## CLINICAL PRACTICE GUIDELINES

# A Core Set of Outcome Measures for Adults With Neurologic Conditions Undergoing Rehabilitation

## A CLINICAL PRACTICE GUIDELINE

Jennifer L. Moore, PT, DHS, NCS,
Kirsten Potter, PT, DPT, MS,
Kathleen Blankshain, PT, DPT,
Sandra L. Kaplan, PT, DPT, PhD,
Linda C. O'Dwyer, MA, MSLIS, and
Jane E. Sullivan, PT, DHS, MS

South Eastern Norway Regional Knowledge Translation Center, Sunnaas Rehabilitation Hospital, Oslo, Norway, and Institute for Knowledge Translation, Carmel, Indiana (J.L.M); Department of Physical Therapy Education, Rockhurst University, Kansas City, Missouri (K.P.); Department of Physical Therapy and Human Movement Sciences (K.B, J.E.S) and Galter Health Sciences Library and Learning Center (L.C.O), Feinberg School of Medicine, Northwestern University, Chicago, Illinois; and Department of Rehabilitation and Movement Sciences, Rutgers School of Health Professions, Newark, New Jersey (S.L.K).

## ABSTRACT

**Background:** Use of outcome measures (OMs) in adult neurologic physical therapy is essential for monitoring changes in a patient's status over time, quantifying observations and patient-reported function, enhancing communication, and increasing the efficiency of patient care. OMs also provide a mechanism to compare patient and organizational outcomes, examine intervention effectiveness, and generate new knowledge. This clinical practice guideline (CPG) examined the literature related to OMs of balance, gait, transfers, and patient-stated goals to identify a core set of OMs for use across adults with neurologic conditions and practice settings.

**Methods:** To determine the scope of this CPG, surveys were conducted to assess the needs and priorities of consumers and physical therapists. OMs were identified through recommendations of the Academy of Neurologic Physical Therapy's Evidence Database to Guide Effectiveness task forces. A systematic review of the literature on the OMs was conducted and additional OMs were identified; the literature search was repeated on these measures. Articles meeting the inclusion criteria were critically appraised by 2 reviewers using a modified version of the COnsensus-based Standards for the selection of health Measurement INstruments. (COSMIN) checklist. Methodological quality and the strength of statistical results were determined. To be recommended for the core set, the OMs needed to demonstrate excellent psychometric properties in high-quality studies across neurologic conditions.

**Results/Discussion:** Based on survey results, the CPG focuses on OMs that have acceptable clinical utility and can be used to assess change over time in a patient's balance, gait, transfers, and patient-stated goals. Strong, level I evidence supports the use of the Berg Balance Scale to assess changes in static and dynamic sitting and standing balance and the Activities-specific Balance Confidence Scale to assess changes in balance confidence. Strong to moderate evidence supports the use of the Functional Gait Assessment to assess changes in dynamic balance while walking, the 10 meter Walk Test to assess changes in gait speed,

and the 6-Minute Walk Test to assess changes in walking distance. Best practice evidence supports the use of the 5 Times Sit-to-Stand to assess sit to standing transfers. Evidence was insufficient to support use of a specific OM to assess patient-stated goals across adult neurologic conditions. Physical therapists should discuss the OM results with patients and collaboratively decide how the results should inform the plan of care.

**Disclaimer:** The recommendations included in this CPG are intended as a guide for clinicians, patients, educators, and researchers to improve rehabilitation care and its impact on adults with neurologic conditions. The contents of this CPG were developed with support from the APTA and the Academy of Neurologic Physical Therapy (ANPT). The Guideline Development Group (GDG) used a rigorous review process and was able to freely express its findings and recommendations without influence from the APTA or the ANPT. The authors declare no competing interest.

**Video Abstract available** for more insights from the authors (see Video, Supplemental Digital Content 1, available at: http://links.lww.com/JNPT/A214).

**Key words:** *gait disorders, human movement system, nervous system diseases, neurodegenerative diseases, neurologic, neurologic examination, neurologic rehabilitation, outcome and process assessment (health care), outcome assessment (health care), patient care planning, patient outcome assessment, postural balance, practice guideline, psychometrics, reproducibility of results*

Correspondence: Kirsten Potter, PT, DPT, MS, Department of Physical Therapy Education, Rockhurst University, Kansas City, MO 64110 (Kirsten.Potter@rockhurst.edu).

# TABLE OF CONTENTS

# LEVELS OF EVIDENCE AND GRADE OF RECOMMENDATIONS

This clinical practice guideline (CPG) is intended to be a guide for rehabilitation management of adults with neurologic conditions and to inform outcome measurement research. The CPG applies to all adult patients with neurologic conditions, including those with acute (ie, <6 months since onset/diagnosis), chronic stable (ie, >6 months since onset/diagnosis, but not expected to worsen with time), and chronic progressive (ie, > 6 months since onset/diagnosis, but with the potential to experience additional symptoms or functional changes). Clinicians and organizations should interpret these recommendations in the context of the patient's situation, clinical practice, and potential for harm. The methodology used in this CPG, including the critical appraisal and assignment of levels of evidence and strength of the recommendations, was derived from the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist,[1-5] recommendations from the APTA, and is in accordance with internationally accepted methodologies for evidence-based practice. This CPG is organized to present the level of evidence definitions and the grades of recommendations (Tables 1 and 2, respectively), clear and implementable recommendations in the form of 9 action statements, an introduction and description of the need for this CPG, and a standardized profile for each action

statement that meets the Institute of Medicine's criteria for transparency of the CPG.[6] The 9 action statements include recommendations for the core set of measures, use of the core set, and collaborative decision-making. Research recommendations are included in the action statement profiles and summarized at the end of the CPG.

Each article included in this CPG was appraised by 2 reviewers, and assigned a level of evidence by the guideline development group (GDG). The grading criteria to determine the level of evidence that supports the recommendations are described in Table 1. These criteria, recommended by the Academy of Neurologic Physical Therapy (ANPT), were modified to accommodate descriptions of studies of psychometric properties. Levels I and II differentiate stronger from weaker studies by integrating the quality of the research design and/or reporting of the study,[7] as well as the strength of the psychometric data.[8,9] The criteria for the grades of recommendation assigned to each action statement are provided in Table 2. The grade reflects the overall strength of the evidence available to support the action statement. Throughout the CPG, each action statement is preceded by a letter grade indicating the strength of the recommendation, followed by the statement and summary of the supporting evidence.

**TABLE 1.** Levels of Evidence

| I | *Level of evidence rating I:* Evidence obtained from at least one high-quality (>50% critical appraisal score) study of psychometric properties |
| II | *Level of evidence rating II:* Evidence obtained from multiple, lesser quality (<50% critical appraisal score) studies of psychometric properties |
| III | *Level of evidence rating III:* Evidence obtained from one lesser quality (<50% critical appraisal score) study of psychometric properties |
| IV | Not applicable to studies of psychometric properties |
| V | Expert opinion (or best practice) |

**TABLE 2.** Grades of Recommendations

| GRADE | RECOMMENDATION | STRENGTH OF RECOMMENDATION |
|---|---|---|
| A | Strong evidence | A preponderance of level I studies, but least 1 level I study directly on the topic supports the recommendation |
| B | Moderate evidence | A preponderance of level II studies, but at least 1 level II study directly on the topic supports the recommendation |
| C | Weak evidence | A preponderance of level III studies, but at least 1 level III study directly on the topic supports the recommendation |
| P | Practice | Best practice based on expert opinion (review papers, white papers, consensus documents) developed by various methodologies (e.g., Delphi and RAND) and the clinical experience of the guideline development group |
| R | Research | An absence of research on the topic, or conclusions from existing studies on the topic are in disagreement |

# SUMMARY OF ACTION STATEMENTS

**A. Action Statement 1: STATIC AND DYNAMIC SITTING AND STANDING BALANCE ASSESSMENT.** Clinicians should use the Berg Balance Scale (BBS) for adults with neurologic conditions who have goals to improve static and dynamic sitting and standing balance and have the capacity to change in this area. The BBS should be administered under the same test conditions using the protocol recommended by the CPG Knowledge Translation (KT) Committee at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**B. Action Statement 2: WALKING BALANCE ASSESSMENT.** Clinicians should use the Functional Gait Assessment (FGA) for adults with neurologic conditions who have goals to improve balance while walking and have the capacity to change in this area. The FGA should be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: moderate

**A. Action Statement 3: BALANCE CONFIDENCE ASSESSMENT.** Clinicians should use the Activities-specific Balance Confidence (ABC) Scale to assess self-reported changes in balance confidence in adults with neurologic conditions who have goals and the capacity to change in this area. The ABC should be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, discharge, and, when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**B. Action Statement 4: WALKING SPEED ASSESSMENT.** Clinicians should use the 10 meter Walk Test (10mWT) for adults with neurologic conditions who have goals to improve walking speed and have the capacity to change in this area. The 10mWT should be administered (per the protocol by Steffen and Seney,[10] as adapted by the CPG KT Committee) under the same test conditions at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: V; recommendation strength: best practice
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**B. Action Statement 5: WALKING DISTANCE ASSESSMENT.** Clinicians should use the 6-Minute Walk Test (6MWT) for adults with neurologic conditions who have goals to improve walking distance and the capacity to change in this area. The 6MWT should be administered (per the Quinn et al protocol,[11] as adapted by the CPG KT Committee) under the same test conditions at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: V; recommendation strength: best practice
- Chronic stable conditions: Evidence quality: I; recommendation strength: moderate
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**P. Action Statement 6: TRANSFER ASSESSMENT.** Clinicians should document the transfer ability of adults with neurologic conditions who have goals to improve transfers and have the capacity to change. Documentation should include the type of transfer, level of required assistance, equipment or context adaptations, and time to complete. In patients who have goals and the capacity to improve sit-to-stand transfers, the 5 Times Sit-to-Stand (5TSTS) may be used. The 5TSTS and documentation of other transfers may be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, and discharge, and when feasible, between these periods for adult patients with neurologic conditions (Evidence quality: V; recommendation strength: best practice).

**P. Action Statement 7: DOCUMENTATION OF PATIENT GOALS.** Clinicians should document patient-stated goals and monitor changes in individuals with neurologic conditions, using an outcome measure (OM) such as the Goal Attainment Scale (GAS), reporting the task, the performance conditions, and the time to complete or level of independence desired. Patient goals should be documented at least 2 times, at admission and discharge, and, when feasible, between these testing periods (Evidence quality: V; recommendation strength: best practice).

**B. Action Statement 8: USE OF THE CORE SET OF OUTCOME MEASURES.** Clinicians should use and document the OMs in the core set to assess changes over time. The core set includes the BBS, FGA, ABC, 10mWT, 6MWT, and 5TSTS, and the recommended patient goal assessment for adults who are undergoing neurologic physical therapy. The core set should be administered with patients who have goals and the capacity to improve transfers, balance, and/or

gait. In cases when a patient cannot complete one or more core set OMs (eg, a patient who is unable to walk; thus, cannot complete the 10mWT or the 6MWT), a score of 0 should be documented. The core set should be administered under the same test conditions at least 2 times, at admission and discharge, and when feasible between these periods (Evidence quality: II; recommendation strength: moderate).

**P. Action Statement 9: DISCUSS OUTCOME MEA-SURE RESULTS AND USE COLLABORATIVE/ SHARED DECISION-MAKING WITH PATIENTS.** Clinicians should discuss the purpose of OMs, OM results, and how these results influence treatment options with patients undergoing neurologic physical therapy. Collaboratively, the clinician and the patient should decide how these data should inform the plan of care (Evidence quality: V; recommendation strength: best practice).

These guidelines were issued in 2018 based on the scientific literature published before March 2016. These guidelines will be considered for review by 2023, or sooner if new evidence becomes available. The ANPT will oversee the process and methodology for updating the CPG. The GDG will work collaboratively with the ANPT Evidence-Based Guideline Committee. Any updates to the guidelines in the interim period will be noted on the ANPT Web site.

## INTRODUCTION

### Purpose of Clinical Practice Guidelines

The APTA and the ANPT support the use of CPGs, as they provide therapists with evidence-based recommendations to guide clinical decision-making.[12] This CPG pertains to the examination of patients with neurologic conditions. Per the *Guide to Physical Therapist Practice,*[13] the physical therapy examination consists of 3 components: history, systems review, and tests and measures. Using standardized tests and measures is recommended, and selection of these measures is informed by their psychometric properties and clinical utility. Standardized tests and measures may be used to predict and diagnose, discriminate, and assess changes over time. Measuring outcomes is also emphasized in the *Guide to Physical Therapist Practice.*[13] The term "outcome measure" is used to refer to a standardized test or measure that is used to monitor changes in a specific construct (eg, gait function) during an episode of care. Various terms are used in the literature related to OMs, including standardized assessments, instruments, and tools. OMs exist and can be used for assessment at any level of the International Classification of Function, Disability, and Health (ICF),[14] including body function and structure, activity, and participation. The focus of this CPG is to describe evidence that supports the use of specific standardized measures (both performance-based and self/patient-reported),[15,16] and the term "OM" is used to describe these measures. Furthermore, this CPG identifies gaps in the research related to OMs that may be used in adult neurologic rehabilitation.

The recommendations presented in this CPG follow the efforts of the ANPT to develop measurement recommendations as part of the Evidence Database to Guide Effectiveness (EDGE) initiative. From 2009 to 2015, 6 ANPT EDGE task forces identified standardized tests and measures, including OMs, for use in several patient populations (stroke, multiple sclerosis, Parkinson disease, traumatic brain injury, spinal cord injury, and vestibular dysfunction). These task forces aimed to enhance the quality of care and decrease unwarranted variation in practice by recommending standardized tests and measures for each condition. The EDGE process included a literature review, and a synthesis of psychometric properties and clinical utility data. Using a modified Delphi process, recommendations were made for the use of 243 standardized measures in clinical practice, education, and research. Each task force developed recommendations for specific patient subgroups (eg, acute, subacute, and chronic stroke) and across a variety of health care settings.[17-21] This work may have enhanced the quality of rehabilitation by providing clinicians with a substantial amount of summarized information for each OM for the target patient population. However, due to the large number of OMs reviewed and recommended, it is unlikely that the goal of decreasing unwarranted variation in practice was achieved. Furthermore, the recommendations provided by each task force were focused on specific patient populations and not intended for use across all populations of patients with neurologic conditions.

### Background and Need for a Core Set of OMs

In 2012, the Institute of Medicine recommended that health care organizations build a learning health system that collects and analyzes standardized measurement data in clinical practice to measure patients' perspectives, improve care delivery, increase transparency of outcomes, link clinicians' performance to patient outcomes and internal and external benchmarks, manage patient care, improve processes, strengthen public health, and generate knowledge.[22] The core set of OMs recommended in this CPG provides a first and necessary step toward achieving the learning health system vision in neurologic physical therapy. Using OMs throughout a patient's episode of care is considered good clinical practice[23] and may enhance care by contributing to a more thorough examination and tailored care plan.[24] OMs can be used to monitor changes in a patient's status over time, quantify observations and patient-reported function over time, enhance communication between care settings,[15,16] and increase the efficiency of the delivery of patient care.[25] OMs can also help managers measure costs,[26] identify "at-risk" patients,[27] enhance reimbursement,[28] and compare outcomes among clinicians and settings.[27] Use of a common set of OMs promotes best practice by allowing direct comparisons of outcomes associated with different interventions.[29] Widespread adoption of a core set of OMs across clinical settings would support the Institute of Medicine recommendations, and may enable robust data collection efforts to rapidly advance clinical practice through the development of practice-based evidence.[30]

Despite reports describing the benefits of routine use of OMs, they are inconsistently used in rehabilitation.[23,24,31] Reported barriers include time, available equipment, perceptions of patient burden, clinician attitude/knowledge/skill, lack of financial compensation, and poor availability of measures.[24,32-36] Current practice is characterized by great variation in the use of OMs, few mandates for the use of specific OMs, and a lack of recommendations for a core set of OMs across neurologic conditions. With the exception of the Functional Independence Measure, which is required in inpatient rehabilitation, no measure (or group of measures) is required for all patients with neurologic conditions receiving physical therapy. Yet, the Centers for Medicare & Medicaid Services (CMS) now requires the use of objective measures of function in outpatient physical therapy practice.[37] The APTA, through PTNow, has identified multiple OMs that can be used to meet the requirements set by CMS.[38] However, to date, a core set of OMs has not been identified for use in neurologic physical therapy practice; thus, the primary purpose of this CPG is to identify a core set of OMs for use with adults who have neurologic conditions.

### Scope

This CPG aims to standardize practice by providing rehabilitation clinicians with recommendations for a core set of OMs for adults with neurologic conditions that should be routinely used in all settings. Based on input provided by physical therapists (PTs) and consumers of physical therapy, the core set focuses on the highest priority constructs of balance, gait, transfers, and patient-stated goals. Use of the core

set should increase standardization of OM selection and administration and provide the ability to measure changes in a patient's status over time. In addition, greater standardization of OMs should enhance effective communication among providers and with patients/caregivers, facilitate intervention effectiveness analysis and programmatic assessment within and among facilities, and may improve reimbursement.

This CPG focuses on adult patients (older than 18 years), of either sex, who are undergoing physical therapy services for treatment of a neurologic condition (eg, an injury or disease to the central or peripheral nervous system). The CPG action statements apply:

- When examining balance, gait, transfers, and when setting patient goals.
- In all health care settings or contexts, across the continuum of care settings, including but not limited to acute care hospitals, inpatient and outpatient rehabilitation, skilled nursing facilities, and home health care.

The specific goals of this CPG are to:

1. Standardize the use of a core set of OMs to assess changes over time in neurologic physical therapy within and among facilities.
2. Facilitate comparison of outcomes across interventions, providers, and patients within and among diagnostic groups through the use of a common set of measures.
3. Facilitate the development of practice-based evidence by standardizing the use of OMs for patients with neurologic conditions to enable the creation and analysis of large data sets.
4. Improve quality of care by standardizing data elements to answer important clinical questions (eg, identification of treatment responders vs nonresponders).
5. Ensure systematic and standardized documentation of OMs to help justify a patient's need for therapy and to inform policy. Improved documentation of OMs could be used to clarify and improve policies related to reimbursement and access to care.
6. Identify gaps in the literature related to OMs in adult neurologic rehabilitation. This may prompt researchers to rigorously study the psychometric properties of untested OMs or develop new measures to meet clinical needs.
7. Enhance the education of future rehabilitation providers by informing curricular decisions about the core set of OMs to include in entry-level and residency physical therapy education.

## Statement of Intent

Primarily intended for application in adult neurologic rehabilitation, this CPG may be useful to rehabilitation professionals including PTs, physical therapist assistants (PTAs), occupational therapists, and occupational therapy assistants who select and administer OMs; therapeutic recreation therapists, physicians, and nurses who are interested in understanding the use of OMs in rehabilitation; educators who make decisions about academic curricula; researchers who select or study OMs; regulatory bodies and policy makers; professional associations (eg, the APTA, APTA Academies of Neurology and Geriatrics, Canadian Physiotherapy

Association, and World Confederation of Physical Therapy); consumer organizations and associations (eg, the National Stroke Association and the Multiple Sclerosis Society); health care administrators, and third-party payers. This CPG does not serve as a legal standard of care or mandate. It provides recommendations for the use of a core set of OMs in clinical practice, based on a rigorous systematic review and critical appraisal process. Adherence to these guidelines will not guarantee a positive outcome in care; however, it is anticipated that the CPG will improve quality of care when implemented. Furthermore, this CPG does not provide a comprehensive review of all OMs. Rather it focuses exclusively on OMs in the constructs of balance, gait, transfers, and patient-stated goals. The appropriate use of the recommended OMs in clinical practice is ultimately the decision of each clinician and patient/significant other. If these OMs are not used, the rationale for the use of other OMs should be documented. We intend for the OM results to be shared with patients and significant others during adult neurologic rehabilitation. Collaboratively, clinicians and patients should decide how the results should guide the plan of care.

