CLINICAL ENDPOINTS REVIEW SUMMARY: With table edits corresponding to review by three external experts

Devices to Manage Tremor in Parkinson's Disease and Essential Tremor

Contents

Overview

Objectives: 1) To identify the most common clinical endpoints among patients who use medical devices for the management/treatment of medication-refractory Parkinson's disease (PD) and Essential Tremor (ET), focusing on tremor/motor aspects. 2) To identify the commonly used instruments/tools to assess those outcomes and clarify clinically meaningful differences for each instrument/tool where possible.

Literature considered: 1) Peer-reviewed systematic reviews of clinical studies of devices for patients with Parkinson's disease PD or ET and two professional guidance documents, i.e., clinical practice guidelines and consensus statements.

Important clinical endpoints identified:
Table A1 in this document lists the most important clinical endpoints according to a prioritization process based on the frequency with which a clinical endpoint was measured in the primary studies included in the selected systematic reviews (see the table footnotes). A complete list of the clinical endpoints identified by the systematic reviews appears in Table A2 in this document.

Table B lists the clinical endpoints emphasized in selected professional guidance documents pertaining to PD. This set of clinical endpoints overlaps considerably with the prioritized set derived from systematic reviews and presented in Table A1. However, behavioral and mood disorders, motor complications, and general safety endpoints, all emphasized by three or four of the five professional guidance documents, do not appear in the systematic review-based prioritized list. Clinical practice guidelines published jointly by the European Academy of Neurology and the European section of the Movement Disorders Society considered mood and behavioral disorders to be "important" clinical endpoints, motor complications to be "critical", and safety endpoints to be "critical". The same guidelines also judged all the types of endpoints that do appear in Table A1 to be either "important" or "critical". No relevant professional guidance documents were identified for ET.

Devices represented by the systematic reviews: deep brain stimulation (DBS) (the most frequently covered by systematic reviews), robot-assisted gait training, repetitive transcranial magnetic stimulation (rTMS), wearable devices/sensors, subthalamotomy, magnetic resonance-guided focused ultrasound (MRgFUS).

Measurement instruments: The CER's list of prioritized clinical endpoints (Table A1) is generally measured according to formal, validated scales, and the CER identified published definitions of minimal clinically important

differences (MCIDs) for the scales most frequently used to measure the prioritized clinical endpoints (Table C). Full descriptions of the prioritized measurement tools and associated validation studies are in Appendix Table C1 of the full CER.

Information unavailable from the CER: Typical follow-up duration specific to each clinical endpoint was not provided in the CER. Nor did the CER analyze the frequency of clinical endpoints according to device type or patient clinical history.

Applicability to the Medicare beneficiary population: None of the systematic reviews or professional guidance documents related to PD focused on older adults, but the mean age of study groups ranged from 54 to 77 years across the systematic reviews. The reviews described most study groups as representing a mix of disease severity according to the Hoehn and Yahr scale. The reviews did not provide sufficient information to allow assessment of the generalizability of the CER's findings concerning patient demographics.

Table A1. Prioritized Clinical Endpoints Identified in Peer-Reviewed Systematic Reviews (see Table 2 in the full CER)

Green-shaded boxes reflect feedback from one or more of three experts who reviewed the completed CER. Instruments in these boxes were taken from Appendix Table B2 of the CER or were recommended by a reviewer.

