part III ($P = 0.02$). Using a composite score (including components of the UPDRS III and IV), 54% of the ODT group and 27% of the DBS + ODT group experienced clinically important worsening; a 50% relative risk reduction for the DBS + ODT group versus ODT group ($P = 0.25$). When applied to the focused cohort of participants in the DBS in early PD pilot trial on medication 1-4 years at enrollment = 80% reduction in the risk of clinically important worsening experienced by the DBS + ODT group versus ODT group ($P = 0.07$). 	<p><i>“These results suggest that DBS in early PD reduces the relative risk of clinically important worsening and is superior to medication alone for treating motor symptoms, reducing the complications of BMT, and providing better quality of life. Additional study is needed to confirm these findings, and the FDA has approved this focused inclusion criteria and patient-centered outcome in a prospective, pivotal, phase III, multicenter, double-blind, placebo-controlled clinical trial designed to definitively discern whether DBS is superior to medication alone in early PD.” (p 1182)²³</i></p>
Tramontana, 2015 ²⁵	
<p>All outcomes comparing DBS+ODT (intervention) to ODT only (control). Given the number of outcomes, significant differences are reported with their associated p-values and non-significant values are reported as “NS.”</p> <p>Manual dexterity</p> <ul style="list-style-type: none"> <u>Purdue Pegboard</u> <ul style="list-style-type: none"> 12 months: NS 24 months: NS <p>Visual-spatial orientation (non-motor)</p> <ul style="list-style-type: none"> <u>Benton Judgment of Line Orientation Test</u> <ul style="list-style-type: none"> 12 months: NS 24 months: NS <p>Confrontation naming</p> <ul style="list-style-type: none"> <u>Boston Naming Test</u> <ul style="list-style-type: none"> 12 months: NS 24 months: NS <p>Rapid word production</p> <ul style="list-style-type: none"> <u>Verbal fluency (phonemic, semantic)</u> <ul style="list-style-type: none"> 12 months: <i>Phonemic fluency</i> ODT (0.9 ± 2.9) performed better than DBS+ODT (-1.1 ± 2.7; $P = 0.047$); <i>Semantic fluency</i> NS 24 months: NS <p>Attention/working memory</p> <ul style="list-style-type: none"> <u>WAIS-III Digit Span</u> <ul style="list-style-type: none"> 12 months: ODT (0.6 ± 1.0) performed better than DBS+ODT (-1.3 ± 2.5); $P = 0.004$ 	<p><i>“The results of this trial provide novel data regarding the effects of DBS on cognitive function in early PD. We believe that the findings and insights from this trial can help guide the safety analysis and risk-benefit evaluations in future discussions of DBS in early stage PD.” (p 151)²⁵</i></p> <p><i>“This study yielded initial safety and tolerability data of DBS in early stage PD with a focus on cognitive and neurobehavioral outcomes. AE rates in this trial were similar to what was reported in a large trial of DBS in subjects with advanced PD. Of the 15 subjects in our DBS+ODT group, two of them experienced serious AEs that resulted in some persistent cognitive changes. There were few differences between the two groups at 12 months and they were further diminished at 24 months, especially when data from these two subjects were removed. Nevertheless, the AEs reported in this trial must be factored into future discussions regarding the risk-benefit of offering DBS to subjects with early PD.” (p 160)²⁵</i></p> <p><i>“Important questions remain regarding the relative merits of applying STN DBS in early stage PD. The potential benefits of STN DBS in early PD are discussed in the primary report on this trial, which reports motor and quality-of-life outcomes [10]. Future studies are necessary to determine whether any advantages derived from the early use of STN DBS, especially with patients still responsive to standard medical treatment, outweigh any added risks that may be involved. We believe that</i></p>