## METHODS

The steps outlining the process of review and determination of the core set are shown in Table 3. The GDG consisted of 3 PTs (J.M., K.P., and J.S.) with expertise in outcome measurement. Two of the team leaders (J.S. and K.P.) served as Chair of the ANPT's EDGE task forces for stroke[18] and multiple sclerosis,[17] respectively. The third (J.M.) led the development of the Rehabilitation Measures Database[39] and has expertise in knowledge translation. The GDG proposed the CPG on the core set of OMs to the ANPT's Board of Directors, who approved the proposal. The GDG attended the APTA Clinical Practice Guideline Workshop in July 2013 and received funding from the APTA in December 2013 to support the CPG's development.

The GDG recruited 2 consultants including a methodologist (S.K.) to provide advice on conducting the systematic review and writing the CPG, and a psychometrician (C.H.C.—see Acknowledgments) to assist with survey development, modifying COSMIN to create a critical appraisal tool, and data analysis. A medical reference librarian (L.O.) led the literature search process and assisted with writing the CPG. A doctor of physical therapy student (K.B.) functioned as a graduate assistant who assisted with the development and management of article and data storage systems, coordinated communication between the GDG and article reviewers, and assisted with data analysis and writing of the CPG.

The GDG also recruited an expert panel consisting of an international and diverse group of stakeholders who provided feedback about the scope, process, and final CPG recommendations. The expert panel, identified in the Acknowledgments, included consumers (ie, patients) who had received neurologic physical therapy; PTs (novice and experienced) who were members of the ANPT; other rehabilitation professionals (neurologists, occupational therapists, speech/language pathologists, neuropsychologists); representatives of professional associations; health care administrators; journal

**TABLE 3.** Outline of the CPG Process

| STEP | GENERAL PROCESS | SPECIFIC TASKS |
|---|---|---|
| 1 | Team recruitment | Recruitment of consultants, medical librarian, graduate assistant, and expert panel members. |
| 2 | Identification of CPG scope and focus | Development of surveys to identify scope and focus of the CPG; IRB approval obtained. |
| 3 | | Administered surveys to consumers and PTs; analyzed data to determine the CPG scope in areas of balance, gait, transfers, and patient-stated goals. |
| 4 | Initial identification of OMs considered for the CPG | Identification and evaluation of OMs reviewed by the 6 EDGE task forces. The measures met the following criteria to be considered for inclusion in the core set: (a) received a rating of ≥2 by the EDGE task force(s); (b) generic (eg, not condition-specific); (c) relevant to the scope of the CPG (balance, gait, transfers, and patient-stated goals); (d) availability of data in at least 2 neurologic populations; (e) able to track patient change over time; (f) high clinical utility (ie, free, <20 min to administer, and no specialized equipment); (g) published data on reliability and data to assess change (coincided with the literature review, described later). |
| 5 | Literature reviews, identification of additional OMs, and review of OMs for inclusion in the CPG | Literature search for articles pertaining to the OMs reviewed by EDGE task force(s) with databases searched from article inception through April 2015. |
| 6 | | Title and abstract review. Two members of the GDG reviewed each article (third member serving as tie breaker when needed). Inclusion criteria included English language, subjects older than 18 years with adult-onset neurologic condition, studied reliability and/or psychometric properties that assess change, and sample size ≥30 (or power analysis conducted and sample size met). |
| 7 | | Full-text article review using inclusion criteria described previously. One member of the GDG reviewed each article, sorting each into folders in accordance with the OM studied (eg, Berg Balance Scale) and the psychometric property studied. Inclusion criteria described previously. |
| 8 | | Review of literature for newly identified OMs. Literature search for articles pertaining to the additional OMs identified with databases searched from article inception through April 2015. |
| 9 | | Title and abstract review of articles pertaining to additional OMs identified; process described previously. |
| 10 | | Full-text article review of articles pertaining to additional OMs identified; process described previously. |
| 11 | | Final literature search to identify articles published between April 2015 and March 2016; articles reviewed using the process described previously. |
| 12 | Use and modification of COSMIN to rate article methodological quality | COSMIN selected and modified to meet the needs of the CPG and process for scoring COSMIN was established. |
| 13 | Article reviewer recruitment and training | Article reviewers recruited. |
| 14 | | Online program developed to train article reviewers to use modified COSMIN. Each potential reviewer completed the training program and analyzed one article using modified COSMIN; those achieving a score of 80% were invited to serve as an article reviewer. |
| 15 | Critical appraisal of articles | Two reviewers scored each article using modified COSMIN administered via SurveyMonkey. |
| 16 | | Graduate assistant exported data from SurveyMonkey to Excel spreadsheet and compared data from the 2 reviewers. Inconsistencies addressed by initial reviewers when able; if continued inconsistencies existed, the GDG member solved the conflict. |
| 17 | Scoring of article methodological quality | Scores for each section of COSMIN were calculated. Total article methodological score calculated based on lowest section score received for the given article. Level I or II article rating determined. |
| 18 | Scoring of psychometric properties | COSMIN recommendations for scoring strength of psychometrics adopted and used to score reliability and measurement error from each article for each OM. |

(*continues*)

**TABLE 3.** Outline of the CPG Process (*Continued*)

| STEP | GENERAL PROCESS | SPECIFIC TASKS |
|---|---|---|
| 19 | | The GDG identified process for review of OM data for inclusion in core set and set criteria for strong vs moderate recommendations. |
| 20 | Analysis of OM data across articles | Combined information from all articles on a given measure, as related to psychometric properties studied, strength of psychometric data, patient population studied, and category (acute, chronic stable, and chronic progressive). |
| 21 | Recruitment of KT team | KT team recruited. |
| 21 | Action statement profile generation | Action statements creation using BridgeWiz. |
| 22 | Review of CPG | CPG reviewed by ANPT Evidence-Based Documents Advisory Committee and CPG Expert Panel, and will be reviewed APTA PTNow using Appraisal of Guidelines, Research, and Evaluation (AGREE II) tool. |
| 23 | | CPG reviewed by KT committee using Guideline Implementability Appraisal Tool (GLIA) |
| 24 | | Public review of the CPG with feedback submitted through SurveyMonkey. |

Abbreviations: AGREE, Appraisal of Guidelines for Research and Evaluation; COSMIN, Consensus-based Standards for the Selection of Health Measurement Instruments; CPG, clinical practice guideline; EDGE, Evidence Database to Guide Effectiveness; GDG, guideline development group; IRB, institutional review board; KT, Knowledge Translation; OM, outcome measure; PT, physical therapist.

editors; and experts in OMs, knowledge translation, policy, and reimbursement.

## Methods to Determine the CPG Scope

To identify the scope and focus of the CPG, the GDG developed and administered separate online surveys to consumers of neurologic physical therapy services and to ANPT members. Surveys were administered via SurveyMonkey and focused on the use of OMs during physical therapy examination and care. Before dissemination, the surveys were approved by the Institutional Review Boards at Northwestern University (Chicago, Illinois) and Rockhurst University (Kansas City, Missouri).

## Consumer Survey

An invitation to consumers of neurologic physical therapy was distributed through the Clinical Neuroscience Research Registry at the Rehabilitation Institute of Chicago and Northwestern University, Heartland Chapter of the National Parkinson's Disease Foundation, and the Mid America Chapter of the National Multiple Sclerosis Society. Participants included individuals with email access who were registered in the research and/or email databases of these organizations. Approximately 828 people with stroke, 395 with spinal cord injury (SCI), 11 635 with multiple sclerosis (MS), and 2500 with Parkinson disease (PD) received an invitation to participate. The invitation provided a link to the survey, and indicated that participation was optional. To be eligible, consumer participants were required to have a medically diagnosed neurologic condition, have received physical therapy services, be 18 years or older, English-speaking, and have email access. Participants confirmed that they met these inclusion criteria and provided informed consent on the first page of the survey.

The 21-item survey included questions pertaining to neurologic physical therapy, including the:

- reason for seeking services;
- frequency, duration, and setting of services;
- perceived importance of improving function in various areas (eg, gait and decrease fatigue);
- constructs (eg, balance) examined using tests performed by the PT;
- formats of tests used in clinical settings (eg, questionnaires and performance tests);
- frequency and duration of testing;
- information provided by the PT regarding the purpose and results of tests;
- perceived importance of the tests;
- recommendations for therapy time that should be dedicated to testing; and
- satisfaction with services and information received about the tests conducted.

## Academy of Neurologic Physical Therapy Member Survey

Approximately 5000 PT and PTA members of the ANPT were invited to participate in the survey. Inclusion criteria required that the PT or PTA be licensed, college educated, and have email access. A link to the survey was sent via e-mail through the ANPT's listserv and electronic newsletter. Survey participation was optional and the respondent provided informed consent prior to survey initiation.

The ANPT member survey included a maximum of 65 questions; the number and type of questions answered varied by the participant's responses. Survey logic ensured that questions received by each respondent were relevant to the individual's role (eg, clinician or educator/researcher/other). The survey consisted of 3 sections: demographic data, a core set needs assessment, and use of OMs in practice. Demographic data included primary and professional roles, experience (eg, number of years of experience, certifications, and training on OMs), APTA and ANPT membership, education, primary employment setting, and willingness to use a core set of OMs. The core set needs assessment questions captured the respondent's understanding of core sets and their use; importance of having a core set; types of OMs recommended for

the core set; representation of the ICF domains and specific items (eg, aerobic capacity) in the core set; time and money to support use of the core set; and benefits and potential impact of the core set. Lastly, questions inquired about use of OMs in practice. Clinician respondents were asked about current use of OMs in practice, whereas the educators, researchers, and other respondents (eg, managers) were asked to provide their thoughts on what should be measured in practice.

De-identified aggregate data from both surveys were analyzed using descriptive statistics. Data were used to inform the scope and focus of the CPG, particularly to identify the highest priorities for each sample group.

## Survey Results

A total of 518 individuals completed the survey (303 PTs and 215 consumers). The PT respondents reported their primary position as either a clinician (69%) or educator (24%). They were experienced, with 45% having greater than 15 years of experience and 54% holding American Board of Physical Therapy Specialties certification. The majority were employed either in an outpatient (46%) or inpatient rehab setting (28%). The neurologic conditions experienced by the consumers included MS (49%), stroke (34%), or SCI (14%). Most received outpatient physical therapy (70%), and some received services in inpatient rehabilitation (21%).

Survey results showed that 94% of clinicians use OMs in clinical practice. The majority reported having 30 to 60 minutes to conduct examinations at admission (78%), interim (53%), and discharge (52%). Almost all (98%) reported that a core set is either *essential* (65%) or *desirable* (33%), and 91% indicated they were *very willing* (58%) or *willing* (33%) to incorporate a core set of OMs into practice. Regarding the maximum amount of time that should be used to administer OMs, the greatest number (43%) answered 15 to 29 minutes. All stated the core set should include OMs related to the ICF domain of activity, with 98% scoring this as *essential*. Clinicians scored the following constructs as *essential* to include in the core set: balance (98%), gait (95%), patient-stated goals (82%), and transfers (81%).

Results from the consumer survey showed that they also value the use of tests in their care; 59% scored tests as *very important* and 35% as *somewhat important*. Of note, consumers identified that they were referred to physical therapy due to walking (83%) and balance difficulties (68%), with approximately 90% indicating it was *very important* to improve walking and balance.

Survey results indicated that OMs that assess changes in balance and gait are important to both clinicians and consumers and should be included in the core set. In addition, the PT survey indicated OMs related to patient-stated goals and transfers were also important for inclusion in the core set.

## Selection of Measures to Consider for the CPG

Two sets of measures were evaluated for the inclusion in the CPG—(1) all measures (n = 243) that had been reviewed by the 6 ANPT EDGE task forces,[17-21] and (2) new measures (n = 67) identified during the literature search—that were not originally reviewed by the EDGE task forces and were studied in any adult neurologic population. During each step of the review process, the GDG reached consensus on decisions about measure inclusion.

Appendix 1 provides a list of measures reviewed for inclusion in the CPG (see Supplemental Digital Content 2, Appendix 1, available at: http://links.lww.com/JNPT/A215). Details about the literature search are provided in the section titled *Literature Search*.

### EDGE-Reviewed Measures

#### Step 1. Identification of Standardized Measures With EDGE Ratings of 2 to 4/4

All 243 standardized measures reviewed by the ANPT EDGE task forces were considered for inclusion in the CPG. The EDGE task forces used a 1- to 4-point rating scale to make recommendations for measures in categories such as condition acuity, severity, and site of care.[21] A rating of "4" indicated that the measure had excellent psychometric properties and clinical utility in the target condition; a "1" rating indicated poor psychometrics (inadequate reliability or validity) or limited clinical utility (extensive testing time, unusual or expensive equipment, costs to administer, etc).[17,18] In step 1, measures that received a "1" rating across all categories and EDGE groups were eliminated. A total of 222 standardized measures were retained.

#### Step 2. Identification of Generic/Not Condition-Specific Standardized Measures

To identify measures that could be used across neurologic populations, condition-specific measures (eg, Stroke Impact Scale) were eliminated. One hundred forty-six of the 222 standardized measures were retained.

#### Step 3. Identification of Standardized Measures That Address the CPG Target Constructs

The remaining measures were evaluated relative to the constructs of balance, gait, transfers, and patient-stated goals. A measure was eliminated if fewer than 75% of the items or questions assessed these constructs. Fifty-four of the 146 measures were retained.

#### Step 4. Identification of Standardized Measures Used in 2 or More Neurologic Populations

To identify OMs that were appropriate for use across neurologic conditions, measures were eliminated that did not have published psychometric data in at least 2 neurologic populations. Forty-one of the 54 standardized measures were retained.

#### Step 5. Identification of Standardized Measures That Evaluate Change

Each measure was evaluated to determine whether it could be used to demonstrate changes over time. The availability of psychometric properties that assess changes (eg, minimum detectable change and minimum clinically important difference) for each measure was ascertained. All 41 standardized measures were retained.

#### Step 6. Identification of Measures With Excellent Clinical Utility

Approximately 85% of PT survey respondents indicated that 45 minutes or less should be spent on OM administration, with 63% indicating the maximum time spent on measure

administration should be less than 30 minutes. In addition, 71% indicated the OM should cost $100 or less. Therefore, the GDG decided that, to be included, an OM had to be free, require equipment commonly available in a clinic, and take 20 minutes or less to administer. Thirty-five of the 41 OMs were retained.

## Step 7. Identification of Candidate OMs

Step 7 followed a literature search of the 35 OMs that met the criteria described in steps 1 through 6. Following the literature search, title/abstract screening, and full-text review, each OM was evaluated to determine whether reliability and data to support interpretation of results (eg, minimal detectable change [MDC] and minimal clinically important difference [MCID]) were available in at least one article that met inclusion criteria for the CPG. The remaining 16 measures and relevant literature proceeded to a critical appraisal with data extraction via the modified COSMIN checklist[5] by the trained reviewer pool.

## New Measures

During the initial literature search (including the title/abstract and full-text review), the GDG identified 67 additional measures that were not previously reviewed by EDGE. These measures were reviewed using the process described in steps 2 through 7 previously. The measures retained during each step are described next.

*Step 1*: Not applicable because these measures were not reviewed by the EDGE task forces.

*Step 2*: 65 of the 67 new measures were retained; 2 were excluded because they were condition-specific.

*Step 3*: 52 of the 65 measures were retained; 13 were excluded because fewer than 75% of the test items pertained to gait, balance, transfers, and patient-stated goals.

*Step 4*: 13 of the 52 measures were retained; 39 were excluded because there were no published data in 2 or more neurologic populations.

*Step 5*: 12 of the 13 measures were retained; 1 was excluded because there were no data on psychometric properties that indicated the measure could detect changes over time.

*Step 6*: 10 of the 12 measures were retained; 2 were excluded because they did not meet the clinical utility criteria.

*Step 7*: 2 of the 10 measures were retained and relevant literature proceeded to a full-text review and data extraction by the trained reviewer pool using the modified COSMIN checklist.[5] Eight were eliminated because they lacked data demonstrating reliability and supporting interpretation of the results (eg, MDC and MCID).

## Literature Search

A medical librarian (L.O.) collaborated with the GDG to develop the search strategies to identify articles related to each of the OMs of interest. The study types included meta-analyses, systematic reviews, and psychometric studies in the following databases: PubMed MEDLINE, Embase, Cochrane Central Register of Controlled Trials, and CINAHL. Search strategies for the Embase, CENTRAL, and CINAHL databases were adapted from the PubMed MEDLINE search strategy. A validated search filter, developed by COSMIN for finding studies on OMs, in conjunction with the search strategies in PubMed, was used.[40] A validated version of the filter was also used for the Embase search (developed by E. P. Jansma, Medical Library, VU University, Amsterdam, the Netherlands). The search strategy is depicted in Appendix 2 (see Supplemental Digital Content 3, available at: http://links.lww.com/JNPT/A216).

The initial searches focused on articles pertaining to the EDGE-reviewed OMs and were performed in April 2015, October 2015, and December 2015, resulting in a total of 18 007 articles. All databases were searched back to their inception, and no language or date limits were applied. This literature review is depicted in Appendix 2. After duplicates were removed, 12 088 articles remained. To be included, the study was published in English, studied the English language version of the OM, and assessed reliability and or values support interpretation of the results (eg, standard error of measurement [SEM], MDC, and MCID). In accordance with COSMIN, the sample size needed to be a minimum of 30; articles with a sample size less than 30 were acceptable if a power analysis was done and the required sample size was met. Lastly, study participants needed to be adults (18 years or older) with a neurologic condition. Table 4 outlines the inclusion and exclusion criteria.

The titles and abstracts of the 12 088 articles were reviewed by 2 of 3 GDG members, and reviewer pairs were rotated within the GDG. The third member played the role of tie breaker where disagreement on an article's inclusion occurred between the 2 initial reviewers. Following the title and abstract review, 11 548 articles were excluded. Full-text reviews were conducted on the remaining 540 articles; each was reviewed by 1 GDG member using the same criteria. A second GDG member assessed articles if questions or concerns about an article were identified. Lastly, the graduate assistant reviewed the reference lists in each article to identify any additional relevant articles. None was identified.

Follow-up literature searches using the strategies described previously were performed in March 2016 to identify any new articles published since April 2015; 403 articles were identified after duplicate removal. After title and abstract review, 365 articles were excluded, leaving 38 additional articles for review. The PRISMA diagram (Figure) illustrates the article search processes used; 64 articles were included for full-text review (see Supplemental Digital Content 4, Figure, available at: http://links.lww.com/JNPT/A218).

## Critical Appraisal Tool Development

To determine the methodological quality of the articles, the original version of the COSMIN[8,9,41,42] was modified (COSMIN-M). COSMIN[1,3-5] provides a standard for evaluation of the study design and statistical analysis of the psychometric properties, including sections representing these psychometric properties: internal consistency, reliability, measurement error, content validity, construct validity, structural validity, hypothesis-testing, cross-cultural validity, criterion validity, responsiveness, and interpretability. During an article review using COSMIN-M, only the sections appraising properties assessed in the study were completed by reviewers, using a dichotomous (eg, yes or no) scale. For example, if a study only reported on reliability, reviewers only completed COSMIN-M sections on reliability and general methodology.