Domain	Clinical endpoints (citation volume)*	Instruments (citation volume)†
PARKINSON'S DISEASE		
Clinician-Assessed Health Outcomes	Global change in motor symptom severity (65%)	UPDRS III (97.8%)
	Gait function, postural instability, balance, falls (40%)*	FOG-Q (60.3%) and TUG (60.3%)
		Relevant items from MDS-UPDRS III and MDS-UPDRS II
	Motor complications (particularly important for disease duration >5-7 years)	UPDRS or MDS-UPDRS IV
	Reduction in dyskinesia	Relevant items from MDS-UPDRS Part IV
	Cognitive function (42%) (particularly important in adults aged ≥65 years because of greater risk)	Stroop test (85.7%)
		MoCA instead of or in addition to Stroop test. MDS-UPDRS I (relevant subscale)
	1 reviewer pointed out that cognitive decline and mood changes can be adversely affected by some treatments. Only 1 systematic review in the CER covered mood changes.	
	Motor-related ADL (48%)	UPDRS II or MDS-UPDRS II (47.5%)
	Nonmotor symptoms	MDS-UPDRS-II, MDS-NMSS
	Speech function and swallowing	Relevant items from MDS-UPDRS II and MDS-NMS, Speech Intelligibility Test
Patient-Reported Outcomes	PD-related QoL (53%)	PDQ-39 (43.6%)
	Sleep quality (10%)	PDSS-2, relevant items from MDS-UPDRS I
Device-Related Safety	Not addressed in the CER. Reviewers commented on variability across devices. More serious: 1) dysarthria or ataxia following MRgFUS or DBS; risk of intracranial hemorrhage, infection, or need for additional surgery following DBS surgery; speech/swallowing difficulties, mood (including suicidality); and cognition decline after DBS, MRgFU or subthalamotomy. Less serious: discomfort in implementation (for example, skin issues, or discomfort from wrists and other non-invasive wearables for tremor)	
ESSENTIAL TREMOR		
Clinician-Assessed Health Outcomes	Reduction in tremor (89%)	FTM-TRS (66.1), Hand Tremor score (30.8), CRST (15.7), and TETRAS (2.3%)‡
Patient-Reported Outcomes	Tremor-related QoL OR tremor-related ADL (44%)	QUEST (42.2%) and FTM-TRS Part C (24.4%)
One reviewer cautioned that the FTM-TRS is not commonly used clinically.		
Device-Related Safety	Same as those listed for PD	

Abbreviations: ADL: Activities of daily living; CRST: Clinical Rating Scale for Tremor; FOG-Q: Freezing of Gait Questionnaire; FTM-TRS: Fahn Tolosa Marin Tremor Rating Scale; MDS-NMSS: MDS Non-Motor Symptom Scale; MDS-UPDRS: Movement Disorder Society–Sponsored Revision of the UPDRS; MoCA: Montreal Cognitive Assessment; PDQ-39: Parkinson Disease's Questionnaire-39; PDSS-2: Parkinson's Disease Sleep Scale 2nd version; QoL: Quality of life; QUEST: Quality of Life Essential Tremor Questionnaire; TETRAS: Tremor Research Group Essential Tremor Rating Assessment scale; TUG test: Time Up and Go Test; UPDRS: Unified Parkinson's Disease Rating Scale

*A clinical endpoint was considered prioritized if it was cited by ≥40% of the studies included in the systematic reviews that discussed that endpoint. For example, the systematic reviews that discussed change in motor severity collectively included 278 primary studies and 180 of those studies (64.78%) investigated change in motor severity; thus, motor severity was prioritized.

†Citation volume was calculated in a manner analogous to that used for clinical endpoint citation volume.

‡Due to high heterogeneity in the instruments used for assessing prioritized clinical endpoints, instruments for some prioritized endpoints did not reach the cut-off value of ≥40%.

Table A2. All Clinical Endpoints Investigated in the Primary Studies Included in Systematic Reviews (Derived from Appendix Table B2 in the CER)

Clinical Endpoint	Citation Volume
PARKINSON'S DISEASE	
Prioritized	
Global change in motor severity	65%
QoL	53%
PD-related ADL	48%
Cognitive function	42%
Gait function, postural instability and balance	40%
Did not meet prioritization threshold	
PD medication usage	31%
Dyskinesia	29%
Ataxia	10%
Sleep quality	10%
Nonmotor symptoms	16%
Physical function (e.g., balance confidence)	10%
Fatigue	5%
Fear of falling	5%
ESSENTIAL TREMOR	
Reduction of tremor	89%
Impact of motor symptoms on ADL and QoL	44%

Table B. Clinical Endpoints Derived from Professional Guidance Documents on Parkinson's Disease (Table 3 in the CER)