Table 11: Summary of Findings of Included Clinical Studies

Main Study Findings	Authors' Conclusion
<ul style="list-style-type: none"> ○ 24 months: NS Attention/working memory/processing speed • <u>Paced Auditory Serial Addition Test</u> <ul style="list-style-type: none"> ○ 12 months: ODT performed better than DBS+ODT for slowest paced rates ($P = 0.009$) and fastest paced rates ($P = 0.005$) ○ 24 months: NS Immediate and delayed verbal memory • <u>WMS-III Word List Learning I & II</u> <ul style="list-style-type: none"> ○ 12 months: NS ○ 24 months: NS Immediate and delayed visual memory (non-motor) • <u>WMS-III Memory for Faces I & II</u> <ul style="list-style-type: none"> ○ 12 months: NS ○ 24 months: NS Executive functioning/flexible problem-solving • <u>Wisconsin Card Sorting Test</u> <ul style="list-style-type: none"> ○ 12 months: ODT (1.5 ± 2.1) performed better than DBS+ODT (0.3 ± 1.8) for perseverative errors ($P = 0.053$) ○ 24 months: ODT performed better than DBS+ODT for categories achieved (0.4 ± 1.1 versus -0.3 ± 1.0; $P = 0.053$) perseverative errors (1.7 ± 2.0 versus 0.1 ± 2.3; $P = 0.051$) Processing speed/control of competing responses • <u>Stroop Color and Word Test</u> <ul style="list-style-type: none"> ○ 12 months: ODT (0.4 ± 2.1) performed better than DBS (-2.6 ± 2.1; $P = 0.035$) ○ 24 months: NS Emotional status and personality • <u>Minnesota Multiphasic Personality Inventory-2</u> <ul style="list-style-type: none"> ○ 12 months: NS ○ 24 months: NS 	<p><i>these initial insights can help guide and inform discussions regarding future trials investigating the safety and efficacy of DBS in early stage PD.” (p 161)²⁵</i></p>

AE = adverse event; BMI = body mass index; cZi = caudal zona incerta; DBS = deep brain stimulation; FDA = Food and Drug Administration; LEDD = levodopa-equivalent daily doses; NS = non-significant; ODT = optimal drug therapy; PD = Parkinson's disease; PDQ-39 = Parkinson's disease Questionnaire; SE = standard error; STN = subthalamic nucleus; UPDRS = Unified Parkinson's Disease Rating Scale; WAIS-III = Wechsler Adult Intelligence Scale – Third Edition; WMS-III = Wechsler Memory Scale – Third Edition

Table 12: Summary of Findings of Included Economic Evaluations

Main Study Findings	Authors' Conclusion
Dams, 2016 ²⁶	
<p>Life-long time horizon</p> <p>Mean discounted direct costs for patients with early PD</p> <ul style="list-style-type: none"> • STN-DBS: 151,800 EUR/patient • BMT group: 115,400 EUR/patient <p>Discounted quality-adjusted life expectancy at age of 52 years</p> <ul style="list-style-type: none"> • STN-DBS group: 13.84 QALYs • BMT group: 12.25 QALYs <p>ICUR = 22,710 EUR/QALY for STN DBS compared with BMT</p>	<p><i>“Based on our decision-analytic cost-utility analysis, STN DBS should be considered to be cost-effective in younger patients with earlier stages of PD (H & Y I– III). An ICUR of 22,700 EUR/QALY was assessed for a life-long time horizon comparing STN DBS with BMT. The sensitivity analyses showed a major effect of time to battery exchange. Including values for the recently introduced rechargeable device system resulted in an ICUR of 37,600 EUR/QALY. Nevertheless, this system has not yet been used for PD patients with early motor complications</i></p>