**TABLE 4.** Inclusion and Exclusion Criteria for Article Review

| CRITERION | INCLUSION | EXCLUSION |
|---|---|---|
| Language of article | Published in English | Published in language other than English |
| Language of OM | OM studied is not the English version | OM studied is the non-English version |
| Sample size | n = ≥30 or n < 30, but power analysis done and sample size met | Sample size <30 and no power analysis done<br>Sample size <30 but insufficient to meet power analysis requirements |
| Conditions | Acquired neurologic conditions for entire sample<br>If a mixed (neuro; nonneuro) must report data separately for neuro and must meet sample size requirement for neuro subset | Adults with congenital neuro conditions<br>Study focuses on nonneuro populations<br>Study includes subjects with neuro and nonneuro conditions, but data reported in aggregate<br>Dementia<br>Study includes only community-dwelling elders |
| Age | Minimum of 18 y old | Sample includes individuals younger than 18 y |
| Purpose of article | OM on our list<br>Psychometrics assessed | Systematic review<br>Meta-analysis<br>Intervention study |
| Psychometrics | Study examined one or more of the following:<br>    Internal consistency<br>    Reliability<br>    MDC<br>    MCID<br>    SEM<br>    Ceiling and/or floor effects | Does not assess one of the target psychometric properties |

Abbreviations: MCID, minimal clinically important difference; MDC, minimal detectable change; OM, outcome measure; SEM, standard error of measurement.

Although the original COSMIN rating scale has been modified to incorporate a 4-point scale (poor, fair, good, and excellent), the GDG selected the original version to facilitate ease of scoring and higher reliability of the reviewers.

In consultation with the methodologists, to focus on the purpose and intent of this CPG, the following modifications were made to the COSMIN tool by the GDG. We retained COSMIN questions about statistical techniques used and results, and questions about the presence of potential study flaws. However, the sections on internal consistency, reliability, interpretability, and generalizability were modified to reduce the number of items and include only those that were of utmost importance to determining the methodological quality of the study. Questions relevant to the development of the core set were also retained. For example, questions pertaining to psychometric variables that measure changes, such as MCID, MDC, and SEM, were retained, as these can be used to set goals and determine treatment effectiveness. Additional questions about specific psychometric values, such as intraclass correlation coefficients (ICCs), and the location of that data in the manuscript were added. Appendix 3 provides a list of measurement terms used in the CPG with definitions (see Supplemental Digital Content 5, Appendix 3, available at: http://links.lww.com/JNPT/A217). The COSMIN-M generalizability section included questions pertaining to the neurologic condition of the population studied (eg, stroke and PD), acuity and stability (progressive and nonprogressive) of the condition, age and sex, and the setting in which the study took place. A new section, labeled "general methodology," related to sample size, missing data, and rater training and experience was included. Reviewers completed the COSMIN-M via an online survey Web site (SurveyMonkey).[43] Appendix 4 provides a copy of the COSMIN-M. Two members of the GDG reviewed each article to determine and document any reported adverse events (see Supplemental Digital Content 6, Appendix 4, available at: http://links.lww.com/JNPT/A219).

## Reviewer Selection and Training

Article reviewers were recruited at the 2015 APTA Combined Sections Meeting and via postings on the ANPT's e-newsletter and listserv. All applicants completed an online reviewer training course developed by the GDG using Articulate Storyline 2™. The training program consisted of an overview of the CPG and the COSMIN-M, followed by a detailed description of the methods for completing each section of the COSMIN-M (internal consistency, reliability, interpretability, generalizability, and general methodology). Lastly, information was provided outlining the CPG process and reviewer expectations.

The GDG selected one article for reviewer training and testing, and 2 GDG members first completed the online COSMIN-M for the article. The third GDG member served as a tiebreaker to resolve any conflicts. The GDG's final ratings were used as a basis for the testing score agreement with article reviewers. Each potential reviewer completed the COSMIN-M review for 2 measures studied in this article. To successfully complete the training and begin reviewing articles, a reviewer needed to score 80% or more agreement with the GDG score. If needed, reviewers were allowed

a second chance to resubmit the review on the same article (without any feedback on the previous review) and achieve a score of 80% or more; 23 individuals successfully completed the training to review articles.

## Scoring of Methodological Quality

Two reviewers assessed the methodological quality of each article using the online COSMIN-M (Appendix 4), for each OM reported in the article. To avoid redundancy, each reviewer completed the general methodology section only once for each article. The graduate assistant exported COSMIN-M data into an Excel spreadsheet to compare data from the 2 reviewers. When inconsistencies were identified, reviewers were asked to reevaluate the question and confirm or change the original response. When inconsistencies continued, a GDG member resolved the conflict.

Once the results were finalized, the score for each section was calculated using the percentage of "yes" responses to the questions. Section scores were compared to inform the overall article quality score, which reflected the score received by the lowest scoring section. For example, if an article received 80% for reliability and 60% for measurement error, the article would receive an overall quality score of 60%. If the overall quality score was 50% or more, the article received a level I rating. If the score was < 50%, the article could not receive higher than a Level II rating.

The strength of the psychometric data was determined in accordance with COSMIN (Table 5). Relevant statistical results from each article were evaluated to determine whether they exceeded the threshold established by COSMIN (Table 5). If the article received a level I rating and had strong psychometric properties, the article received a psychometric property rating of *strong* ($+++$). A rating of strong ($---$) was used for level I studies where the psychometric properties were below the COSMIN threshold. Level II articles received a score of *moderate* ($++$) if the psychometric properties met the psychometric threshold and a *moderate* ($--$) if the psychometric properties were below the threshold. Ratings of *strong* (?) or *moderate* (?) were assigned if specific psychometric properties were not studied (eg, where MDC was calculated, but not minimal important change [MIC]). After this step, each article was assigned a level of evidence and statistical strength score.

Finally, information from multiple articles on each OM was combined, including level of evidence, strength of psychometric property, the patient population studied, and the condition category (acute, chronic stable, and chronic progressive) as depicted in Table 6 (step 4). The *acute* category was defined as participants who had the condition for less than 6 months; this applied to individuals with new conditions that were expected to improve (eg, peripheral vestibular hypofunction) or to those with potentially long-lasting, but recently diagnosed conditions (eg, stroke, SCI, and brain injury). The *chronic stable* category was defined as more than 6-month duration, but not expected to progress with time, applying to participants with conditions such as stroke, SCI, or brain injury diagnosed more than 6 months ago. The *chronic progressive* category was defined as more than 6 months in duration, but with potential to experience additional symptoms or functional decline (eg, amyotrophic lateral sclerosis, MS, or PD).

## Recommended Action Statements

Using BridgeWiz for APTA 3.0, action statements were generated that include clear and implementable recommendations, consistent with the Institute of Medicine recommendations for transparency.[44] The first step was to identify OMs that demonstrated level I evidence of excellent internal consistency and/or reliability and SEM/MDC data in 2 or more populations and 3 condition categories (acute, chronic stable, and chronic progressive). If a construct area did not have an OM that met this first criterion, other OMs that demonstrated level I evidence of excellent internal consistency and/or reliability and SEM/MDC data in 2 or more populations and 2 categories were considered. Because the aim of this CPG was to recommend a core set of OMs for use in adult neurologic conditions, when more than one OM in a construct area had substantial supporting evidence, the OM with the strongest psychometric properties across diagnostic groups was selected. For the construct of gait, measures of speed and endurance were considered separately, as these represent 2 different, yet important, aspects of gait performance. Similarly, for balance, both performance-based and patient-reported measures were considered separately. Only one OM for the construct of transfers met the criteria for consideration in the core set. Because this was a priority area identified in our surveys, and the OM had some data to support inclusion in the core set, a best practice recommendation was made and documentation standards were recommended for other types of transfers.

For patient-stated goals, no OMs were identified with sufficient literature for recommendation in the core set. Instead, general recommendations for documentation standards were developed. To standardize administration of OMs in clinical practice, recommendations related to the general OM use and OM timing were also generated. Lastly, recommendations were made related to the sharing of OM-related information and decisions with patients. Research recommendations (designated by R) were generated to identify missing or conflicting evidence related to using the psychometric variables studied in the CPG, for OMs that should be studied across more condition categories, and regarding study of recommended administration protocols.

## Guideline Review

1. This CPG underwent 4 formal reviews. The first review was conducted by the GDG using 2 tools:
   - The Appraisal of Guidelines for Research and Evaluation (AGREE II)[45] an instrument used to assess CPG quality with 23 items in 6 domains. Each item is rated using a 7-point rating scale that includes specific rating criteria.
   - The Guideline Implementability Appraisal v 2.0 (GLIA)[46] to assess each action statement across 8 dimensions of implementability including executability, decidability, validity, flexibility, effect on care processes, measurability, novelty/innovation, and computability.
2. A second review included completion of the AGREE II by the ANPT Evidence-Based Documents committee and CPG expert panel. Eight reviewers completed the AGREE II. The aggregate score was 94%. The GLIA tool was completed by each member of the ANPT-appointed Knowledge Translation Task Force

**TABLE 5.** COSMIN Ratings for Strength of Statistics[a8]

| LEVEL | COSMIN RATING | CRITERIA |
|---|---|---|
| I | Strong (+++) | *Level of evidence rating I:* Evidence obtained from a high-quality (≥50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (+++):* |
| | | Internal consistency: Cronbach α ≥ 0.7 |
| | | Reliability: ICC or weighted κ ≥ 0.70; Pearson's $r$ ≥ 0.80 |
| | | Measurement error: MIC > SDC *or* MIC outside the LOA |
| | | Responsiveness: Floor or ceiling effect <15% |
| | Strong (− − −) | *Level of evidence rating I:* Evidence obtained from a high-quality (≥50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (− − −):* |
| | | Internal consistency: Cronbach α < 0.7 |
| | | Reliability: ICC or weighted κ < 0.70; Pearson's $r$ < .80 |
| | | Measurement error: MIC ≤ SDC *or* MIC inside the LOA |
| | | Responsiveness: Floor or ceiling effect ≥ 15% |
| | Strong (?) | *Level of evidence rating I:* Evidence obtained from a high-quality (≥50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (?):* |
| | | Internal consistency: Cronbach α not determined |
| | | Reliability: Neither ICC or weighted κ, nor Pearson's $r$ determined |
| | | Measurement error: MIC not defined |
| II | Moderate (++) | *Level of evidence rating II:* Evidence obtained from a lesser quality (<50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (++):* |
| | | Internal consistency: Cronbach α ≥ 0.7 |
| | | Reliability: ICC or weighted κ ≥ 0.70; Pearson's $r$ ≥ 0.80 |
| | | Measurement error: MIC > SDC *or* MIC outside the LOA |
| | | Responsiveness: Floor or ceiling effect <15% |
| | Moderate (− −) | *Level of evidence rating II:* Evidence obtained from a lesser quality (<50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (− −):* |
| | | Internal consistency: Cronbach α < 0.7 |
| | | Reliability: ICC or weighted κ < 0.70; Pearson's $r$ < 0.80 |
| | | Measurement error: MIC ≤ SDC *or* MIC inside the LOA |
| | | Responsiveness: Floor or ceiling effect ≥15% |
| | Moderate (?) | *Level of evidence rating II:* Evidence obtained from a lesser quality (<50% critical appraisal score) study of psychometric properties |
| | | *Psychometric property rating (?):* |
| | | Internal consistency: Cronbach α not determined |
| | | Reliability: Neither ICC or weighted κ, nor Pearson's $r$ determined |
| | | Measurement error: MIC not defined |

Abbreviations: COSMIN, Consensus-based Standards for the Selection of Health Measurement Instruments; ICC, intraclass correlation coefficient; LOA, limits of agreement; MIC, minimal important change; SDC, smallest detectable change.
[a]From Terwee.[8]

**TABLE 6.** Process Used to Make Recommendations

| STEP | DESCRIPTION | CRITERIA |
|---|---|---|
| 1. Score articles | Review and score methodological quality for the study of psychometric properties (internal consistency, reliability, measurement error, and responsiveness) | Quality of each psychometric property scored<br>Article assigned score of the lowest scoring section<br>Level I if ≥50% criteria met<br>Level II if <50% criteria met |
| 2. Score strength of psychometric properties | Review statistical results from articles, score the psychometric property while considering the article level of evidence | Statistical strength criteria listed in Table 5. |
| 3. Combine results by OM | Compile data by OM to view amount and quality of literature, and strength of psychometric property | Considered data for each OM for level of evidence, strength of psychometric, condition, and category (acute, chronic progressive, chronic stable) |
| 4. Select OMs for consideration of core set | Compare the amount and strength of literature available for each OM. If an OM met the criteria listed, it was compared with other OMs in the same construct area. | Prioritized OMs that met the following criteria:<br>Level I evidence in ≥2 populations and 3 categories (acute, chronic stable, and chronic progressive)<br>AND<br>Internal consistency and/or reliability (strong +++) in 2 populations and *3 categories*<br>AND<br>Standard error of measurement and/or minimum detectable change data (strong +++)/ (strong ?) in 2 populations and *3 categories*<br>In cases in which a measurement construct did not have an OM with this level of evidence, we considered OMs that met the following criteria:<br>Level I evidence in ≥2 populations and 3 categories (acute, chronic stable, and chronic progressive)<br>AND<br>Internal consistency and/or reliability (strong +++) in 2 populations and *2 categories*<br>AND<br>Standard error of measurement and/or minimum detectable change data (strong +++)/(strong ?) in 2 populations and *2 categories* |

Abbreviation: OM, outcome measure.

(n = 8). The aggregate score was 88%. Feedback from the reviewers on the AGREE II and GLIA reviews was integrated in the final CPG. It is anticipated that a further review would result in a comparable/higher score.

3. A revised draft of the CPG was posted for public comment on the ANPT, APTA, and Academy of Geriatric Physical Therapy Web sites by the ANPT Director of Practice. Notices of the public comment period were distributed via email to CPG reviewers and others who inquired about the CPG while it was in development. An electronic newsletter and social media posting disseminated the public comment notice to ANPT members. The posting was also made available on a web-based listserv of PTs who treat individuals with neurologic conditions. Listserv subscribers included members and nonmembers of the ANPT. During the public comment period, reviewers identified the following strengths of the CPG: usefulness, value, clarity, comprehensiveness of the literature review, and format. There were some comments for improvement that the GDG determined were beyond the scope of the CPG. Numerous suggestions for dissemination were forwarded to the CPG KT Committee.

4. The fourth review was completed by 2 *Journal of Neurologic Physical Therapy* peer reviewers prior to publication.

## Document Structure

The action statements are organized under the following headings: the core set of OMs, discussing results of OMs, and shared decision-making. After the action statement profiles, a section that describes implementation recommendations for all action statements is included. Lastly, acknowledgments and references are provided.

## THE CORE SET OF OUTCOME MEASURES FOR NEUROLOGIC PHYSICAL THERAPY

**A. Action Statement 1: STATIC AND DYNAMIC SITTING AND STANDING BALANCE ASSESSMENT.** Clinicians should use the BBS for adults with neurologic conditions who have goals to improve static and dynamic sitting and standing balance and have the capacity to change in this area. The BBS should be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**Aggregate Evidence Quality and Strength**: Level I; strong. Based on 16 level I studies (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220).

**Benefits:**

- The BBS demonstrates excellent internal consistency and reliability, and data exist to assist in interpretation and measuring changes, in individuals with acute, chronic progressive, and chronic stable neurologic conditions. Floor and ceiling effects and information to assist in OM result interpretation, such as MDC and MCID, are available for individuals with acute, chronic stable, and chronic progressive neurologic conditions.
- The BBS has high clinical feasibility, as it requires minimal equipment, is free, and requires less than 20 minutes to administer.
- Ninety-seven percent of PTs surveyed reported that a balance assessment is an *essential* component for the core set.
- Initial costs of purchasing equipment (eg, stopwatches and measuring device) are minimal and the required equipment is commonly available in clinical settings. The time cost to administer the test is less than 20 minutes.

**Risk, Harm, and Cost:**

- No adverse events were documented in research studies.

**Benefit-Harm Assessment:** Preponderance of benefit.

**Value Judgments:** The GDG emphasizes the importance of using standardized administration and scoring procedures for measuring patients in the clinic. While there is not a universally accepted protocol for the BBS, we recommend that each clinical site adopt the testing protocol developed by the CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standardized procedures and, on an annual basis, establishing consistency within and among raters using the BBS.

**Intentional Vagueness:** The BBS has demonstrated a ceiling effect in individuals with acute,[47-50] chronic stable,[50,51] and chronic progressive conditions.[52] The BBS only includes one item that assesses sitting balance. Therefore, if a patient has a primary goal to improve sitting balance, the BBS should be administered in addition to a sitting balance measure.

**Role of Patient Preferences:**

- Sixty-eight percent of consumers surveyed reported that balance was a common reason for seeking a PT referral.
- Clinicians should consider the degree to which improvements in balance are achievable and important to each patient.

**Exclusions:**

- For patients who do not have explicit goals to improve static and dynamic sitting and standing balance, the clinician should document that the BBS was not administered and provide a rationale (eg, not applicable due to the patient's current and expected functional capability or not applicable due to a lack of related patient goals).
- Patients who have a high level of balance ability (eg, able to walk without an assistive device at a gait speed >1.0 m/s) may experience a ceiling effect on the BBS.

**Quality Improvement:**

- Organizations may use BBS results to assess balance outcomes of individuals and groups with neurologic conditions receiving rehabilitation.
- The physical therapy profession may use BBS scores to describe the effectiveness of physical therapy services for adults with neurologic conditions.

**Implementation and Audit:**

- The measurement error of the BBS may vary throughout the scale. It may be more difficult to achieve high reliability on individuals who score between 20 and 44.[53,54] Measurement error has not been established for individuals with an average score of less than 20, thus it is unknown.[54] Additional efforts may be needed to standardize and improve reliability of BBS administration in clinical practice for patients who score less than 44.
- The BBS has demonstrated a ceiling effect in individuals with acute,[47-50] chronic stable,[50,51] and chronic progressive conditions.[52] In patients who perform well on the BBS, and score near the top of the scale, it may not be necessary to readminister the test.
- Clinics and organizations should establish administration consistency within and among clinicians prior to using the BBS, and this should be repeated annually.