Clinical Endpoints	**Fox et al., (2018) (IPMDS).** Updates to evidence-based int'l CPG, broad range of therapies	**Odin et al., (2018) (IPMDS).** Expert panel recommendations regarding how to assess use of wearable sensor devices	**Rughani et al., (2018) (CNS/ASSFN).** Evidence-based CPG on comparative effectiveness of DBS to the STN vs. GPi	**Deuschl et al., (2022) (EAN and European section of the MDS).** Evidence-based CPG on invasive therapies for advanced PD, including invasive administration of pharmaceuticals	**Saba et al., (2022) (BAN).** Evidence-based Brazilian CPG, pharmaceutical and DBS treatment of motor symptoms
Clinician-Assessed Health Outcomes					
Disease severity				Staging according to Hoehn and Yahr ("important")	
Global measures of motor symptom severity	*Discussed without reference to measurement method except for a single reference to UPDRS endpoints in a pharmaceutical treatment*	UPDRS *(cited as a scale for purposes of assessing correlation with sensor measurements)*	• UPDRS III • Reduction in dopaminergic medications	• UPDRS III ("critical") • Daily dosage of anti-PD medication (levodopa mg equivalent) ("important")	• UPDRS III
Specific aspects of motor function	*Discussed without reference to measurement method; gait and balance are mentioned*	• AIMS *(cited as a scale for purposes of assessing correlation with sensor measurements)* • Measured by sensor: Bradykinesia, tremor, dyskinesia, immobility during sleep	*Measurement methods not specified*	• Gait ("important") • Speech ("important") *No tools recommended, but UPDRS III includes gait and speech scales.*	
Reduction in or delay of motor complications*	*Discussed without reference to measurement method*		*Discussed without reference to measurement method*	• According to UPDRS IV ("critical") • Motor fluctuations reported by patient diary ("critical")	*Both reduction and delay are mentioned.*
Cognitive function				UPDRS-I, MoCA, MMSE ("important")	
Mood and behavioral disorders		Risk marker for impulsivity measured by sensor (NOTE: This actually belongs in the Surrogate Marker domain.)	Depression (HDRS) Suicidality (UPDRS-I)	• HDRS ("important") • Apathy ("important") (No tool recommended, but UPDRS III includes an apathy scale)	

Clinical Endpoints	**Fox et al., (2018) (IPMDS).** Updates to evidence-based int'l CPG, broad range of therapies	**Odin et al., (2018) (IPMDS).** Expert panel recommendations regarding how to assess use of wearable sensor devices	**Rughani et al., (2018) (CNS/ASSFN).** Evidence-based CPG on comparative effectiveness of DBS to the STN vs. GPi	**Deuschl et al., (2022) (EAN and European section of the MDS).** Evidence-based CPG on invasive therapies for advanced PD, including invasive administration of pharmaceuticals	**Saba et al., (2022) (BAN).** Evidence-based Brazilian CPG, pharmaceutical and DBS treatment of motor symptoms
				• Ardouin-scale, QUIP, QUIP-RS, MIDI, BIS ("important")	
Composite of mood, cognitive, and behavioral effects			*Measurement methods not specified*		
ADL			UPDRS-II	UPDRS-II ("critical")	UPDRS-II
Patient-Reported Health Outcomes					
QoL		*Assumed to improve under continuous assessment*	PDQ-39	PDQ-39 or PDQ-8 ("critical")	
Mood and behavioral disorders			Depression (BDI)	Depression (BDI) ("important")	
Other	*Reduction in falls (mentioned in the context of pharmaceutical treatment and physical therapy)*				
Safety (clinician-assessed or patient-reported)					
Harmful effects	• "Acceptable risk with/without specialized monitoring" *(no identification of specific risks to be considered)* • Unacceptable risk		Adverse events	Serious adverse medical and surgical events ("critical")	Discontinuation of treatment (tolerability)
Other					
					• Duration of response • Discontinuation of treatment (lack of efficacy)

Clinical Endpoints	**Fox et al., (2018) (IPMDS).** Updates to evidence-based int'l CPG, broad range of therapies	**Odin et al., (2018) (IPMDS).** Expert panel recommendations regarding how to assess use of wearable sensor devices	**Rughani et al., (2018) (CNS/ASSFN).** Evidence-based CPG on comparative effectiveness of DBS to the STN vs. GPi	**Deuschl et al., (2022) (EAN and European section of the MDS).** Evidence-based CPG on invasive therapies for advanced PD, including invasive administration of pharmaceuticals	**Saba et al., (2022) (BAN).** Evidence-based Brazilian CPG, pharmaceutical and DBS treatment of motor symptoms

Abbreviations: ADL: Activities of Daily Living; ASSFN: American Society for Stereotactic and Functional Neurosurgery; BAN: Brazilian Academy of Neurology; BIS: Barrett Impulsivity Scale; CNS: Congress of Neurological Surgeons; CPG: clinical practice guidelines; DBS: deep brain stimulation; EAN: European Academy of Neurology; HDRS, Hamilton Depression Rating Scale; MDS: Movement Disorder Society; IPMDS: International Parkinson and Movement Disorder Society; MIDI: Minnesota Impulsivity-Compulsivity Disorders Interview; MoCA: Montreal Cognitive Assessment; MMSE: Mini-Mental State Examination; PD: Parkinson's disease; PDQ: Parkinson's Disease Questionnaire; QoL: quality of life; QUIP(-RS): Questions for Impulsivity-Compulsivity Disorders (-Rating Scale); STN: subthalamic nucleus

*Motor complications are symptoms that occur after chronic treatment of PD. Key examples are motor fluctuations due to the wearing off of treatment effectiveness and levodopa-induced dyskinesia (Freitas et al., 2017).