Table 12: Summary of Findings of Included Economic Evaluations

Main Study Findings	Authors' Conclusion
<p>PDQ-39 summary index (effectiveness measurement)</p> <ul style="list-style-type: none"> STN-DBS group: 1,270 points BMT group: 1,670 points <p>ICER = 89 EUR per PDQ-39 summary index point gained for patients treated with STN DBS versus BMT.</p> <p><u>EQ-5D index estimate</u></p> <ul style="list-style-type: none"> STN-DBS group: average EQ-5D index value of 0.79 at baseline, 0.88 at the 12-month follow-up, and 0.86 at the 24-month follow-up. BMT group = average EQ-5D index value of 0.80 at baseline, 0.82 at the 12-month follow-up, and 0.79 at the 24-month follow-up <p>Sensitivity analyses</p> <ul style="list-style-type: none"> One-way sensitivity analyses ICURs range: 7,500 – 45,700 EUR/QALY (life-long time horizon) The outcome was most sensitive to battery exchange, but never exceeded the threshold of 50,000 EUR/QALY Costs for battery exchange and surgery, drug costs, and variations in improvement in motor complications resulted in changes between 15,000 and 30,000 EUR/QALY. Changes less than 8,000 EUR/QALY were observed for costs of adverse events, improvement in HrQoL, severity of disease, age, and mortality. No changes were observed for uncertainties in transition probabilities Varying assumptions about the model structure resulted in minor changes 	<p><i>treated with STNDBS. Therefore, data concerning the effects of the rechargeable device system are required. In addition, reliable data for the calculation of QALYs using different multi-attribute instruments are urgently needed to provide trustworthy information on the denominator for an appropriate health economic evaluation of DBS in the future.” (p 1190)²⁶</i></p>
<p>Fundament, 2016²⁸</p>	
<p>15-year time horizon</p> <p>Discounted model results using deterministic model</p> <ul style="list-style-type: none"> DBS: £73,077/patient; mean QALYs gained/patient = 6.69 BMT group: £46,278/patient; mean QALYs gained/patient = 5.35 <p>ICER = £19,887/QALY gained for DBS versus BMT</p> <p>Sensitivity analyses</p> <ul style="list-style-type: none"> One-way sensitivity analyses suggests results were not significantly impacted by plausible changes in the input parameter values; no scenario was the computed ICER for the comparison of DBS versus BMT in excess of £30,000 per QALY gained Probabilistic sensitivity analyses suggests a 99% probability of DBS being cost-effective at a UK maximum willingness-to-pay threshold of £30,000/QALY 	<p><i>“In conclusion, we modelled cost-effectiveness of DBS in PD patients with early motor complications over a time horizon of 15 years building on evidence from patients with advanced disease. Using a novel mapping algorithm to link disease progression with health-related quality of life, and two-year follow-up data from the EARLYSTIM trial [36], our analysis concludes that the incremental cost-effectiveness ratio of DBS compared to BMT is acceptable based on current thresholds. Our findings are line with the evidence for cost-effectiveness of DBS in advanced PD and provide support from an economical standpoint for the extension of existing policy recommendations to make DBS available to patients with PD with early motor complications.” (p 12)²⁸</i></p>