## Supporting Evidence and Clinical Interpretation (Table 7)

**Administration and Conditions:** The BBS is a 14-item clinician-rated scale that assesses sitting and standing, static and dynamic balance.[55] Considered one of the most commonly used measures in adult neurologic rehabilitation,[56] the

**TABLE 7.** Evidence Table, Berg Balance Scale

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **Berg Balance Scale, acute samples** | | | | | | | |
| Chinsongkram et al[60] | Stroke | I | NT | NT | NT | 0% | 4.3% |
| Gustavsen et al[47] | Stroke; subacute rehabilitation | I | NT | NT | NT | NT | 36% (admission) |
| Lemay and Nadeau[48] | SCI; AIS D; 15 paraplegia, 17 tetraplegia | I | NT | NT | NT | NT | 37.5% |
| Mao et al[49] | Stroke | I | Cronbach α = 0.92–0.98 | Interrater ICC = 0.95 (total score); weighted k = 0.92 (individual items) | NT | 14 d 35%; 30 d = 17.3%; 90 d = 6.5%; 180 d = 5% | 14 d = 4.9%; 30 d = 11.8%; 90 d = 21.5%; 180 d = 28.8% |
| Pickenbrock et al[57] | Stroke; acute care facility | I | NT | Interrater: Mean difference between raters 0.13; LOA < ±1 point (−0.25, 0.51) | NT | NT | NT |
| Salter et al[59] | Stroke; inpatient rehabilitation | I | NT | NT | NT | 0% | 0% |
| Stevenson[58] | Stroke | I | NT | Test-retest ICC = 0.92 | MDC 90% = 6; MDC 95% = 7; SEM = 2.49 | NT | NT |
| **Berg Balance Scale, chronic stable samples** | | | | | | | |
| Hiengkaew et al[63] | Stroke; outpatient clinic | I | NT | Test-retest ICC = 0.95 | SEM = 1.68 MDC = 4.66 | NT | NT |
| Liaw et al[64] | Stroke | I | NT | Interrater ICC = 0.98 | SEM = 2.4; MDC = 6.7 | NT | NT |
| Tsang et al[51] | Stroke | I | NT | NT | NT | 0% | 32.10% |
| Wirz et al[65] | SCI AIS 4.8% A, 4.8% B, 83.3% C, 7.1% D | I | NT | Interrater ICC = 0.953 | NT | NT | NT |

(*continues*)

**TABLE 7.** Evidence Table, Berg Balance Scale (*Continued*)

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **Berg Balance Scale, chronic progressive samples** | | | | | | | |
| Leddy et al[62] | PD; UPDRS/MDS mean 72.6 (25-135); mean Hoehn and Yahr 2.45 (4-66) | I | NT | Interrater ICC = 0.95; test-retest ICC = 0.79 (student PT), 0.80 (PT) | NT | NT | NT |
| Paltamaa et al[66] | MS EDSS median 2.0 (0-6.5) | NA | NT | NT | MIC-deterioration (clinician anchor) = −0.60; MIC-deterioration (patient anchor) = −1.41 | NT | NT |
| Quinn et al[11] | HD (various stages) | I | NT | Test-retest ICC = 0.86-0.97 | MDC premanifest HD = 1; manifest HD = 5, early-stage HD = 4, middle-stage HD = 5, late-stage HD = 5 | NT | NT |
| Schlenstedt et al[52] | n = 85; PD; Hoehn and Yahr mean 2.4; UPDRS mean 40.1; 33% female, 67% male; mean age 67.2 (40-82) | I | NT | Test-retest = 0.95 | NT | 0% | 17.6% |
| Steffen and Seney[10] | PD; Hoehn and Yahr 1-4 (median 2) | I | Cronbach α = 0.86-0.87 | Test-retest ICC = 0.94 | MDC = 5 | NT | NT |
| **Berg Balance Scale, acute and chronic stable mixed samples** | | | | | | | |
| Knorr et al[50] | Stroke, first data at average of 3.3 mo poststroke, second data collection at average 8.2 mo poststroke | I | NT | NT | NT | 0% at baseline, 0% at follow-up | 15% at baseline, 21% at follow-up |

Abbreviations: AIS, American Spinal Injury Association Impairment Scale; d, day; EDSS, Expanded Disability Status Scale; HD, Huntington's disease; ICC, intraclass correlation coefficient; LOA, limits of agreement; MCID, minimal clinically important difference; MDC, minimal detectable change; MDS, Movement Disorder Society; MIC, minimal important change; MS, multiple sclerosis; NA, not applicable; NT, not tested; PD, Parkinson disease; PT, physical therapist; SCI, spinal cord injury; SEM, standard error of measurement; UPDRS, Unified Parkinson's Disease Rating Scale.

BBS has been well studied in research and widely used in research and clinical practice. A standardized testing form with administration instructions is available, and commonly available equipment (chair, stopwatch, ruler, and step) is used during testing. Each of the 14 items requires that the patient perform a specific activity to challenge balance. The patient's ability to complete each item is rated on a 0- to 4-point scale, with 0 representing the inability to complete the task and 4 reflecting independent item completion. The total score is calculated by summing the scores of the 14 items, with the maximum score of 56 and the minimum score of 0.[56]

**Populations:** The BBS can be applied across adult neurologic conditions. This action statement is based on 16 level I studies that reported data in 7 acute samples (6 stroke)[47,49,57-60] and 1 SCI,[48] 4 chronic progressive samples (1 Huntington's disease [HD][11] and 3 PD),[10,52,62] 4 chronic stable samples (3 stroke[51,63,64] and 1 SCI),[65] and 1 study that included a mixed acute and chronic stable sample (stroke).[50]

**Psychometric Data:** *Reliability:* Three level I studies examined reliability in individuals with acute stroke and demonstrated excellent interrater reliability. Mao et al[49] assessed the total score (ICC = 0.95) and individual item interrater reliability (weighted $\kappa$ = 0.92). Using a Bland-Altman plot, Pickenbrock et al[57] demonstrated a mean difference among raters of 0.13. While this demonstrates high interrater reliability, the article received a *strong* (?) reliability rating because of the statistics used in the study.[57] Excellent test-retest reliability has been demonstrated in individuals with stroke, with an ICC = 0.92.[58]

Three level I studies assessed reliability in chronic stable conditions. Excellent interrater reliability (ICC = 0.953) was demonstrated in individuals with chronic SCI.[65] Test-retest reliability results were also excellent in individuals with stroke, with ICCs of 0.95[63] and 0.98.[64]

Four level I studies examined reliability in individuals with chronic progressive conditions. Quinn et al[11] studied test-retest reliability of the BBS in individuals with HD, which resulted in ICCs of 0.86 to 0.97 across 5 manifestations of HD from premanifest to late-state HD.[11] Three additional studies of the BBS in PD suggest excellent interrater reliability (ICCs of 0.95[62] to 0.98),[52] and good to excellent test-retest reliability (ICCs of 0.94,[10] 0.95,[52] and 0.79-0.80) in PD.[62]

*Internal Consistency:* Two level I studies demonstrated excellent internal consistency of the BBS in acute and chronic progressive conditions, with a Cronbach $\alpha$ of 0.92 to 0.98 in individuals with acute stroke[49] and 0.86 to 0.87 in individuals with PD.[10]

*SEM, MDC, MCID, Ceiling, and Floor Effects:* Five level I studies assessed SEM or MDC for the BBS; however, none simultaneously reported an MIC or MCID. Thus, measurement error was rated as a *strong* (?) across the 5 studies. In participants with acute stroke, the SEM was 2.49 points,[58] whereas in chronic stroke the SEM varied from 2.4[64] to 1.68 points.[63] In individuals with HD (chronic progressive), the SEM was used to calculate the MDC, but was not explicitly stated in the article.[11] In participants with PD (chronic progressive), Hoehn and Yahr classification of 1 to 4 (median = 2), the SEM was used to calculate an MDC, but it was not explicitly reported.[10]

Five studies reported an MDC for the BBS. In participants with acute stroke, Stevenson[58] reported an $MDC_{95}$ of 7. In chronic stroke, the $MDC_{95}$ varied from 4.66[63] to 6.7 points.[64] In chronic progressive conditions, the $MDC_{95}$ varied based on the condition and severity. In participants with HD, the MDC varied from 1 in individuals with premanifest HD to 4 to 5 in individuals with other stages of HD.[11] Similarly, a study of individuals with PD demonstrated an $MDC_{95}$ of 5.[10] Only one study reviewed determined an MIC for the BBS. In participants with MS, the MIC-deterioration with clinician and patient anchors was −0.60 and −1.41, respectively.[66]

Six level I studies assessed the floor effects of the BBS. No floor effects were identified in 2 studies of individuals with acute stroke.[59,60] In contrast, Mao et al[49] identified the presence of a floor effect that varied by time poststroke, depending on the level of acuity as follows: 14 days = 35% (of sample), 30 days = 17.3%, 90 days = 6.5%, and 180 days = 5%. Studies conducted on individuals with chronic stroke and PD (mean Hoehn and Yahr = 2.4) indicated no floor effect.[51,52] Knorr et al[50] did not find a floor effect at 3.3 and 8.2 months poststroke.

Eight level I studies assessed ceiling effects of the BBS. In individuals with acute conditions, the presence of a ceiling effect varied by study. Ceiling effects of 36%[47] and 15%[50] of the sample were identified in subacute stroke, and 37.5%[48] in the SCI-ASIA Impairment Scale D. However, these results conflict with other data that identified 0%[59] to 4.3%[60] ceiling effect in a similar stroke population. A finding by Mao et al[49] may provide a potential reason for these conflicts, as they determined the ceiling effect varies by time poststroke, with 4.9% at 14 days, 11.8% at 30 days, 21.5% at 90 days, and 28.8% at 180 days. In individuals with chronic stroke, ceiling effects of 21%[50] and 32.1%[51] have been identified. A ceiling effect of 17.6% was also identified in individuals with PD.[52]

The strong recommendation for the BBS is based on level I evidence of internal consistency and/or reliability data, availability of information to assist in assessing changes, and floor and ceiling effect data across acute, chronic stable, and chronic progressive conditions.

**Related Outcome Measures:** While several other balance OMs were assessed in this CPG, the only other OM that assessed static and dynamic sitting balance in acute, chronic stable, and chronic progressive conditions was the Trunk Impairment Scale (TIS) (see Supplemental Digital Content 8, Appendix 6, available at: http://links.lww.com/JNPT/A221). This 10-item measure requires that a patient perform various activities in a sitting position. Two publications, including samples of acute and chronic brain injury[67] and MS,[68] demonstrated excellent reliability and established an SEM in MS. Other psychometric properties were not established. Because of the lack of psychometric evidence across categories, the TIS was not included in the core set.

Shorter BBS versions were considered (eg, BBS-3P, BBS 9, and BBS-Short form). While decreasing BBS administration time is desirable, these versions included different items and none had sufficient evidence to support use across patient populations. The FGA and other OMs that assess balance while walking were also reviewed, and have been discussed later in this CPG.

**R. Research Recommendation 1:** Researchers should further examine the BBS to determine its psychometric properties in neurologic conditions other than stroke, SCI, PD, HD, and MS. Properties such as SEMs, MDCs, and MCIDs/MICs should be established for individuals with scores throughout the range of the scale in all adult neurologic conditions. Specific information regarding the functional levels of individuals who may benefit from the BBS, and when to start with or transition to another OM, is needed. Determination of optimal administration timing would assist clinicians in administering the BBS within a reasonable time frame of when "real change" would be expected. Development and comprehensive testing of a BBS-Short form would decrease administration burden.

**R. Research Recommendation 2:** Studies on OMs that provide a comprehensive assessment of sitting balance across acute, chronic progressive, and chronic conditions are needed. These should aim to determine the psychometric properties, including reliability, and to identify information to assist in interpretation, such as MDCs and MIC/MCIDs.

**B. Action Statement 2: WALKING BALANCE ASSESSMENT.** Clinicians should use the Functional Gait Assessment (FGA) for adults with neurologic conditions who have goals to improve balance while walking and have the capacity to change in this area. The FGA should be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: moderate

    **Aggregate Evidence Quality and Strength:** Level I; moderate. Based on 5 level I and 1 level II studies (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220).
    **Benefits:**
- The FGA demonstrates excellent internal consistency in individuals with acute and chronic stable neurologic conditions and excellent reliability in individuals with acute, chronic progressive and chronic stable neurologic conditions. Floor and ceiling effects, and data to assist in interpretation and measuring change, such as MDC and MCID, are available for individuals with acute and chronic stable neurologic conditions.
- The FGA has high clinical feasibility, as it requires minimal equipment, is available for free, and requires less than 20 minutes to administer.
- Initial costs of purchasing equipment (eg, stopwatches and measuring device) are minimal and the required equipment is commonly available in clinical settings. The time to administer the test is less than 20 minutes.
    **Risk, Harm, and Cost:**
- No adverse events were documented in research studies.

**Benefit-Harm Assessment:** Preponderance of benefit.
**Value Judgments:** The GDG emphasizes the importance of using standardized administration and scoring procedures for measuring patients in the clinic. While no single protocol has been validated for the FGA, the GDG recommends that each facility adopt the testing protocol developed by the KT committee for this CPG (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standard procedures and, on an annual basis, establishing consistency within and among raters using the FGA.
**Intentional Vagueness:** The FGA has not been assessed for internal consistency, measures of change (eg, MDC, SEM, and MCID), and floor or ceiling effects in individuals with chronic progressive neurologic conditions.
**Role of Patient Preferences:**
- Sixty-eight percent of consumers surveyed reported that balance was an important goal and a primary reason for seeking physical therapy services.
- Clinicians should consider the degree to which improvements in balance are achievable and important to individual patients when determining whether to administer the FGA.
**Exclusions:**
- Clinicians should use discretion when applying the FGA with patients who do not have explicit goals to improve balance while walking. Dynamic balance may be required to perform other related tasks that are stated in the patient's goals; in these cases, the FGA would be appropriate to administer.
- The FGA should not be administered with patients who do not have the capacity to walk. A score of 0 should be documented in these instances.
**Quality Improvement:**
- Organizations may use FGA data to assess balance outcomes of individuals and groups with neurologic conditions receiving rehabilitation.
- FGA scores may be used to describe the effectiveness of physical therapy services for adults with neurologic conditions.
**Implementation and Audit:**
- The FGA is intended to assess balance while walking, and has demonstrated a ceiling effect in individuals with balance and vestibular deficits seen in a tertiary care center.[69] If a patient demonstrates a high score (near 30 out of 30), or is likely to do so, the clinician may need to select a more challenging OM to assess changes over time.
- If a patient is unable to ambulate, but has goals and capacity to improve balance, a baseline score of 0 should be documented on the FGA.
- For patients who perform well on the FGA and score near the top of the scale, it may not be necessary to readminister the test.
- Clinics and organizations should establish administration consistency within and among clinicians prior to using the FGA, and this should be repeated annually.

**195**

## Supporting Evidence and Clinical Interpretation (Table 8)

**Administration and Conditions:** The FGA is a 10-item clinician-rated test that assesses balance while walking. The items are rated on a 0- to 3-point scale, with 0 indicating severe impairment and 3 indicating normal ambulation. To score the FGA, the items are summed and a maximum total score is 30. A testing form with administration instructions is available,[70] and commonly available equipment (obstacles, stopwatch, and steps) is used during testing.

**Populations:** The majority of the studies reviewed for this CPG examined acute and chronic stable conditions, with only one level I study examining individuals with PD (chronic progressive).[62] Studies reviewed included level I studies on individuals with acute and chronic stroke,[71] acute and chronic vestibular dysfunction,[72] and a level II study on acute vestibular dysfunction.[69]

**Psychometric Data:** *Reliability:* Interrater, intrarater and test-retest reliability were assessed in articles reviewed for this CPG. Leddy et al[62] demonstrated excellent interrater reliability (ICC = 0.93) in patients with PD with a mean Hoehn and Yahr score of 2.45. A lower, but acceptable, interrater reliability (ICC = 0.73) was demonstrated in a mixed sample of individuals with acute or chronic vestibular dysfunction.[72] Excellent intrarater reliability was found in acute and chronic vestibular dysfunction (ICC = 0.94).[72] Leddy et al[62] found that student PTs had a slightly lower, but still excellent interrater reliability, with ICC = 0.80 as compared

with practicing PTs (ICC = 0.90). Excellent test-retest reliability (ICC = 0.95) was also demonstrated in a mixed sample of individuals with acute or chronic stroke.[71]

*Internal Consistency:* Two studies (levels I and II) assessed internal consistency of the FGA. Both studies demonstrated excellent internal consistency, with a Cronbach α of 0.86 in acute vestibular dysfunction[69] and 0.88 in a mixed acute and chronic vestibular population.[70]

*SEM, MDC, MCID, Ceiling, and Floor Effects:* Two studies of levels I and II evidence assessed the MDC and/or MDC% of the FGA, but neither study reported an MCID; the methodological quality ratings were *strong* (?)[71] and *moderate* (?),[69] respectively. In participants with mixed acute and chronic stable conditions, Lin et al[71] calculated an MDC of 4.2. In individuals with acute vestibular dysfunction, the SEM was utilized to determine the MDC of 6; however, the SEM was not explicitly reported.[69]

Two studies (one level I and one level II) assessed the FGA for ceiling and/or floor effects. In individuals with acute vestibular dysfunction, the ceiling effect was 25%.[69] A much lower ceiling effect of 0% to 5.7% and a floor effect of 0% to 2% were found in a mixed sample of individuals with acute or chronic stroke.[71] It is important to note that these studies were both completed in outpatient care settings. The presence of floor or ceiling effects in an inpatient setting has not been assessed.

The core set recommendation for the FGA was based on levels I and II evidence in acute conditions, and level I

**TABLE 8.** Evidence Table, Functional Gait Assessment

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **Functional Gait Assessment, acute samples** | | | | | | | |
| Marchetti et al[69] | Vestibular (tertiary care center) | II | Cronbach α = 0.86 | NT | MDC = 6 | NT | 25% |
| **Functional Gait Assessment, chronic progressive samples** | | | | | | | |
| Leddy et al[62] | PD (mean Hoehn and Yahr 2.45) | I | NT | Interrater ICC = 0.93 Test-retest ICC = 0.80 (student), 0.91 (PT) | NT | NT | NT |
| **Functional Gait Assessment, mixed acute and chronic stable samples** | | | | | | | |
| Lin et al[71] | Stroke (outpatient rehabilitation) | I | NT | Test-retest ICC = 0.95 | MDC = 4.2 MDC% = 14.1 | 0%-2% | 0%-5.7% |
| Nilsagård et al[72] | Vestibular | I | Cronbach α = 0.88 | Intrarater ICC = 0.94 Interrater ICC = 0.73 | NT | NT | NT |

Abbreviations: ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; NT, not tested; PD, Parkinson disease; PT, physical therapist.

evidence in chronic stable and chronic progressive conditions. Data to assist with measuring change are lacking in chronic progressive conditions. Therefore, the FGA received an aggregate recommendation rating of moderate.

**Related Outcome Measures:** Several OMs that assess balance while walking were reviewed for this CPG, and 4 had sufficient evidence to be considered for the core set. While the FGA had the highest-quality evidence across patient categories, the Dynamic Gait Index (DGI), Mini-Balance Evaluation Systems Test (Mini-BESTest), and Timed Up and Go (TUG) were also considered. The level of evidence for each measure is available (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220). The DGI (see Supplemental Digital Content 9, Appendix 7, available at: http://links.lww.com/JNPT/A222) met the criteria for the core set, but there were conflicting results from reliability studies. In a level I study with individuals with acute vestibular deficits, interrater reliability of the DGI was a κ of 0.64, with individual items ranging from 0.35 to 1.0,[74] whereas studies on PD[75] and stroke[71] demonstrated test-retest ICCs of 0.84[75] and 0.94.[71] The FGA was developed as a modification of the DGI; both OMs include the following items: gait level surfaces, changes in gait speed, gait with horizontal head turns, gait with vertical head turns, gait with pivot turn, step over obstacle, and stairs. Unlike the DGI, the FGA includes gait with narrow base of support, gait with eyes closed, and ambulating backward. The DGI includes step around obstacles, not included in the FGA. The FGA provides more specific operational definitions for its items. For example, the DGI indicates that the patient must have "good speed" to achieve a score of 3/3, but the FGA indicates the item must be completed in less than 5.5 seconds. A modified version of the DGI[76] was also assessed in this CPG; however, it did not have enough evidence to be considered for the core set. In summary, the FGA was selected instead of the DGI for inclusion in the core set for the following reasons: better reliability across acute, chronic stable and chronic progressive populations; inclusion of clinically relevant balance items of gait with narrow base of support, gait with eyes closed, and ambulating backward; and improved response categories to facilitate consistency in OM administration.

The Mini-BESTest (see Supplemental Digital Content 10, Appendix 8, available at: http://links.lww.com/JNPT/A223) was considered for inclusion in the core set of OMs; however, it did not meet the established criteria. Data existed from 1 level I study in acute conditions,[60] 2 level I studies in chronic progressive conditions,[52,77] and 1 level I study in a chronic stable condition.[51] No data were available on internal consistency, reliability, and measures of change (eg, MDC and MCID) in participants with acute conditions. Reliability was studied in chronic progressive conditions, but internal consistency and measures of change (eg, MDC and MCID) were not examined.

The TUG (see Supplemental Digital Content 11, Appendix 9, available at: http://links.lww.com/JNPT/A224) was considered for the core set, with a total of 9 level I studies meeting review requirements. Although the majority of the evidence was from participants with chronic progressive conditions (HD,[11] MS,[78] PD,[10,75,79] and postpoliomyelitis),[80] the TUG showed excellent reliability. In participants with stroke,[50,63,81] 3 articles described the reliability, MDC,

or ceiling and floor effects of the TUG. In participants with acute stroke,[50] only floor and ceiling effects of the TUG were established. Furthermore, the TUG includes a sit-to-stand transfer, walking speed, and turning, all of which are represented in other core set measures. Given the lack of reliability data in acute conditions and the overlap with other core set measures, the TUG was not selected for the core set.