REVIEW QUESTIONS

After reading the full CER and Clinical Endpoints Summary, please be prepared to discuss the following questions in the subcommittee meeting.

1. Do you believe that the search strategies used in the CER are likely to have captured the appropriate literature? (See the "Identifying the Literature" section, starting at the bottom of page 2 of the full CER).

2. Are the most important domains and clinical endpoints for studies of PD and ET therapies reflected in the CER findings? (See Tables A1 and B in this document). Are important domains or clinical endpoints missing, and are omissions likely due to deficiencies in the literature search strategy or omissions in the available body of clinical research literature? Please suggest additions and/or deletions to each domain with commentary as appropriate:

    a. Surrogate markers, i.e., indirect assessments using biomarkers, physiologic measures, or imaging and intended to predict or act as a proxy for a target outcome of interest
    b. Clinician-assessed health outcomes
    c. Patient-reported outcomes
    d. Safety *(Your opinion on the appropriateness of assigning motor complications to this domain would be appreciated. Motor complications have been assigned to the safety domain because of the potential for tremor-controlled devices to beneficially reduce the motor complications that accompany pharmaceutical treatment.)*
    e. Other *(Respond here if you wish to recommend any important clinical endpoints that do not fit one of the preceding domains.)*

3. Patient- and technology-specific clinical endpoints

    a. Is there any difference in the clinical endpoints that should be measured in subpopulations of patients with PD or ET versus general adult populations with those diseases? Examples of subpopulations include older (age ≥ 65 years) versus younger adults, more versus less severe disease, women versus men, and patients with versus without high levels of comorbidity?

    b. Is there any difference in important clinical endpoints, particularly safety endpoints, according to type of technology? Examples: invasive versus noninvasive.

4. Measurement instruments

a. Consider the instruments and measurement methods listed in Tables A1 and B in this document. Are these appropriate for studies investigating devices to manage PD- and ET-related tremor? Are there others that should be considered?
b. Consider any additional endpoints you identified in response to question #2. What instruments or methods would you recommend for measuring these endpoints?

5. MCIDs

    a. Do you know of any published definitions of MCID, other than those referenced in Appendix Table C1 of the full CER, for the clinical endpoints that are in that table?
    b. Can you offer MCID definitions for any clinical endpoints you identified in response to Question #2?
6. Based on your clinical expertise, what is the domain-specific minimum duration of follow-up that should be observed in studies of devices to manage PD- and ET-related tremor? If appropriate follow-up intervals vary within domains according to various endpoints, device type, or patient characteristics, please provide that information.

    Clinician-assessed health outcomes:

    Patient-reported outcomes:

    Clinician-assessed or patient-report safety outcomes:

    Patient adherence:

    Device durability:

    Other:

Table C. MCID Definitions for Prioritized Clinical Endpoints (derived originally from Appendix Table C1 of the full CER)

Green-shaded boxes show additions recommended by 1 or more of 3 experts who reviewed the full CER.

Teal-shaded boxes show additions recommended by 2 or more of 5 professional guidance documents.

MCIDs were identified in the original CER, by the expert reviewers or through supplemental literature searches.