Table 12: Summary of Findings of Included Economic Evaluations

Main Study Findings	Authors' Conclusion
Hacker, 2016 ²⁹	
<p>Medication costs (baseline to 24 months)</p> <ul style="list-style-type: none"> DBS+ODT group: decreased 16% (from 13.39 ± \$5.98/day to \$11.31 ± \$7.80/day) ODT group: Increased 72% (from \$13.81 ± \$6.98/day to \$23.82 ± \$27.95/day) <p>This cost difference translates into a cumulative savings for the DBS+ODT group of \$7,150 over the study period (i.e. 24 months; \$14,662 for ODT versus \$7,512 for DBS+ODT).</p> <p>Polypharmacy (baseline to 24 months)</p> <ul style="list-style-type: none"> DBS+ODT participants were 80% less likely to require polypharmacy compared with ODT participants at 24 months (p < 0.05; OR= 0.2; 95% CI: 0.04–0.97) <p>10-year projected medication cost savings</p> <ul style="list-style-type: none"> \$64,590 in favour of DBS+ODT group 	<p><i>“STN-DBS in early PD reduced medication cost over the two-year study period. DBS may offer substantial long-term reduction in medication cost by maintaining a simplified, low dose medication regimen. Further study is needed to confirm these findings, and the FDA has approved a pivotal, multicenter clinical trial evaluating STN-DBS in early PD.”</i> (p 125)²⁹</p> <p><i>“While the cost of DBS therapy is significant, STN-DBS applied in early stage PD may impart a substantial reduction in medication utilization and cost over many more years than when DBS is applied in current standard of care. Even though these savings do not exceed the cost of DBS therapy, they do represent a substantial offset when considered over the course of the disease. Additional investigation into the safety and efficacy of DBS in early stage PD is indicated, and the FDA has approved a pivotal pivotal, multicenter, phase III, clinical trial evaluating STN-DBS in early stage PD.”</i> (p 130)²⁹</p>
McIntosh, 2016 ²⁷	
<p>Year 1</p> <ul style="list-style-type: none"> Number of antiparkinsonian drugs was lower for the DBS group versus BMT group (P < 0.000) Costs higher for DBS group (£19,069, SD = £8,364) versus BMT group (£9,813, SD = £7,479; 95% CI, £7,625 to £10,887) Mean utility gain for DBS group (0.06, 95% CI: 0.003 to 0.12) QALY insignificant gain for DBS group (0.02, 95% CI: -0.015 to 0.05) ICER = £468,528 Sensitivity analyses performed found that DBS surgery in this patient group has a very low probability of being cost-effective at 1 year. <p>5-year projection</p> <ul style="list-style-type: none"> ICER = £45,180 willingness-to-pay threshold for a QALY of £20,000 to £30,000 (based on NICE guidelines) One-way sensitivity analysis: for DBS to be cost-effective at 5 years requires any of the following: a 10-year life span for the original implantable pulse generator and electrodes; surgery QALY gains increase by around 30%; and annual follow-up costs in the BMT arm to increase by around 30% or more. Reducing annual health and personal social services follow-up costs in the DBS group by 50% brings the ICER into the higher range of being cost-effective. Two-way sensitivity analysis: for DBS group to be cost-effective at 5 years requires a 10-year life span for the original implantable pulse generator and electrodes and 	<p><i>“In this patient group, DBS is not cost-effective at 1 year. Extrapolation, however, reveals an increasing likelihood of cost-effectiveness up to 5 years and reducing cost-effectiveness between 5 and 10 years. These models are sensitive to assumptions about future costs and quality-adjusted life years gained.”</i> (p 1173)²⁷</p>

Table 12: Summary of Findings of Included Economic Evaluations

Main Study Findings	Authors' Conclusion
<p>QALY gains in the surgery arm to increase by 10%. One cost-effective scenario at 5 years (ICER of £4,126) = original implantable pulse generator and electrodes having a life span of 10 years and drug costs in the BMT arm increasing by 30%.</p> <p>10-year projection</p> <ul style="list-style-type: none"> • ICER = £70,537; increase versus 5-years sensitive to increased probability of battery replacements (and re-replacements) beyond 5 years and assumptions about changes in QALY gain in the surgery group • One-way sensitivity analysis DBS is likely to be highly cost-effective at 10 years if the IPG and electrodes have a 10-year life span or QALY gains in the DBS group increase by around 30%. • Two-way sensitivity analysis presents reveals a large number of scenarios for surgery to be highly cost-effective at 10 years, including a dominant scenario combining a 10-year IPG life span and increasing drug costs in the BMT group by 30%. 	

BMT = best medical therapy; DBS = deep brain stimulation; EQ-5D = EuroQoL Index; EUR = Euro (currency); H & Y = Hoehn and Yahr scale; HrQoL = health-related quality of life; ICER = incremental cost-effectiveness ratio; ICUR = incremental cost-utility ratio; IPG = implantable pulse generator; OR = odds ratio; PD = Parkinson's disease; PDQ-39 = Parkinson's disease Questionnaire; SD = standard deviation; QALY = quality-adjusted life=years; STN = subthalamic nucleus