**R. Research Recommendation 3:** Specific information regarding the functional levels of individuals who may benefit from the FGA and when to start with or transition to another OM is needed. Determination of optimal administration timing would assist clinicians in administering the FGA within a reasonable time frame of when real change can be expected. Development and psychometric testing of a FGA short-form would decrease administration burden.

**R. Research Recommendation 4:** Studies are needed to examine other OMs, such as the Mini-BESTest and the TUG, in individuals with acute, chronic progressive, and chronic stable neurologic conditions. While the FGA had enough evidence to support its inclusion of the core set, more comprehensive measures of standing and walking balance should be tested to ensure a complete comparison against the FGA. Properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**A. Action Statement 3: BALANCE CONFIDENCE ASSESSMENT.** Clinicians should use the ABC Scale to assess self-reported changes in balance confidence in adults with neurologic conditions who have goals and the capacity to change in this area. The ABC should be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: I; recommendation strength: strong
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**Aggregate Evidence Quality and Strength**: Level I; strong. Based on 3 level I studies (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220).

**Benefits:**
- The ABC demonstrates excellent internal consistency and has data to assist in measuring changes in individuals with acute, chronic progressive, and chronic stable neurologic conditions. Reliability has been assessed in a chronic progressive condition. Floor and ceiling effects, and information to assist in test result interpretation (eg, MDC), are available for individuals with acute, chronic progressive, and chronic stable neurologic conditions.
- The ABC has high clinical feasibility, as it is a patient-reported measure, requires only a writing utensil, is free to administer, and requires minimal time (5-10 minutes[82]).

- The time cost associated with this measure is minimal, as patients may be able to independently complete the ABC prior to their initial clinical visit.
  **Risk, Harm, and Cost:**
- No adverse events or financial costs were documented in research studies.
- There may be a potential burden to patients, as the ABC is a patient-reported measure.
- The tool is available in English, Turkish, and Spanish, so there is a risk of misinterpretation of items for those who are not fluent in these languages.
  **Benefit-Harm Assessment:** Preponderance of benefit.
  **Value Judgments:**
- The GDG emphasizes the importance of using standardized administration and scoring procedures for measuring patients in the clinic. While no single protocol has been used for the ABC, we recommend that each clinical site adopt the testing protocol developed by the CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standard procedures and, on an annual basis, establishing consistency within and among raters using the ABC.
- Standardization procedures should be reviewed on an annual basis.
- Administration of both clinician-rated and patient-reported measures may provide a more comprehensive assessment of balance confidence than administering only a clinician rated measure.[15]
  **Intentional Vagueness:**
- The ABC asks individuals to rate confidence in balance while doing several tasks at home and community. Individuals with a recently diagnosed neurologic condition may not have experience with these specific tasks since the onset of the condition. Clinicians should begin administering the ABC when it is appropriate for the patient.
- Individuals with lack of insight into impairments may have difficulty accurately answering the ABC questions.

In these cases, clinicians should use their judgment to determine appropriateness of administering this test.
- Patients with hand impairments may require assistance with recording their responses to the ABC.
  **Role of Patient Preferences:**
- Sixty-eight percent of consumers surveyed reported that balance was a common reason for seeking a PT referral.
- Clinicians should consider the degree to which improvements in balance are achievable and important to their individual patients when determining whether to administer the ABC.
  **Exclusions:**
- Clinicians should use discretion when applying the ABC with patients undergoing neurologic rehabilitation who do not have goals to improve balance confidence.
  **Quality Improvement:**
- Use of a single measure across clinical settings will facilitate communication among clinicians and more accurately reflect changes in a patient's perceived balance confidence over time.
- Organizations may use data collected from the ABC to assess changes in balance confidence in individuals with neurologic conditions receiving rehabilitation.
- ABC scores may be used to describe the effectiveness of physical therapy services for increasing balance confidence perceptions in adults with neurologic conditions.
  **Implementation and Audit:**
- While the ABC did not demonstrate a substantial ceiling effect, if a patient demonstrates a score near 100%, the clinician may stop using the OM for the purpose of measuring change over time.

## Supporting Evidence and Clinical Interpretation (Table 9)

**Administration and Conditions:** The ABC is a patient-reported OM that assesses a person's perceived confidence in performing functional activities without becoming unsteady or falling. The stem, "How confident are you that you will

**TABLE 9.** Evidence Table, Activities-specific Balance Confidence

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **Activities-specific Balance Confidence, chronic progressive samples** | | | | | | | |
| Jonasson et al[83] | PD % self-rated severity mild 24%, moderate 64%, severe 13% | I | Cronbach $\alpha = 0.98$ | Test-retest ICC = 0.86 | SEM = 11% | 0% | 4.9% |
| Steffen and Seney[10] | PD Hoehn and Yahr 1-4 (median 2) | I | Cronbach $\alpha = 0.95\text{-}0.96$ | Test-retest ICC = 0.94 | MDC = 13; SDD = 30.5% | NT | NT |
| **Activities-specific Balance Confidence, acute and chronic stable mixed samples** | | | | | | | |
| Salbach et al[84] | Stroke | I | Cronbach $\alpha = 0.94$ | NT | SEM = 5.05 | 0% | 0% |

Abbreviations: ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; NT, not tested; PD, Parkinson disease; PT, physical therapist; SEM, standard error of measurement; SDD, smallest detectable difference.

not lose your balance or become unsteady when you …?" leads to 16 items. Each item is rated on a 0% to 100% scale, and the total score is calculated by adding item scores and dividing by 16 (eg, the number of items). The resulting scores range from 0% to 100% and reflect overall perceived confidence. The ABC is a self (patient)-report measure; however, questions can be read to an individual and the responses recorded. One study used a mailed version of the ABC, but did not provide any details about instructions related to the methods to complete the scale.[83] Two studies were conducted in a laboratory setting, but did not provide details about the ABC test administration.[10,84]

**Populations:** The ABC has been tested in individuals with acute, chronic progressive, and chronic stable conditions. Two level I studies examined individuals with PD,[10,83] and 1 level I study included a mixed sample of individuals with acute and chronic stroke.[84]

**Psychometric Data:** *Reliability:* Test-retest reliability was assessed in individuals with PD in 2 level I studies; both demonstrated excellent reliability, with ICCs ranging from 0.86[83] to 0.94.[10] Reliability has not been assessed in acute or chronic stable conditions.

*Internal Consistency:* In a sample with acute or chronic stroke, Salbach et al[84] demonstrated excellent internal consistency (Cronbach $\alpha$ = 0.94). In 2 studies on individuals with PD of various levels of impairment, the Cronbach $\alpha$ ranged from 0.95 to .96[80] to 0.98.[83]

*SEM, MDC, MCID, Ceiling, and Floor Effects:* SEM was assessed in 3 level I studies, with results stated in 2 studies. In individuals with mixed acute and chronic stable conditions, the SEM was 5.05.[84] In PD, Steffen and Seney[10] identified an SEM of 13% and the smallest detectable difference of 30.5%. Jonasson et al[83] calculated an MDC of 11%. While this MDC was relatively close to the SEMs reported in individuals with PD, Steffen and Seney[10] reported a substantially higher MDC of 30% in a similar sample. When applying these data in clinical practice, the patient should be similar to the sample studied.

Floor and ceiling effects of the ABC have been reported in individuals with acute and chronic stroke and in PD. In a mixed sample of individuals with acute or chronic stroke, no floor or ceiling effects were identified (0%).[84] In individuals with PD (self-rated severity, mild 25%, moderate 64%, severe 13%), no floor effects and minimal ceiling effects (4.9%) were identified.[83]

The strong recommendation for the ABC is based on level I evidence of internal consistency and/or reliability data, and availability of data to assist in measuring change across acute, chronic stable, and chronic progressive conditions.

**Related Outcome Measures:** No other patient-reported OMs of balance had sufficient literature to be considered for the core set. The Falls Efficacy Scale-International had evidence to support its use in acute and chronic progressive conditions (see Supplemental Digital Content 12, Appendix 10, available at: http://links.lww.com/JNPT/A225). In 3 separate level I studies, reliability, internal consistency, and data to assist in measuring changes were established.[83,85,86] Floor and ceiling effects and MDCs have also been published. This OM has also been translated and tested in many different languages. Because of the lack of evidence to support the

use of this measure with individuals who have chronic stable conditions, it was not recommended for the core set.

**R. Research Recommendation 5:** Studies are needed to determine the psychometric properties (eg, reliability) of the ABC in acute, chronic progressive, and chronic stable neurologic conditions. Furthermore, information to assist clinicians in interpreting the results of the ABC, such as MDCs and MIC/MCIDs, should be established across neurologic conditions. Specific information regarding the characteristics of individuals who may benefit from the ABC is needed.

**R. Research Recommendation 6:** Studies are needed to examine other OMs, such as the Falls Efficacy Scale International, in individuals with acute, chronic progressive, and chronic stable neurologic conditions. While evidence supports the inclusion of the ABC in the core set, other patient-reported measures of balance should be studied to ensure a comprehensive comparison to the ABC. Properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**B. Action Statement 4: WALKING SPEED ASSESSMENT.** Clinicians should use the 10 meter Walk Test (10mWT) for adults with neurologic conditions who have goals to improve walking speed and have the capacity to change in this area. The 10mWT should be administered (per the protocol by Steffen and Seney[10] as adapted by the CPG KT Committee) under the same test conditions at admission, discharge, and, when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: V; recommendation strength: best practice
- Chronic stable conditions: Evidence quality: I; recommendation strength: strong
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**Aggregate Evidence Quality and Strength:** Level I; strong. Based on 8 level I studies reporting reliability and/or data to assist in measuring changes in acute, chronic stable, and/or chronic progressive conditions, 2 level I studies reporting ceiling and floor effect data in acute, and 1 study reporting only MIC data in a chronic progressive condition (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220).

**Benefits:**
- The 10mWT demonstrates excellent reliability in individuals with chronic progressive and chronic stable neurologic conditions. Data to assist in interpretation and measuring change exists in acute, chronic progressive, and chronic stable populations.
- Floor and ceiling effects have been assessed in individuals with acute neurologic conditions. Information to assist in test result interpretation, such as MDC and MIC, is available for individuals with acute, chronic stable, and chronic progressive neurologic conditions.
- The 10mWT requires minimal equipment (eg, stopwatch and equipment for measuring walkway distance),

which is likely available in clinical settings or can be purchased at a low cost. There is a minimal time cost associated to administer the test (<5 minutes).

**Risk, Harm, and Cost:**

- No adverse events were documented in research studies.
- Administering the 10mWT has minimal risks, provided the patient's vital signs are monitored and appropriate guarding is used.

**Benefit-Harm Assessment:** Preponderance of benefit.

**Value Judgments:**

- The GDG emphasizes the importance of using standardized administration and scoring procedures for measuring patients in the clinic. While no single protocol has been used for the 10mWT, Quinn et al[11] and Steffen and Seney[10] described standardized procedures. The GDG recommends the protocol by Steffen and Seney because both comfortable and fast speeds were tested, providing an assessment of the patient's ability to alter gait speed. In addition, Steffen and Seney used a shorter walkway (the 10-m as compared with the 14m walkway used by Quinn et al), which may be more feasible in smaller spaces. This protocol has also been adapted by the ANPT CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standard procedures and, on an annual basis, establishing consistency within and among raters using the 10mWT.
- Walking safety may be more of a priority in acute and subacute rehabilitation to prepare for discharge, whereas walking speed may be a higher priority thereafter.
- Community ambulation requires the ability to ambulate at various speeds. The 10mWT enables the assessment of comfortable and fast walking; therefore, it is a useful measure to determine a patient's ability to resume community ambulation.

**Intentional Vagueness:** It is possible that authors of the studies reviewed used different administration procedures, resulting in some variability in the 10mWT protocols used among studies.

**Role of Patient Preferences:** Eighty-eight percent of consumers surveyed expressed that it was *important to improve walking* and 83% reported that *difficulty with walking* was a primary reason for seeking physical therapy.

**Exclusions:** The 10mWT is not appropriate for patients who do not have the capacity to walk. The GDG recommends that a score of 0 m/second be documented for patients who are unable to walk at a given point in time, but who have goals and the capacity to walk in the future.

**Quality Improvement:**

- Use of a single measure across clinical settings will facilitate communication among clinicians and enable assessment of changes in a patient's gait speed over time.
- Identifying a patient's capacity to return to specific activities requiring various gait speeds may be enhanced when using the 10mWT.

- Standardizing a gait speed measure for patients with neurologic conditions within and across clinical settings will enable comparative outcomes for quality improvement initiatives. Because scores may differ based on testing protocol, it may be difficult to compare data collected in different facilities unless the protocol is also specified. Individual organizations should use the CPG-recommended standardized protocol by Steffen and Seney[10] to assess aggregate data for their patients. In cases when the protocol cannot be used, the modifications to the OM administration should be documented.

**Implementation and Audit:**

- The GDG recommends that clinicians use the protocol by Steffen and Seney,[10] which has been adapted by the CPG KT Committee.
- For patients who are unable to walk at admission but have goals and the capacity to improve in this area, a score of 0 m/second should be documented to track patient change as ambulatory ability improves.
- The distance of the 10mWT is short and the use of assistive devices is permitted, which facilitates its use across functional levels and environments (eg, home). The type of device must be documented.
- Clinics and organizations should establish administration consistency within and among clinicians prior to using the 10mWT, and this should be repeated annually.

## Supporting Evidence and Clinical Interpretation (Table 10)

**Administration and Conditions:** The 10mWT involves measuring the time it takes for a person to walk the distance, with results typically reported in meters/second (m/s). The patient's ability to walk at both comfortable and fast speeds can be measured, and assistive devices can be used. Quinn et al[11] and Steffen and Seney[10] have described detailed administration procedures. Both used a walkway length of 10 m, but varied in their measurement of the entire walkway[11] versus the central 6 m.[10] Quinn et al[11] also measured the number of steps taken during the test. Both Quinn et al[11] and Steffen and Seney[10] administered 2 trials; Quinn et al[11] reported separate time data on each trial whereas Steffen and Seney[10] averaged the time from the 2 trials.

The 10mWT protocol by Steffen and Seney[10] is recommended by the GDG. This protocol assesses the time to the nearest 100th of a second to walk the central 6 m of a 10-m walkway at the patient's comfortable and fast walking speeds. The time starts when any part of the foot crosses the plane of the tapeline and ends when any part of the foot crosses the plane at the 6-m mark. Two trials are administered at the comfortable speed, with the instruction "walk at your own comfortable speed and stop when you reach the far line," followed by 2 trials at the fast speed, with the instruction "walk as fast as you can safely walk." The 2 trials, for each speed, are averaged and the 2 gait speeds are documented in meters/second. Use of an assistive device is permitted and should be documented. CPG KT Committee adaptations are located online at: http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg/core-measures.

**TABLE 10.** Evidence Table, 10 meter Walk Test

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCS AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **10 meter Walk Test acute samples** | | | | | | | |
| Lemay and Nadeau[48] | SCI; AID D | I | NA | NT | NT | NT | 0% |
| Perera et al[88] | Stroke | I | NA | NT | "Substantial meaningful change" (decline) = 0.01 m/s (anchor-stairs) or 0.10 m/s (anchor-walk block) (depending on anchor); SEM = 0.04 m/s | NT | NT |
| Scrivener et al[87] | Stroke—assessed within 48 h | I | NA | NT | NT | On admission 67%; on discharge 26% | NT |
| **10 meter Walk Test chronic stable samples** | | | | | | | |
| Hiengkaew et al[63] | Stroke | I | NA | Test-retest ICC = 0.96 | MDC = 0.18 m/s | NT | NT |
| Scivoletto et al[89] | SCI; mean time since onset 24 mo; 20 nontraumatic; 12 tetraplegia, 25 paraplegic; 35 AIS D, 2 AIC C | I | NA | Intrarater ICC = 0.98-99; interrater ICC = 0.95-0.98 | NT | NT | NT |
| **10 meter Walk Test chronic progressive samples** | | | | | | | |
| Nilsagård et al[78] | MS; EDSS 3-6 | I | NA | Test-retest ICC = 0.97 | Smallest % difference change = −23/+30 | NT | NT |
| Paltamaa et al[66] | MS; median EDSS 2.0 (0-6.5); 88% relapsing-remitting | NA (MIC only) | NA | NT | MIC self-paced-deterioration (patient anchor) = −0.14 m/s; MIC fast-paced deterioration (clinician anchor) = −0.19 m/s; MIC fast-paced deterioration (patient anchor) = −0.11 m/s | NT | NT |
| Quinn et al[11] | HD from premanifest to late state | I | NA | Test-retest ICC = 0.92-0.97 | MDC-pre-manifest HD = 0.23 m/s; manifest HD = 0.34 m/s; early-stage HD = 0.20 m/s; middle-stage HD = 0.46 m/s; late-stage HD = 0.29 m/s | NT | NT |
| Steffen and Seney[10] | PD; Hoehn and Yahr 1-4 (median 2) | I | NA | Test-retest ICC = 0.96 (comfortable speed) -0.97 (fast speed) | MDC comfortable = 0.18 m/s; fast = 0.25 m/s | NT | NT |

*(continues)*

**TABLE 10.** Evidence Table, 10 meter Walk Test *(Continued)*

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCS AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| Stolwijk-Swuste et al[80] | Postpolio with mean duration of new symptoms 10.3 y | I | NA | Test-retest ICC = 0.95 (preferred and maximum) | SDC = 1.9 m/s (preferred); 1.7 m/s (fast) | NT | NT |
| **10 meter Walk Test mixed chronic stable and chronic progressive samples** | | | | | | | |
| Rossier and Wade[90] | Various neurologic conditions; primarily BI (8); stroke (21); tumor (3); myelopathy (3); and HD (3) | I | NA | Test-retest ICC = 0.93 | NT | NT | NT |

Abbreviations: AIC, Akaike information criterion; AIS, American Spinal Injury Association Impairment Scale; BI, brain injury; EDSS, Expanded Disability Status Scale; HD, Huntington's disease; ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; MS, multiple sclerosis; NA, not applicable; NT, not tested; PD, Parkinson disease; SCI, spinal cord injury; SDC, smallest detectable change.

**Populations:** Ten level I studies on the 10mWT across all categories were reviewed: 3 acute (1 SCI[48] and 2 stroke[87,88]), 4 chronic progressive (1 MS[78], 1 HD[11], 1 PD[10], and 1 postpolio[80]), 2 chronic stable (stroke[63] and SCI[89]), and a mixed sample with acute stable, and chronic progressive conditions.[90] Meaningful change data have been reported in acute (stroke)[88] and chronic progressive (MS[66]) populations. Reliability has not been determined in acute neurologic conditions. Floor and ceiling effects have not been studied in individuals with chronic progressive and chronic stabile neurologic conditions.

**Psychometric Data:** *Reliability:* Intrarater (ICC = 0.98-0.99) and interrater (0.95-0.98) reliabilities were reported in one study in participants with SCI (chronic stable).[89] Test-retest reliability was established (ICC = 0.96) in patients with stroke (chronic stable).[63] Four studies examined test-retest reliability in individuals with chronic progressive conditions, including HD (ICCs ranged from 0.92 to 0.97 across manifestations of HD),[11] MS (ICC = 0.97),[78] PD (ICC = 0.96 and 0.97 for comfortable and fast speeds, respectively),[10] and postpolio (ICC = 0.95 for both preferred and maximum speeds).[80] In a mixed population of chronic stable and chronic progressive participants, the test-retest reliability was ICC = 0.93.[90] Collectively, these studies indicate excellent reliability of the 10mWT.