Clinical Endpoint*	Instruments	MCID
PARKINSON'S DISEASE		
Domain: Clinician-Assessed Health Outcomes		
Global change in motor symptom severity	MDS-UPDRS III	3.25 points for detecting minimal, but clinically pertinent, improvement and 4.63 points for minimal, but clinically pertinent, worsening (Hungary) (Horváth et al., 2016)
Gait function, postural instability and balance	FOG-Q	MCID for improvement based on expert clinician rating was three scale points with a sensitivity of 0.67 and a specificity of 0.96 (Germany) (Fietzek et al., 2020)
	TUG	No MCID identified
Speech function	Speech Intelligibility Test	No MCID identified
Cognitive function	Stroop test	5.5 points (Sweden, Alzheimer's patients) (Borland et al., 2022)
	MoCA	1-2 points, depending on anchor vs. distribution method and choice of anchor (Sweden) (Lindvall et al., 2024)
Depression	HRDS	HRSD-6: 3-6.2 points (Hengartner & Plöderl, 2022) HRSD-17: 3-6 points (Hengartner & Plöderl, 2022) These estimates are from 4 studies conducted in the U.S., Germany and Japan. Estimates varied depending on the clinical scale used as an anchor and whether between- or within-patient differences were being considered
Motor-related ADL	MDS-UPDRS II	3.05 and 2.51 points for improvement and deterioration respectively (Hungary) (Horváth et al., 2017a)
Motor complications	MDS-UPDRS IV	0.9 and 0.8 points for improvement and worsening respectively (Hungary) (Makkos et al., 2019)

Clinical Endpoint*	Instruments	MCID
Nonmotor symptoms	MDS-UPDRS-I	2.64 and 2.45 points for improvement and deterioration respectively (Hungary) (Horváth et al., 2017a)
	MDS-NMSS	No MCID identified
Domain: Patient-Reported Outcomes		
QoL	PDQ-39	For summary indices: -4.72 and +4.22 for detecting minimal clinically important improvement and worsening respectively (Hungary) (Hungary) (Horváth et al., 2017b)
Sleep quality (10%)	PDSS-2	3.44 and 2.07 points for improvement and deterioration respectively (Hungary) Horváth et al., 2015b
Depression	BDI	3-7 points (Hengartner & Plöderl, 2022) These estimates are from 2 studies conducted in Japan. Estimates varied depending on the clinical scale used as an anchor and whether between- or within-patient differences were being considered.
ESSENTIAL TREMOR: No MCIDs were identified for frequently used or expert-recommended MCIDs.		

Abbreviations: ADL: Activities of daily living; FOG-Q: Freezing of Gait Questionnaire; MCID, minimal clinically important difference; MDC: minimum detectable change; MDS-NMS: Movement Disorders Society Nonmotor Symptoms; MDS-UPDRS: Movement Disorder Society-sponsored revision of UPDRS; MoCA, Montreal Cognitive Assessment; NA: Not applicable; PD: Parkinson's disease; PDQ: Parkinson Disease's Questionnaire; PDSS-2: Parkinson's Disease Sleep Scale 2nd version; QoL: Quality of life; TUG: Timed Up and Go Test; UPDRS: Unified Parkinson's Disease Rating Scale

*MCID definitions associated with formal instruments were sought only for the prioritized endpoints derived from the published systematic reviews and for endpoints emphasized by at least two of the professional guidance documents but not belonging to the prioritized list. **NOTE:** Appendix Table C1 in the full CER provides definitions of minimal detectable change (MDC) where MCIDs could not be identified. MDCs are based on statistical considerations, i.e., variability in treatment response, rather than clinical relevance. MCDs have to do with true effect versus random result.

Table D. Follow-up Duration Recommended by 3 External Expert Reviewers

Follow-up information was not captured in the CER. The following recommendations are entirely from the 3 external expert reviewers.

Domain	Reviewer Recommendations
Clinician-assessed health outcomes	Reviewer 1: 3 mos Reviewer 2: • Wearables/sensors and robot-assisted gait trainers, 3 mos • rTMS, 6 mos • DBS and ablations (subthalamotomies and MRg-FUS), 1 yr Reviewer 3: 6 mos and 1 yr
Patient-reported outcomes	The same as for clinician-assessed health outcomes
Safety (clinician-assessed or patient-reported)	Reviewer 1: 1 mo Reviewer 2: • Wearables/sensors and robot-assisted gait trainers, 3 mos • rTMS, 6 mos • DBS and ablations (subthalamotomies and MRg-FUS), 1 yr Reviewer 3: 1, 3 and 6 mos
Device durability	Reviewer 1: 1 mo Reviewer 2: • Wearables/sensors and robot-assisted gait trainers, 6 mos • rTMS, No response • DBS, should last for ≥5 yrs • Ablations (subthalamotomies and MRg-FUS), N/A Reviewer 3: 1 yr and 2 yrs

Abbreviations: DBS: deep brain stimulation; MRg-FUS: MRI-guided focused ultrasound; rTMS: repetitive transcranial magnetic stimulation. **NOTES:** Reviewer 1 is a neurosurgeon; Reviewer 2 are neurologists.