Table 13: Summary of Recommendations in Included Guideline

Recommendations	Strength of Evidence and Recommendations
NICE 2017 ³⁰	
<ul style="list-style-type: none"> • <i>“Offer people with advanced PD BMT, which may include intermittent apomorphine injection and/or continuous subcutaneous apomorphine infusion.</i> • <i>Do not offer DBS to people with PD whose symptoms are adequately controlled by BMT.</i> • <i>Consider DBS for people with advanced PD whose symptoms are not adequately controlled by BMT” (p 215)</i> 	<p>Adverse events</p> <ul style="list-style-type: none"> • Strength of evidence: very low to moderate <p>Symptom severity</p> <ul style="list-style-type: none"> • Strength of evidence: low to moderate <p>Neuropsychological outcomes</p> <ul style="list-style-type: none"> • Strength of evidence: very low to moderate <p>Health related quality of life – patient</p> <ul style="list-style-type: none"> • Strength of evidence: moderate to high <p>Medication load</p> <ul style="list-style-type: none"> • Strength of evidence: moderate

BMT = best medical therapy; NICE = National Institute for Health and Care Excellence; PD = Parkinson's disease

Appendix 5: Overlap of relevant primary studies between Included Systematic Reviews

Table 14: Primary Study Overlap between Included Systematic Reviews

Primary Study Citation	Systematic Review Citation		
	Wyman-Chick 2016 ¹⁸	Xie 2016a ²⁰	Xie 2016b ¹⁹
Adrian 2010			X
Angelo 2011			X
Castelli 2010	X	X	
Cilia 2007		X	
Daniels 2010		X	
Ehlen 2013	X		
Frances 2009			X
Gunther 2006			X
Marshall 2012	X		
Merola 2014		X	
Michael 2012			X
Morrison 2004	X		
Rinehardt 2010	X		
Rothlind 2015		X	
Schubpach 2007			X
Schubpach 2013			X
Smeding 2006	X	X	
Weaver 2009	X		
Williams 2011	X	X	
Witt 2008	X		X
Witt 2013		X	
York 2008		X	
Zangaglia 2009	X	X	

Appendix 6: Additional References of Potential Interest

A systematic review and guideline comparing subthalamic nucleus deep brain stimulation to globus pallidus internus deep brain stimulation for the treatment of patients with PD

Rughani A, Schwalb JM, Sidiropoulos C, et al. Congress of Neurological Surgeons systematic review and evidence-based guideline on subthalamic nucleus and globus pallidus internus deep brain stimulation for the treatment of patients with Parkinson's disease: executive summary. *Neurosurgery*. 2018;82(6):753-756..

A systematic review of hardware-related complications of DBS

Jitkrisadukul O, Bhidayasiri R, Kalia SK, Hodaie M, Lozano AM, Fasano A. Systematic review of hardware-related complications of deep brain stimulation: do new indications pose an increased risk? *Brain Stimul*. 2017 Sep 1;10(5):967-76.

A randomized controlled trial that was the primary trial for three of our included secondary analyses

Charles D, Konrad PE, Neimat JS, et al. Subthalamic nucleus deep brain stimulation in early stage Parkinson's disease. *Parkinsonism Relat Disord*. 2014;20(7):731-737.

A cost-minimization analysis comparing DBS-related costs for rechargeable and non-rechargeable devices

Akazawa M, Konomura K, Shiroiwa T. Cost-minimization analysis of deep-brain stimulation using national database of Japanese health insurance claims. *Neuromodulation*. 2018;21(6):548-552.

Guidelines with Unclear Methodology

Anderson DG, Van Coller R, Carr J. South African guideline on deep brain stimulation for Parkinson's disease. *S Afr Med J*. 2017;107(10):1027-1032.

Fox SH, Katzenschlager R, Lim SY, et al. International Parkinson and movement disorder society evidence-based medicine review: update on treatments for the motor symptoms of Parkinson's disease. *Mov Disord*. 2018;33(8):1248-1266.

A summary of a Canadian DBS conference

Panisset M, Picillo M, Jodoin N, et al. Establishing a standard of care for deep brain stimulation centers in Canada. *Can J Neurol Sci*. 2017;44(2):132-138.