Only one study assessed interrater and intrarater reliability; this emphasizes the importance of establishing the consistency within and among clinicians within their own practice.[89] The high test-retest reliability across individuals with various neurologic conditions suggests that the 10mWT can be administered with consistent results across 2 time periods. No article established the reliability of the 10mWT in individuals with acute neurologic conditions. The reason for the lack of focus on speed in the acute phase may be related to a higher priority and emphasis on walking recovery and patient safety.

*SEM, MDC, MCID, Ceiling, and Floor Effects:* Data to assist in interpretation and measuring changes were reported in chronic stable (stroke; MDC = 0.18 m/s)[63] and chronic progressive conditions, including MS (smallest % difference change = −23/+30),[78] HD (MDC = 0.20 m/s to 0.46 m/s across HD manifestations),[11] PD (MDC = 0.18 m/s for comfortable and 0.25 m/s for fast speeds),[10] and postpolio (smallest detectable change [SDC] = 1.9 m/s for preferred and 1.7 m/s for fast speeds).[80] A measurement error rating score of *strong* (?) was assigned to each study, due to the lack of MIC/MCID data. "Substantial meaningful change" and SEM data were established in acute stroke ("substantial meaningful change" decline = 0.01-0.10) depending on the anchor used.[88] MIC was determined in MS[78] (chronic progressive) (MIC = −0.11 to −0.19 m/s) depending on the anchor. Values for MDC vary across patient populations and within a given neurologic condition as can be seen by reviewing our evidence table. Similarly, MIC values vary depending on the selected anchor.[66] Thus, clinicians should avoid generalizing the results of one patient population to another when considering MDC and MIC. These data can assist clinicians when interpreting results of a patient's 10mWT.

**Related Outcome Measures:** The Rivermead Mobility Index (RMI) (see Supplemental Digital Content 13, Appendix 11, available at: http://links.lww.com/JNPT/A226) is a

measure that examines balance, transfers, and gait. It includes 1 performance-based item and 14 self-report items. Five level I studies on the RMI included 2 in acute stroke, reporting on internal consistency (Cronbach α = 0.93)[91] and interrater reliability (ICC = 0.92).[92] In acute stroke, there is a floor effect (30%) at admission to inpatient rehab,[91] but not at 5 weeks. Hsueh et al[92] reported a floor effect at 14 days (40.4%), but not at 30 and 90 days; no ceiling effect was found. Test-retest reliability has been established in chronic stable (stroke; ICC = 0.96),[93] chronic progressive (HD; ICC ranged from 0.81 to 0.98 across HD manifestations),[11] and a mixed chronic stable and chronic progressive group (ICC = 0.96).[90] A smallest real difference of 2.2 was reported in stroke (chronic stable)[91] and chronic progressive populations, with MDCs ranging from 1 to 5 across HD manifestations.[11] One level II study established an SEM of 0.49 in MS.[94] Although RMI data are available across categories, the RMI is composed of 15 items, only 5 of which pertain to gait (on level, unlevel, and stair surfaces). Thus, the RMI is not solely a measure of gait. Because consumers reported that gait was of importance, the GDG selected a gait-specific measure for the core set. Hence, the RMI was not included.

The Timed 25 Foot Walk (see Supplemental Digital Content 14, Appendix 12, available at: http://links.lww.com/JNPT/A227) is a measure of gait speed (eg, the time to walk 25 ft). Eight level I studies on persons with MS (chronic progressive) establish its reliability in this population, with intrarater and interrater ICC values of 0.98 and 0.99, respectively.[95] Six studies[96-101] established test-retest ICC values ranging from 0.92[96] to 0.991.[97] In addition, MIC values (ranging from −0.01 to −3.55 seconds)[102] have been reported, as have SEM, MDC, and MDC% (= 1 second, 2.7 seconds, and 36%, respectively).[97] While the Timed 25 Foot Walk could have broad applicability, there is less evidence overall to support its use across populations as compared with the 10mWT.

The Walk-12 is a self-report walking measure that assesses the impact of a person's neurologic condition on walking capability. One level I study reported internal consistency (Cronbach α = 0.94), and floor (21.7% at admission and 0.9% at discharge) and ceiling effects (0.9% at admission and 0% at discharge) in a mixed chronic stable and chronic progressive sample.[103] Further research would be beneficial, as the Walk-12 would complement the performance-based measures of gait included in the CPG.

**R. Research Recommendation 7:** Studies are needed to explore the reliability and clinically important change (eg, MCID) of the 10mWT in individuals with acute neurologic conditions. Clinically important change should also be determined in chronic stable conditions. Studies to determine the presence of floor and ceiling effects should be conducted in persons with chronic progressive and chronic stable conditions.

**R. Research Recommendation 8:** Studies are needed to examine the Walk-12 in individuals with acute, chronic progressive, and chronic stable neurologic conditions. Properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**B. Action Statement 5: WALKING DISTANCE ASSESSMENT.** Clinicians should use the 6 Minute Walk Test (6MWT) for adults with neurologic conditions who have goals to improve walking distance and the capacity to change in this area. The 6MWT should be administered (per the Quinn et al[11] protocol as adapted by the CPG KT Committee) under the same test conditions at admission, and discharge, and when feasible, between these periods for patients with:

- Acute conditions: Evidence quality: V; recommendation strength: best practice
- Chronic stable conditions: Evidence quality: I; recommendation strength: moderate
- Chronic progressive conditions: Evidence quality: I; recommendation strength: strong

**Aggregate Evidence Quality and Strength:** Level I; moderate. Based on 5 level I studies, reporting both reliability and/or data to assist in measuring changes in chronic progressive conditions; 3 level I studies in chronic stable populations that reported reliability, but no data to assist in measuring change; and, in acute populations, 1 level I study reporting "substantial meaningful change" and SEM, but no studies that examined reliability (see Supplemental Digital Content 7, Appendix 5, available at: http://links.lww.com/JNPT/A220).

**Benefits:**
- The 6MWT demonstrates excellent reliability in patients with chronic progressive and chronic stable neurologic conditions.
- Data to assist in measuring change (eg, MIC, SEM, and MDC) have been assessed in individuals with chronic progressive neurologic conditions, and "substantial meaningful change" and SEM are available for individuals with acute conditions.
- The 6MWT has high clinical feasibility: it requires minimal equipment typically available in most settings and can be used for patients who walk with assistive devices. Only one trial is needed, limiting the time to administer the 6MWT.[104] Standardized procedures for test administration exist, as discussed later.
- Initial costs of purchasing equipment (eg, stopwatches, cones, and distance measuring device) are minimal and equipment is likely available in most settings. The time to instruct the patient and administer the test is less than 10 minutes, which can be minimized if the location and landmarks for conducting the test are standardized within each clinical setting.

**Risk, Harm, and Cost:**
- No adverse events were documented in research studies reviewed for this CPG.
- Administering the 6MWT has minimal risks, provided the patient's vital signs are monitored and appropriate guarding is used.

**Benefit-Harm Assessment:** Preponderance of benefit.

**Value Judgments:**
- The GDG emphasizes the importance of measurement reliability. Various protocols have been used for the 6MWT. To standardize administration and scoring, the GDG recommends the protocol described by Quinn et al.[11] This protocol has also been adapted by the

ANPT CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standard procedures and, on an annual basis, establishing consistency within and among raters using the 6MWT.

- Home and community ambulation requires the ability to walk for lengthy periods and distances. The 6MWT can help determine a patient's ability to resume activities requiring home and community ambulation.

**Intentional Vagueness:** The GDG assigned an aggregate quality rating of *moderate* to this action statement because of the lack of data to assist in measuring change (eg, MDC, SEM, and MCID) in acute and chronic stable conditions, and the absence of reliability data in acute populations.

**Role of Patient Preferences:** Eighty-eight percent of consumers surveyed reported that it was *important to improve walking*, and 83% reported that gait difficulties were a primary reason for seeking physical therapy.

**Exclusions:**

- The 6MWT is not appropriate for patients who do not have the capacity to walk. The GDG recommends that a score of 0 m be documented for patients who are unable to walk at a given point in time, but who have goals and the capacity to walk.
- The 6MWT may have limited feasibility in certain settings, such as a hospital room or home environment with limited walkway space or fixed environmental barriers. Thus, clinicians will need to determine the feasibility and appropriateness of the 6MWT in specific situations. If unable to administer due to limited feasibility, the clinician should document "unable to administer" and provide an explanation in the patient's medical record.

**Quality Improvement:**

- Use of a single measure across clinical settings will facilitate communication among clinicians and more accurately reflect changes in a patient's walking endurance over time.
- A clinician's ability to determine a patient's capacity to return to activities requiring ambulation over long distance (eg, community settings) may be enhanced by using the 6MWT.
- Standardizing a walking endurance OM for patients with neurologic conditions within and across clinical settings will enable comparative outcomes for quality improvement initiatives. Because scores may differ based on pathway, it may be difficult to compare data collected in different facilities.

**Implementation and Audit:**

- The GDG recommends that clinicians use the protocol described by Quinn et al[11] and adapted by the CPG KT Committee described later under Administration and Conditions. The recommended walkway length of 12 m is recommended for use by the GDG as longer walkways (eg, 30-m recommended by the American Thoracic Society)[105] are unlikely to be feasible in all environments (eg, small clinics or a patient's home). A shorter walkway length may facilitate continued administration of the 6MWT as a patient transitions from one service to another (eg, inpatient rehabilitation to home).

- Any deviation from the recommended protocol, including use of encouragement and physical assistance, should be documented.[106]
- For patients who are unable to walk at admission but have goals and the capability to improve ambulatory capability, a score of 0 m should be documented. This will capture changes over time as the patient's ambulatory capability improves.
- Only one trial of the 6MWT is necessary, as there is no practice effect when administering 2 trials.[104]
- Clinics and organizations should establish administration consistency within and among clinicians prior to using the 6MWT, and this should be repeated annually.

## Supporting Evidence and Clinical Interpretation (Table 11)

**Administration and Conditions:** The 6MWT measures the distance an individual can walk in 6 minutes. A systematic review of timed walking tests for persons with stroke identified 36 protocols for the 6MWT.[106] Studies varied in regard to walkway lengths (ranging from 10 to 85 m), shape (rectangular, oval, and circular), and tested speed (fast vs comfortable). The use of encouragement during the administration of the 6MWT varied and the impact is unclear.

Only Quinn et al[11] described standardized procedures for the 6MWT and the protocol recommended by the GDG. The test is performed in a 12-m-long straight and unobstructed walkway located in a quiet hallway or open area. A turn-around point should have clear markings at each end, about 124-cm wide (eg, 2 cones width). The patient should be well rested before this test. With the patient seated, the test is explained as specified by Quinn et al,[11] contraindications are checked, and resting heart rate is measured. The patient is instructed to walk up and down the walkway continuously without slowing, as able, for 6 minutes. Mobility aids may be used and must be documented. The patient stands and resting dyspnea (using the Borg scale) is measured. Encouragement (eg, "you're doing a good job and you have 5 minutes left) is given after each minute of the test; no other communication should occur during the test. The patient may rest at any time, but the stopwatch remains running and the number of rests and the total rest time are recorded. Distance in meters, walked at 1, 3, and 6 minutes, is recorded, as is the patient's heart rate before and after the test.

Various walkway lengths, ranging from 10 to 50 m, have been used.[89] Pathway distance has been shown to impact distance walked, with longer walkways resulting in greater distances walked,[89] suggesting the importance of using a consistent pathway within and across patients in a given clinical setting.

Administration procedures for the 6MWT are clinically feasible with minimal low-cost equipment required (eg, stopwatch and equipment for measuring walkway distance), typically available in most clinical settings. Patients may use assistive devices during the 6MWT, which enables use of the measure across patients at various functional levels. Only one trial is required, as there is no practice effect of 2 trials.[104]

**TABLE 11.** Evidence Table, 6-Minute Walk Test

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| **6-min Walk Test acute samples** | | | | | | | |
| Perera et al[88] | Stroke; subacute | I | NA | NT | "Substantial meaningful change" = 21 m (anchor-stairs); = 54 m (anchor-walk block); SEM = 22 m | NT | NT |
| **6-min Walk Test chronic stable samples** | | | | | | | |
| Liu et al[104] | Stroke | I | NA | Test-retest ICC = 0.98 | NT | NT | NT |
| Ng and Hui-Chan[81] | Stroke | I | NA | Test-retest ICC = 0.98 | NT | NT | NT |
| Scivoletto et al[89] | SCI | I | NA | Intrarater ICC = 0.99; Interrater ICC = 0.99 | NT | NT | NT |
| **6-min Walk Test chronic progressive samples** | | | | | | | |
| Baert et al[102] | MS; mild to severe | I | NA | NT | MIC improvement (patient anchor): 21.56 m (whole group); 26.86 m (EDSS ≤4; 17.39 m (EDSS 4.5-6.5); $SRC_{individual}$ 67.22 m (whole group); 42.86 m (EDSS ≤4; 75.42 m (EDSS 4.5-6.5). MIC improvement (clinician anchor) = 9.06 m (whole group); 6.90 m (EDSS ≤ 4; 9.87 m (EDSS 4.5-6.5); $SRC_{individual}$ = 68.32 m; (whole group); 56.53 m (EDSS ≤4; 73.98 m (EDSS 4.5-6.5) | NT | NT |
| Learmonth et al[97] | MS (79% relapsing remitting; 12% secondary progressive; 9% primary progressive); mean disease duration 11.8 (0.5-32) y; mean EDSS 3.5 (range 0-6.5) | I | NA | Test-retest ICC = 0.959 | SEM = 32 m; MDC = 88 m; MDC = 20% | NT | NT |
| Motl et al[98] | MS; 82% RR | I | NA | Test-retest ICC = 0.959 | NT | NT | NT |

*(continues)*

**TABLE 11.** Evidence Table, 6-Minute Walk Test (*Continued*)

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| Paltamaa et al[66] | MS; median EDSS 2.0 (0-6.5); 88% relapsing-remitting | NA (MIC only) | NA | NT | MIC–deterioration (clinician anchor) = −55.06 m; MIC–deterioration patient anchor) = −53.35 m | NT | NT |
| Quinn et al[11] | HD from premanifest to late state | I | NA | Test-retest ICC = 0.86-0.98 | MDC–premanifest HD = 39.22 m; manifest HD = 86.57 m; early-stage HD = 56.6 m; middle-stage HD = 126.14 m; late-stage HD = 70.65 m | NT | NT |
| Steffen and Seney[10] | PD; Hoehn and Yahr 1-4 (median 2) | I | NA | Test-retest ICC = 0.96 | MDC = 82 m | NT | NT |

Abbreviations: EDSS, Expanded Disability Status Scale; HD, Huntington's disease; ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; MIC, minimal important change; MS, multiple sclerosis; NA, not applicable; NT, not tested; PD, Parkinson disease; RR, risk ratio; SCI, spinal cord injury; SEM, standard error of measurement; SRC, smallest real change.

**Populations:** The 6MWT is appropriate for use in patients with any neurologic condition. Nine level I studies reported data on the 6MWT, including 5 samples with chronic progressive conditions (1 HD,[11] 1 PD,[10] and 3 MS[97,98,102]), 3 samples with chronic stable conditions (1 SCI[89] and 2 stroke[81,104]), and 1 in acute populations.[88] One study reported "substantial meaningful change" and SEM in acute (stroke) populations;[88] another reported MIC in chronic progressive (MS)[66] populations.

**Psychometric Data:** *Reliability:* Intrarater and interrater reliability (both ICCs = 0.99) were reported in participants with SCI (chronic stable).[89] Test-retest reliability has been established in chronic progressive conditions, including HD (ICCs ranged from 0.86 to 0.98 across manifestations of HD),[11] PD (ICC = 0.96),[10] and MS (ICC = 0.959).[97,98] Two studies established ICCs = 0.98 in participants with stroke (chronic stable).[81,104] Collectively, these studies indicate excellent reliability of the 6MWT, with the great majority achieving the preferred reliability of 0.90 or better.

Only one study[89] assessed both interrater and intrarater reliability; this emphasizes the importance of establishing the administration consistency within and among clinicians within their own practice. The high test-retest reliability across participants with various neurologic conditions suggests that the 6MWT can be administered with consistent results across 2 time periods. The reliability of the 6MWT in individuals with acute neurologic conditions was not assessed in any study.

*SEM, MDC, MCID, Ceiling, and Floor Effects:* "Substantial meaningful change" and SEM data have been reported in participants with acute stroke.[88] SEM and/or smallest real change (SRC) data have been reported in individuals with chronic progressive conditions, including HD,[11] PD,[10] and MS.[97,102] Paltamaa et al[66] reported MIC data in persons with MS. Only one study[102] was rated *strong*, as both SRC and MIC data were reported; the other studies[10,11,97] were rated s*trong* (?) due to the lack of MIC/MCID data. Nevertheless, data exist to assist clinicians when determining changes in a patient's 6MWT score. Values for interpreting change (eg, MDC and MIC) can vary across patient populations, within a given neurologic condition, or depending on the anchor used, as is seen in Table 11. This suggests that clinicians should avoid generalizing the results of one patient population to another population when considering data to assess patient change.

Data for use in assessing patient change have not been reported in individuals with chronic, stable neurologic conditions. No studies reported data for floor or ceiling effects in any category, or reliability in acute populations, although "substantial meaningful change" and SEM data exist in persons with acute stroke.[88] Therefore, the 6MWT should be used with caution in individuals with chronic stable neurologic conditions.

**Related Outcome Measures:** The 2-Minute Walk Test (2MWT) (see Supplemental Digital Content 15, Appendix 13, available at: http://links.lww.com/JNPT/A228) was reviewed, as it is a clinically feasible measure of walking distance and has applicability across patients with neurologic conditions, especially those with fatigue (eg, persons with MS). Four

level I studies provide data on persons with stroke (chronic stable), including test-retest reliability (ICC = 0.98) and MDC (13.4 m).[63] In chronic progressive samples, excellent test-retest reliability (ICC = 0.95) exists in persons with postpolio,[80] and MIC (6.81 m) and SRC (26.64 m) have been established in MS.[102] Rossier and Wade[90] established the test-retest reliability (ICC = 0.97) in a mixed chronic stable and chronic progressive sample.[90] No studies reported data on the 2MWT in acute populations. The 2MWT has comparable test-retest reliability and the availability of data to interpret change, but there was less evidence overall to support its use across populations than the 6MWT.

**R. Research Recommendation 9:** Studies are needed to determine the intrarater and interrater reliability, and clinically important change (eg, MCID), of the 6MWT in individuals with acute neurologic conditions. Data to assist in measuring change (eg, MDC, SEM, and MCID) are needed in individuals with acute and chronic stable neurologic conditions.

**P. Action Statement 6: TRANSFER ASSESSMENT.** Clinicians should document the transfer ability of patients who have goals to improve transfers and have the capacity to change. Documentation should include the type of transfer, level of required assistance, equipment or context adaptations, and time to complete. In patients who have goals and the capacity to improve sit-to-stand transfers, the 5 Times Sit-to-Stand (5TSTS) may be used. The 5TSTS and documentation of other transfers may be administered under the same test conditions using the protocol recommended by the CPG KT Committee at admission, discharge, and, when feasible, between these periods for adult patients with neurologic conditions. (Evidence quality: V; recommendation strength: best practice).

> **Aggregate Evidence Quality and Strength:** Level V; best practice. Based on the GDG clinical expertise, informed by related evidence and the results of the clinician survey.
> **Benefits**:

- Use of the 5TSTS will standardize one aspect of transfer skill across patients and may provide information about the methods a patient uses to complete the sit-to-stand transfer.
- Initial costs of purchasing equipment (eg, stopwatches) are minimal and the required equipment (eg, standard chair) is commonly available in clinical settings. The time to administer the test is less than 5 minutes.

> **Risk, Harm, and Cost**:

- No adverse events relative to the use of the 5TSTS were documented in studies reviewed for this CPG.
- Using an OM of transfers may extend the length of the session.

> **Benefit-Harm Assessment:** Preponderance of benefit.
> **Value Judgments**:

- 77% of clinicians surveyed indicated that transfers are an *important* construct to measure.
- Transfers (ie, moving from one position to another, such as sit to stand or wheelchair to mat) are a fundamental skill for daily life and an important component of the physical therapy care provided to patients with neurologic conditions.

- The use of OMs of transfers to assess and monitor changes in individuals with neurologic conditions reflects best practice and is consistent with the APTA Guide to PT Practice.
- The GDG emphasizes the importance of using standardized administration and scoring procedures for measuring patients in the clinic. While there is not a universally accepted protocol for the 5TSTS, we recommend that each clinical site adopt the testing protocol developed by the CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg). We recommend review of the standard procedures and, on an annual basis, establishing consistency within and among raters using the 5TSTS.

> **Intentional Vagueness**:

- No single transfer OM had sufficient literature to support a strong or moderate recommendation for the core set; the 5TSTS received a best practice recommendation.
- Clinicians and organizations need to determine the feasibility and utility of using an OM to measure transfers in view of their patient population, facility-specific requirements and resources, and payer requirements.

> **Role of Patient Preferences**: Consumers of neurologic physical therapy surveyed indicated that the use of standardized OMs is *very important* (58%) or *important* (35%) to their care.
> **Exclusions**: None.
> **Quality Improvement:** Consistent use of a transfer OM may enable clinicians and administrators to monitor the patient's change at an individual, unit, organization, or system level.
> **Implementation and Audit**:

- Procedures for administering the 5TSTS should be standardized for use by clinicians in the facility. The GDG recommends the standard procedure developed by the CPG KT Committee for administration of the 5TSTS. The procedure is located on the ANPT Web site (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg).
- Clinics and organizations should establish administration consistency within and among clinicians prior to using the 5TSTS, and this should be repeated annually.

### Supporting Evidence and Clinical Interpretation (Table 12)

**Administration and Conditions:** The 5TSTS measures the time it takes an individual to transfer from a seated to a standing position and back to sitting 5 times. A patient is instructed to sit with arms folded across their chest and with back against the chair. Patients with stroke may have their impaired arm at their side or in a sling. Chair heights of 43 to 45 cm have been reported in the literature. The patient is instructed to stand up and return to sitting 5 times as quickly as possible. Timing starts when the therapist says "go" and ends when the patient's body touches the chair following the fifth repetition. Administration procedures for the 5TSTS are clinically feasible with minimal low-cost equipment required (eg, stopwatch and chair), typically available in most clinical settings.

**TABLE 12.** Evidence Table, 5 Times Sit-to-Stand

| AUTHOR | PRIMARY POPULATION AND IMPAIRMENT LEVEL (IF AVAILABLE) | LEVEL OF EVIDENCE | INTERNAL CONSISTENCY | RELIABILITY (TYPE, RESULTS) | STANDARD ERROR; MDCs AND MCIDs | FLOOR EFFECTS | CEILING EFFECTS |
|---|---|---|---|---|---|---|---|
| 5 Times Sit-to-Stand chronic progressive samples | | | | | | | |
| Paul et al[79] | Parkinson disease | I | NA | Test-retest ICC = 0.91 | SEM = 0.6 s | NT | NT |
| Abbreviations: ICC, intraclass correlation coefficient; MCID, minimal clinically important difference; MDC, minimal detectable change; NA, not applicable; NT, not tested; SEM, standard error of measurement. | | | | | | | |

**Populations:** The 5TSTS has been studied in individuals with chronic progressive conditions (PD).[79]

**Psychometric Data:** *Reliability:* One level I study reported test-retest reliability (ICC = 0.91) in chronic progressive conditions (PD).[79] Reliability has not been assessed in individuals with acute or chronic stable populations; therefore, the 5TSTS should be used with caution in these groups.

*SEM, MDC, MCID, Ceiling, and Floor Effects:* SEM was reported to be 0.6s in individuals with chronic progressive conditions (PD)[79]; however, data are lacking to assist with measuring changes in acute or chronic stable neurologic conditions. No studies reported data for floor or ceiling effects in any category. Therefore, the 5TSTS should be used with caution in individuals with acute and chronic stable neurologic conditions.

**Related Outcome Measures:** The Rivermead Mobility Index-Modified (RMI-Mod) and the 30-second Chair Stand Test (30SCST) were reviewed for this CPG. The 30SCST was excluded because it did not have at least one article on reliability and data to interpret changes in neurologic populations.

Three articles supported the RMI-Mod,[107-109] and these included participants with acute stroke[107,109] and a mixed population of adults with acute and chronic progressive, but not chronic stable neurologic conditions (see Supplemental Digital Content 16, Appendix 14, available at: http://links.lww.com/JNPT/A229).[108] All articles examining the RMI-Mod were level I articles and reported internal consistency values between 0.80 and 0.96 and reliability between 0.93 and 0.99. Data to assist with measuring the change is lacking. While the RMI-Mod met the initial criteria of at least 75% of the test items matching the constructs of interest, only 50% of the test items matched the construct of transfers. For these reasons, the RMI-Mod was not recommended as a transfer OM.

**R. Research Recommendation 10:** Studies are needed that explore the feasibility and psychometric properties of the 5TSTS to objectively describe the transfer abilities of adults with neurologic conditions, especially those other than individuals with PD, across the continuum of care and spectrum of acuity. Further study of the 30SCST is warranted, particularly relative to reliability and data to interpret changes in individuals with neurologic conditions.

**P. Action Statement 7: DOCUMENTATION OF PATIENT GOALS.** Clinicians should document patient-stated goals and monitor changes in individuals with neurologic conditions using an OM such as the Goal Attainment Scale (GAS), reporting the task, the performance conditions, and the time to complete or level of independence desired. Documentation of patient goal measures should be administered under the same test conditions at least 2 times, at admission and discharge, and, when feasible, between these testing periods. (Evidence quality: V; recommendation strength: best practice)

**Aggregate Evidence Quality and Strength:** Level V; best practice. Based on the clinical expertise of the GDG and informed by related evidence and the results of the clinician survey.

**Benefits:**
- Seventy-nine percent of PTs surveyed for this CPG indicated that patient-stated goals are an *important* construct to measure.
- Using an OM of patient-stated goals will provide an opportunity for patients and clinicians to share their beliefs and values.
- An OM that assesses a patient's goals may capture activities or constructs not included in other OMs, but are important to the patient.
- Use of an OM of patient-stated goals may assist clinicians in identifying and addressing discrepancies between perceived and actual performance.

**Risk, Harm, and Cost:** No adverse events were documented in studies reviewed for this CPG.

**Benefit-Harm Assessment:** Preponderance of benefit.

**Value Judgments:** The GDG believes that the use of OMs that assess and monitor changes in patient-stated goals in patients with neurologic conditions:
- Facilitates a patient-centered approach by integrating the patient's goals, priorities, and values into the plan of care.
- Will encourage patient engagement in the rehabilitation process.

**Intentional Vagueness:** No patient-stated goal OM had sufficient literature to support use across adults with neurologic conditions.

**Role of Patient Preferences:** Using an OM of patient-stated goals will allow patients to clearly state their preferences for the focus of physical therapy.

**Exclusions:** In some situations, such as patients with impaired consciousness, cognition, and/or communication,

it may be challenging to ascertain the patient's goals. A caregiver may be able to provide a proxy response.[110-112]
**Quality Improvement:** Consistent use of a patient-stated goal OM may enable clinicians to monitor the patient's perspective of change, and administrators to monitor the degree to which patients perceive change at an individual, unit, organization, or system level.
**Implementation and Audit:**
- Because a specific patient-stated goal OM was not recommended, the GDG recommends that each organization select an appropriate OM to assess patient-stated goals in regard to its patient population, facility-specific requirements, and resources. The GAS, a measure that was assessed during the CPG review process, has been studied in other populations (eg, pediatric and geriatric) and may be applied to adults with neurologic conditions.
- Administration procedures (eg, interview structure and use of a proxy) for the organization's chosen patient-stated goal OM could be standardized for use in the facility. Standardization regarding assessment and documentation of this construct should include reporting the task, the performance conditions, and the time to complete or level of independence desired. Patient goals should be assessed at least 2 times, at admission and discharge, and preferable in between these time periods under the same test conditions.
- When a discrepancy exists between perceived goals and actual performance or capacity, clinicians should provide education for the patient and caregiver and review the goal expectations.

## Supporting Evidence and Clinical Interpretation

**General Overview:** Patients' and clinicians' health beliefs frequently lack agreement, affirming the need for discussions about goals and shared decision-making with patients.[113] Many OMs make the theoretical assumption that all clients have similar goals leading to the challenge of capturing the unique goals of individual clients.[114] OMs have been developed, which allow the clinician and the patient to collaboratively and systematically establish individualized goals and reach agreement on the scaling of these goals.

**Patient-Stated Goals OM Considered in This CPG:** Three measures of patient-stated goals, the GAS, Canadian Occupational Performance Measure, and Patient-Specific Functional Scale, were reviewed for this CPG. The Canadian Occupational Performance Measure was excluded because it is proprietary and requires payment for use. The Patient-Specific Functional Scale was excluded because it did not have at least one citation each to support reliability and assessment of change over time. Two citations for the GAS were identified. One citation was excluded, as the subject population included a mixed geriatric population, rather than participants exclusively with neurologic conditions.[115] A final citation used the GAS with a neurologic population (brain injury and stroke); the standardized response mean (2.2) was reported, but data were lacking for reliability.[116] One article reported on participants with MS, but failed to meet the sample size required for inclusion in this CPG;[117] others did not focus on adults with neurologic conditions.[118,119]

**R. Research Recommendation 11:** Studies should explore the feasibility and psychometric properties, including reliability and data to assist in interpreting change (eg, MDC and MCID/MIC) of the GAS and other OMs that capture the individual goals of adults with neurologic conditions across the continuum of care and spectrum of acuity.

**B. Action Statement 8: USE OF THE CORE SET OF OUTCOME MEASURES.** Clinicians should use and document the OMs in the core set to assess change over time. The core set includes the Berg Balance Scale (BBS), Functional Gait Assessment (FGA), Activities-specific Balance Confidence Scale (ABC), 10 meter Walk Test (10mWT), 6 Minute Walk Test (6MWT), and 5 Times Sit-to-Stand (5TSTS) and the recommended patient goal assessment for adults who are undergoing neurologic physical therapy. The core set should be administered with patients who have goals and the capacity to improve transfers, balance, and/or gait. In cases when a patient cannot complete one or more core set OMs (eg, a patient who is unable to walk; thus, cannot complete the 10mWT or 6MWT), a score of 0 should be documented. The patient goal assessment should be administered to all adults undergoing neurologic physical therapy. The core set should be administered under the same test conditions at least 2 times, at admission and discharge, and when feasible between these periods (Evidence quality: II; recommendation strength: moderate).

**Aggregate Evidence Quality and Strength:** Level I; moderate. Based on 41 level I studies for the 6 OMs collectively (ABC, Berg, FGA, 6MWT, 10mWT, and 5TSTS) and 1 level II moderate study (FGA). Level I studies provide moderate to strong evidence, supporting the use of the BBS, FGA, ABC, 10mWT, and 6MWT for patients with chronic stable and chronic progressive conditions. Best practice recommendations support the use of the 10mWT (2 level I studies) and the 6MWT (1 study reporting MIC) in patients with acute conditions. A best practice recommendation was made for the 5TSTS based on 1 level I study in patients with chronic progressive conditions. In addition, a best practice recommendation was made that clinicians document patient-stated goals and monitor changes using an OM. In the survey to determine the scope of the core set, the PTs indicated that balance (97%), gait (94%), patient-stated goals (79%), and transfers (77%) were important to address, and 94% of PTs indicated they were *willing* or *very willing* to use a core set of OMs.
The aggregate strength of moderate was given because the core set measures have not been studied collectively.
**Benefits:**
- Consumers of PT and clinicians were in agreement that the constructs of gait, balance, transfers, and patient-stated goals are important to assess. In addition, the recovery of balance, gait, and transfers facilitate improved independence for adults with neurologic conditions. Therefore, a core set of OMs that captures these constructs addresses the needs of patients and practitioners. A comprehensive examination of all constructs, for which a patient has goals and the capacity to improve in these goals, reflects best practice.

- Use of the core set OMs for all patients with neurologic conditions and in all settings will facilitate collection of practice-based evidence to compare interventions and programs.
- Use of the core set OMs across settings will facilitate measurement of patient progress over time and across the continuum of care. For example, as a patient moves from acute care to inpatient rehabilitation to outpatient services, or as a patient's neurologic condition changes over time due to recovery or its progressive nature, the core set will reflect performance changes for the highest priority domains.
- Results of the core set of OMs can facilitate a comprehensive examination of balance, gait, and transfers to assist with clinical decision-making, including the selection of treatment interventions, modification of the plan of care, and discharge decisions.
- Standardization of entry-level DPT and residency education that includes training on the core set.

**Risk, Harm, and Cost:**
- No adverse events relative to the use of any of the measures in the core set were reported in studies reviewed for this CPG.
- Organizational costs to administer the core set of OMs may include the cost to alter the medical record to include data fields, time for staff training and test administration, and the cost of testing forms and equipment.

**Benefit-Harm Assessment:** Preponderance of benefit.

**Value Judgments:** The GDG believes that the use of a core set of OMs will enhance patient outcomes because they will standardize measures across settings. The core set will contribute to the advancement of neurologic physical therapy through the development of a learning health system and the ability to do comparative effectiveness research.

**Intentional Vagueness:**
- The time frames for administration of the core set (eg, admission, interim, and discharge) may vary depending on facility-specific requirements and length of stay.
- The GDG recommends administration of the core set and sharing the measurement results with providers at the next level of care. This is particularly important when it is not feasible to administer the core set more than once within a given setting.
- The measures in the core set were assessed primarily in patients with central nervous system conditions. Therefore, clinicians should use caution when applying these measures to patients with peripheral nervous system conditions.
- Although evidence supports the use of each measure in the core set, the use of the measures collectively has not been studied.

**Role of Patient Preferences:**
- Consumers surveyed reported that OMs were *very important* (60%) and *somewhat important* (36%) to their care.
- Selection of the appropriate OMs for an individual patient should be based on a patient's prognosis and rehabilitation goals.

**Exclusions:**
- The OMs in the core set were assessed for reliability and the ability to measure change over time. They were not assessed for other purposes (eg, prediction or impairment classification).
- In an acute care setting, in situations where a patient's length of stay is short, or when the patient is abruptly discharged from a given setting, administration of the core set at interim and discharge time frames may not be feasible.
- If a patient does not have goals or a prognosis to improve in specific construct areas, OMs should not be collected in the specific goal areas. When an OM in the core set cannot be administered (eg, due to a patient's current abilities or the patient does not have the capacity to improve or goals in the area), the clinician should document that the OM was not administered and provide a rationale (eg, not applicable due to the patient's current and expected functional capability or not applicable due to a lack of related patient goals).

**Quality Improvement:**
- The core set will facilitate monitoring of an individual patient's status across time and settings, and the degree to which patients change in aggregate. The data collected could be used to increase transparency of outcomes; study clinician performance relative to patient outcomes and internal and external benchmarks; improve health care processes; and generate new knowledge.

**Implementation and Audit:**
- The leadership of health care facilities and organizations should prioritize use of the core set and actively support implementation.[120]
- Clinical facilities and organizations should standardize the administration procedures (eg, equipment, instructions, and scoring) of the core set. Efforts should be taken to standardize administration procedures and to determine the consistency within and among clinicians prior to using the core set OMs.
- Documentation of the core set should be standardized to incorporate the following designated fields into electronic health records: the BBS, FGA, Activities-specific Balance Confidence Scale, 10mWT, 6MWT, and 5TSTS. Fields to document the total score and individual items on the OM should be included. In addition, the following items may be documented when assessing transfers: transfer ability of patients who have goals to improve transfers and have the capacity to change, inclusive of type of transfer, level of required assistance, equipment or context adaptations, and time to complete. When documenting patient goals, the following items should be included: the task, the performance conditions, and the time to complete and/or level of independence desired.
- When a patient continues care at another level of service, the core set results should be shared between facilities/organizations.
- Organizations should audit documentation regularly to determine adherence to core set recommendations. If adherence levels are not acceptable, audit and feedback, use of other knowledge translation interventions, or quality improvement initiatives may improve routine administration of the core set.

## Supporting Evidence and Clinical Interpretation

The concept of a core set of OMs for use in neurologic rehabilitation has been discussed for over 10 years. The APTA EDGE task forces made condition-specific recommendations for use of OMs in practice.[121] The development, use, and benefits of core sets, including those organized by condition and construct, have been described.[16] Measurement core sets have been described/developed for clinical and research use with individuals with stroke,[106] MS,[122] cerebral palsy,[123] vertigo and dizziness,[124] and cerebellar ataxia.[125] Other authors have advocated for OM core sets organized by construct such as balance[126] or gait.[106] Most published core sets have been developed by a consensus approach, such as a Delphi process.[123,126,127] While a modest amount has been written in support of the development of OM core sets, the literature on the demonstrated benefits of use in physical therapy is extremely limited.[128] Therefore, research is needed on the impact of the core set on patients, organizations, and the profession.

The use of OMs, including a core set of OMs, will create the foundation for learning health care in adult neurologic physical therapy, as recommended by the Institute of Medicine.[22] The OMs in the core set have value individually as well as when used collectively in the care of adults with neurologic conditions. All OMs, with the exception of the 5TSTS, have documented evidence of strong internal consistency/reliability and data to assist in measuring change (eg, SEM, MDC, and MCID) from multiple level I articles across neurologic conditions and categories. Collectively, the core set OMs capture the client's status across constructs that both PTs and consumers indicated to be important or related to primary reasons for seeking physical therapy services. Furthermore, the use of patient goal assessment will provide standard reporting guidelines for patient goals. The core set will facilitate a comprehensive examination of important constructs in a patient's care and support decision-making, plan of care development, and achievement of outcomes collaboratively set by the patient and the clinician.

**R. Research Recommendation 12:** Studies are needed that explore the impact of using the core set of OMs on rehabilitation outcomes, including factors related to implementation (eg, time and cost). Studies should explore the impact of using the core set of OMs to support clinical decision-making across neurologic conditions and categories. Future measurement studies should be designed to meet the COSMIN requirements for excellent methodology with regard to sample size, design, and rigor of statistical analysis of psychometric properties.[8,9,40,42]

**R. Research Recommendation 13:** The CPG KT Committee is developing standardized administration procedures for all 6 OMs in the core set. Studies are needed to determine the psychometric properties of these protocols across acute, chronic progressive, and chronic conditions in clinical practice.

**P. Action Statement 9: DISCUSS OUTCOME MEASURE RESULTS AND USE COLLABORATIVE/ SHARED DECISION-MAKING WITH PATIENTS.** Clinicians should discuss the purpose of OMs, results, and how these results influence treatment options with patients undergoing neurologic physical therapy. Collaboratively, the clinician and the patient should decide how these data should inform the plan of care (Evidence quality: V; recommendation strength: best practice).

> **Aggregate Evidence Quality and Strength:** Level V; best practice. Based on the GDG clinical expertise and informed by the consumer survey results and references in other medical fields.
> **Benefits:** Discussing the results of OMs with patients may result in:
> - Patients being more informed and engaged in rehabilitation.
> - Better alignment of the plan of care with the patient's goals, preferences, and measurement results.
>
> **Risk, Harm, and Cost:**
> - No adverse events relative to the discussion of the results OMs were documented in the reviewed studies or in a Cochrane review on the use of decision aids (eg, interventions that support patients in shared decision-making) to inform patients about care.[129]
> - A discussion of the OM results may extend the length of the session. Decision aid use to support shared decision-making has been shown to mildly increase (<3 minutes) the length of a patient's consultation with a health care provider.[113]
> - When the results of OMs are not positive and/or patients have difficulty understanding the results, patients may experience stress/discomfort and the discussions may add time to the treatment session.
>
> **Benefit-Harm Assessment:** Preponderance of benefit.
> **Value Judgments:**
> - In a Cochrane review on decision aids (eg, interventions that support patients in shared decision-making), some benefits identified include increased participants' knowledge, accuracy of risk perceptions, improved alignment of values and care choices, and decreased decisional conflict from feeling uninformed.[129]
> - The GDG believes discussing the OM results and sharing (eg, collaboratively) decision-making would benefit patients undergoing neurologic physical therapy.
>
> **Intentional Vagueness:** The time frames (eg, admission, interim, and discharge) for clinicians discussing the results of OMs and sharing decisions with patients who have neurologic conditions may vary depending on facility-specific requirements, patient length of stay, etc.
> **Role of Patient Preferences:** The majority of the consumers surveyed reported that test results were *very important* (60%) or *important* (35%) to them.
> **Exclusions:** In some situations (eg, a patient with an impaired level of consciousness, cognition, or communication impairment), it may be challenging to discuss the results of OMs with a patient. A caregiver may be able to participate in these discussions and decisions as a proxy.
> **Quality Improvement:**
> - Mechanisms (eg, time and space for conversation) should be developed to enable clinicians to share OM-related information with patients and caregivers.

- Sharing OM results and their impact on the plan of care may help to:
- Engage and motivate a patient in his/her physical therapy.
- Facilitate shared decision-making regarding goals and the plan of care.
  **Implementation and Audit:**
- Organizations should develop procedures and documentation for the discussion of OM between the clinician and the patient. Articles are available to guide implementation of shared decision-making in clinical practice, and may be applied to rehabilitation clinics.[130]
- Education and training on methods to discuss OM results and share decision-making may be required.
- A routine audit and feedback of documentation should be performed to ensure adherence to the recommendations of sharing OM results and decision-making with the patient.

## Supporting Evidence and Clinical Interpretation

Shared decision-making is an approach in which patients and clinicians make decisions collaboratively using the patient's health information, their values and preferences, and the best available evidence. Patients are encouraged to consider examination and treatment options and communicate preferences. The clinician should collaborate with the patient to assist in selecting the best plan of care. This approach differs from one in which a clinician makes decisions on behalf of patients, and is intended to respect patient autonomy and promote engagement.[130,131] Sixty-percent of consumers surveyed for this CPG reported that test results were "*very important*" to them. However, 13% did not recall whether their PT conducted tests and 25% reported that tests were conducted only at admission and discharge, but not in between these 2 periods. It is possible that OMs were not consistently used in the patients' care, but these data may also indicate that the consumers were not consistently informed about the use of OMs. The majority of consumers reported that the PT discussed the purpose (80%) and results (76%) of the OMs used and that the PTs explained how the OM results informed the plan of care (53%). Only 37% reported being "*very satisfied*" with the information they received. The consumers were not asked whether they shared decision-making regarding the plan of care.

These data suggest that there is a need to improve the provision of OM-related information to patients and to share decision-making about the plan of care. Providing meaningful information and sharing decisions throughout each patient's episode of care ensure that needs are met and the patient understands the role of physical therapy in his/her health care. This is particularly important, as patients' and clinicians' health beliefs may lack agreement, confirming the need for shared decision-making between clinicians and patients.[113]

A recent Cochrane review[129] concluded that decision aids, which provide evidence-based information to inform patients and support shared decision-making, can have a positive effect on communication between the provider and the patient. Decision aids can inform patients and improve knowledge (high-quality evidence), increase the patient's involvement in care (moderate-quality evidence), and integrate a patient's values with care decisions (low-quality evidence). Although this review focused on decision aids for medical interventions, it may have relevance for rehabilitation practice. Similar outcomes (eg, enhanced patient involvement and knowledge) may be achieved by providing patients with explicit information about their OM results and collaboratively making decisions about their care.

**R. Research Recommendation 14:** Research is needed on the impact of discussing OM results and shared decision-making with patients receiving neurologic physical therapy, including the development and impact of OM-related information (eg, OM-related decision aids) on the understanding and involvement of a patient in his/her care and on the achievement of patient goals. Furthermore, studies should develop and test the use of decision aids that incorporate the core set.

## Limitations

There are several limitations to this CPG. As stated, this CPG focused only on OMs to assess patient change over time. Thus, other OM uses (eg, prediction) were not considered. When critically appraising the articles, the focus was on the strength of the psychometric properties of OMs, not available administration protocols. Our review of OMs reflected the name of the measure (eg, BBS and 10mWT), not the construct (ie, of balance or gait speed). Thus, it is possible that some articles that may have been identified by construct, rather than OM name, were not identified and reviewed. In addition, it is possible that authors of the studies reviewed used different administration procedures, resulting in some variability in the protocols used among studies.

## GUIDELINE IMPLEMENTATION RECOMMENDATIONS

*Overview:* Implementation of the action statements contained in this guideline is integral to the process of knowledge translation (KT). KT has been defined as "the dynamic and iterative *process* that includes the *synthesis, dissemination, exchange and ethically sound application* of knowledge to improve health, provide more effective health services and products, and strengthen the health care system."[132] This complex process is impacted by many variables and is most effective when efforts are multifaceted and sustained, and when they target barriers to the recommended practice. Efforts at the individual, organizational, and societal levels to support KT are critical to ensure rapid and successful CPG implementation. Organizations and clinicians should assess their own barriers and facilitators to using the CPG action statements and develop a KT plan that is tailored to overcome the identified barriers. The GDG considered the literature and input from key stakeholders related to barriers for the CPG (eg, time, cost, and training needed to administer the core set; equipment) when selecting OMs for the core set. The recommendations given next may facilitate adoption and successful use of the core set in practice. Use of KT frameworks can provide a theoretical foundation for implementation, and may lead to successful KT initiatives.[133-137]

*Recommendations for Health Care Organizations and Clinicians:* The GDG recommends that organizations adopt specific standardized practices related to use of this core set of OMs and documentation of patient goals in clinical practice.

- First, the core set should be used when a person undergoing neurologic physical therapy has goals and potential to improve balance, gait, or transfers.
- Patient goal documentation should adhere to the CPG recommendations. The OMs should be administered to a patient when evaluated in any setting. If a patient is unable to perform a test, but will likely be able to perform some or all of the OM at some point in the future, the patient should receive a zero on the initial test. This provides an opportunity to capture data at a later point in time, reflecting change that occurred.
- Follow-up measures should be administered at least twice, with ideal administration time being the middle of treatment and at discharge. While it is recommended to collect the core set at least once between the admission and discharge assessment, the decision to use the OMs for interim measurements is left to the discretion of the clinician and the organization. Factors such as length of stay, facility requirements, and reimbursement may impact the ability to administer the core set at times other than admission and discharge. However, an interim assessment will provide important information about whether the patient's status is changing during the episode of care and may inform intervention modifications. In cases when administration of the OMs multiple times is not feasible (eg, in acute care), the GDG recommends that the clinician administers the OM once and provide the measurement results to the next level of care.
- In health systems with several levels of care, the core set should be used throughout a patient's episode of care and measurement results should be provided to the next level of care.
- Clinicians should utilize data from the core OM set to describe progress to other health professionals (eg, letters to insurance companies, physicians, and team conference reporting).
- Documentation of the OMs should be standardized within the facility based on the recommended methods and incorporated into designated fields in the electronic health record.
- Information, such as SEMs, MDCs, and MCIDs, should be used to support decisions to alter the course of treatment and discharge from care.
- Implementation of this core set may require time for learning about the CPG and the recommended practices, comparing current practice with recommended actions, and creating a plan for CPG implementation within the organization.[138,139] Specific protocols for administering the core set have been recommended by the GDG and CPG KT Committee (http://www.neuropt.org/professional-resources/anpt-clinical-practice-guidelines/core-outcome-measures-cpg).

Organizations and clinicians should determine interrater and intrarater reliability of each core set measure annually and strive to achieve an agreement of more than 0.90 reliability.[140-142] OMs with a test-retest or interrater reliability of less than 0.70 should not be used for individual patients.[140,141] Establishing the reliability of clinicians in a clinical setting should facilitate consistent measurement of a patient's performance (eg, when more than one clinician conducts a given test on an individual patient) or when measures are taken over time (eg, at admission and discharge), and enhance a clinician's and organization's confidence in the OM results. In addition, increased reliability when using OMs may improve the clinician's ability to identify changes in function, reduce measurement error, and improve the development and modification of the plan of care. Training to ensure standardization of OM administration and skills assessment may enhance reliability.

To promote adoption of the core set, organizations should consider the use of KT interventions.[143] A copy of the CPG action statements should be kept in a location that is easy to reference. Equipment and space to administer the core set should be kept in an easily accessible location. Examination forms should be adapted to include facility-specific information, such as the location of equipment and local adaptation to testing paths, and electronic and printed versions should be made available. Initial training on administration of the core set, how to use data to guide decision-making, and methods to use core set data to collaboratively determine a plan of care with patients (eg, shared decision-making) may be required. This content should also be provided during new hire orientation. Audit and feedback[144] may facilitate adherence to the recommendation that OM administration occurs at admission and discharge, and preferably, at least once in between. Audit criteria should include adherence to recommended administration timing and documentation of OM interpretation and shared decision-making. Tools to assist with auditing will be developed by the CPG KT Committee, and added to the ANPT Web site. Incorporating a requirement to adhere to use of the core set into performance appraisals will promote the use of the core set as a clinical and professional expectation. Whenever possible, core set reminder systems and decision-support tools should be integrated into the electronic health system. These and other KT strategies may be used to promote adoption throughout a health care organization.

*Integration With EDGE Recommendations:* Six ANPT EDGE task forces predated the development of this CPG. The OM recommendations from those groups were focused on individuals with a specific neurologic condition (eg, stroke). It is the intent of the GDG that, when caring for an individual with a specific condition, clinicians integrate the core set with the recommendations from the relevant EDGE task force. The core set may be viewed as a "starting point" for measure selection, with additional condition-specific measures as recommended by the EDGE task force used to provide insight into issues specific to their patient's health condition.

*ANPT KT Taskforce Will Support CPG Implementation:* In collaboration with the GDG, the ANPT has developed a KT task force made up of PTs practicing in different levels of care: experts, early career PTs, supervisors, researchers, patients, and educators. Their role is to support clinicians and organizations in the dissemination and implementation

of CPGs. The primary objective of the core set KT task force is to develop implementation packages that will include KT processes, products, and tools for organizations, clinicians, and educators to use to implement the core set.

The GDG and the Practice Committee of the ANPT jointly developed and disseminated the previous objectives with an invitation to apply for membership on the task force. Interested stakeholders were asked to submit a statement of interest and a curriculum vita. The ANPT Director of Practice,

Practice Committee Chair, and GDG reviewed applications and selected members. Two task force cochairs and 7 members agreed to participate.

The process of collaboration between the task force and the GDG has begun and is anticipated to continue through 2019. As this process evolves, the KT task force, in conjunction with the GDG and the leadership of the ANPT, will finalize plans and develop multiple and diverse implementation recommendations and strategies.

# SUMMARY OF RESEARCH RECOMMENDATIONS

**R. Research Recommendation 1:** Researchers should further examine the BBS, to determine its psychometric properties in neurologic conditions other than stroke, SCI, PD, HD, and MS. Properties such as SEMs, MDCs, and MCID/MICs should be established for individuals with scores throughout the range of the scale in all adult neurologic conditions. Specific information regarding the functional levels of individuals who may benefit from the BBS, and when to start with or transition to another OM, is needed. Determination of optimal administration timing would assist clinicians in administering the BBS within a reasonable time frame when "real change" would be expected. Development and comprehensive testing of a BBS short-form would decrease administration burden.

**R. Research Recommendation 2:** Studies on OMs that provide a comprehensive assessment of sitting balance across acute, chronic progressive, and chronic conditions are needed. These should aim to determine the psychometric properties, including reliability, and to identify information to assist in interpretation, such as MDCs and MIC/MCIDs.

**R. Research Recommendation 3:** Specific information regarding the functional levels of individuals who may benefit from the FGA and when to start with or transition to another OM is needed. Determination of optimal administration timing would assist clinicians in administering the FGA within a reasonable time frame when real change can be expected. Development and psychometric testing of an FGA short-form would decrease administration burden.

**R. Research Recommendation 4:** Studies are needed to examine other OMs, such as the Mini-BESTest and the TUG, in individuals with acute, chronic progressive, and chronic stable neurologic conditions. While the FGA had enough evidence to support its inclusion of the core set, more comprehensive measures of standing and walking balance should be tested to ensure a complete comparison against the FGA. Properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**R. Research Recommendation 5:** Studies are needed to determine the psychometric properties (eg, reliability) of the ABC in acute, chronic progressive, and chronic stable neurologic conditions. Furthermore, information to assist clinicians in interpreting the results of the ABC, such as MDCs and MIC/MCIDs, should be established across neurologic conditions. Specific information regarding the characteristics of individuals who may benefit from the ABC is needed.

**R. Research Recommendation 6:** Studies are needed to examine other OMs, such as the Falls Efficacy Scale International, in individuals with acute, chronic progressive, and chronic stable neurologic conditions. While evidence supports the inclusion of the ABC in the core set, other patient-reported measures of balance should be studied to ensure a comprehensive comparison to the ABC. Properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**R. Research Recommendation 7:** Studies are needed to explore the reliability and clinically important change (eg, MCID) of the 10mWT in individuals with acute neurologic conditions. Clinically important change should also be determined in chronic stable conditions. Studies to determine the presence of floor and ceiling effects should be conducted in persons with chronic progressive and chronic stable conditions.

**R. Research Recommendation 8:** Studies are needed to examine the Walk-12 in individuals with acute, chronic progressive, and chronic stable neurologic conditions. Psychometric properties such as reliability, internal consistency, measurement error, floor and ceiling effects, MDCs, and MIC/MCIDs should be established across neurologic conditions.

**R. Research Recommendation 9:** Studies are needed to determine the intrarater and interrater reliability, and clinically important change (eg, MCID), of the 6MWT in individuals with acute neurologic conditions. Data to assist in measuring change (eg, MDC, SEM, and MCID) are needed in individuals with acute and chronic stable neurologic conditions.

**R. Research Recommendation 10:** Studies are needed that explore the feasibility and psychometric properties of the 5TSTS to objectively describe the transfer abilities of adults with neurologic conditions, especially those other than individuals with PD, across the continuum of care and spectrum of acuity. Further study of the 30SCST is warranted, particularly relative to reliability and data to interpret changes in individuals with neurologic conditions.

**R. Research Recommendation 11:** Studies should explore the feasibility and psychometric properties, including reliability and data to assist in interpreting change (eg, MDC and MCID/MIC) of the GAS and other OMs that capture the individual goals of adults with neurologic conditions across the continuum of care and spectrum of acuity.

**R. Research Recommendation 12:** Studies are needed that explore the impact of using the core set of OMs on rehabilitation outcomes, including factors related to implementation (eg, time and cost). Studies should explore the impact of using the core set of OMs to support clinical decision-making across neurologic conditions and categories. Future measurement studies should be designed to meet the COSMIN requirements for excellent methodology with regard to sample size, design, and rigor of statistical analysis of psychometric properties.[8,9,40,42]

**R. Research Recommendation 13:** The CPG KT Committee is developing standardized administration procedures for all 6 OMs in the core set. Studies are needed to determine the psychometric properties of these protocols across acute, chronic progressive, and chronic conditions in clinical practice.

**R. Research Recommendation 14:** Research is needed on the impact of discussing OM results and shared decision-making with patients receiving neurologic physical therapy, including the development and impact of OM-related information (eg, OM-related decision aids) on the understanding and involvement of a patient in his/her care and on the achievement of patient goals. Furthermore, research should develop and test the use of decision aids that incorporate the core set.

## ACKNOWLEDGMENTS

## REFERENCES

1. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol*. 2006;6:2.
2. Mokkink LB, Terwee CB, Gibbons E, et al. Inter-rater agreement and reliability of the COSMIN (COnsensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol*. 2010;10:82.
3. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol*. 2010;10:22.
4. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.
5. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
6. Institute of Medicine. *Clinical Practice Guidelines We Can Trust*. Washington, DC: The National Academies Press; 2011.
7. Fetters L, Tilson J. *Evidence Based Physical Therapy*. Philadelphia, PA: FA Davis; 2012.
8. Terwee CB. *Protocol for Systematic Reviews of Measurement Properties*. Amsterdam, the Netherlands: Knowledge Center Measurement Instruments; 2011. http://www.cosmin.nl/images/upload/files/Protocol%20klinimetrische%20review%20version%20 nov%202011.pdf
9. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42.
10. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther*. 2008;88(6):733-746.
11. Quinn L, Khalil H, Dawes H, et al. Reliability and minimal detectable change of physical performance measures in individuals with pre-manifest and manifest Huntington disease. *Phys Ther*. 2013;93(7):942-956.
12. American Physical Therapy Association. Clinical Practice Guidelines. http://www.apta.org/EvidenceResearch/EBPTools/CPGs/. Accessed October 10, 2017.
13. American Physical Therapy Association. Guide to Physical Therapist Practice 3.0. http://guidetoptpractice.apta.org/. Accessed June 24, 2017.
14. *International Classification of Functioning, Disability, and health: ICF* [computer program]. Version 1.0. Geneva, Switzerland: World Health Organization; 2001.
15. Potter K, Fulk GD, Salem Y, Sullivan J. Outcome measures in neurological physical therapy practice: part I. Making sound decisions. *J Neurol Phys Ther*. 2011;35(2):57-64.
16. Sullivan JE, Andrews AW, Lanzino D, Perron AE, Potter KA. Outcome measures in neurological physical therapy practice: part II. A patient-centered process. *J Neurol Phys Ther*. 2011;35(2):65-74.
17. Potter K, Cohen ET, Allen DD, et al. Outcome measures for individuals with multiple sclerosis: recommendations from the American Physical Therapy Association Neurology Section task force. *Phys Ther*. 2014;94(5):593-608.
18. Sullivan JE, Crowner BE, Kluding PM, et al. Outcome measures for individuals with stroke: process and recommendations from the American Physical Therapy Association neurology section task force. *Phys Ther*. 2013;93(10):1383-1396.
19. Kahn JH, Tappan R, Newman CP, et al. Outcome measure recommendations from the spinal cord injury EDGE task force. *Phys Ther*. 2016;96(11):1832-1842.
20. McCulloch KL, de Joya AL, Hays K, et al. Outcome measures for persons with moderate to severe traumatic brain injury: recommendations from the American Physical Therapy Association

Academy of Neurologic Physical Therapy TBI EDGE Task Force. *J Neurol Phys Ther*. 2016;40(4):269-280.

21. Academy of Neurologic Physical Therapy. Academy of Neurologic PT Outcome Measures Recommendations. http://www.neuropt. org/professional-resources/neurology-section-outcome-measures-recommendations. Accessed June 25, 2017.

22. Institute of Medicine. *Best Care at Lower Cost: The Path to Continuously Learning Health Care in America*. Washington, DC: Institute of Medicine; 2013.

23. Haigh R, Tennant A, Biering-Sorensen F. The use of outcome measures in physical medicine and rehabilitation in Europe. *Eur J Rehabil Med*. 2001;33:273-278.

24. Jette DU, Halbert J, Iverson C, Miceli E, Shah P. Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Phys Ther*. 2009;89(2):125-135.

25. Andrews AW, Folger SE, Norbet SE, Swift LC. Tests and measures used by specialist physical therapists when examining patients with stroke. *J Neurol Phys Ther*. 2008;32(3):122-128.

26. Thier SO. Forces motivating the use of health status assessment measures in clinical settings and related clinical research. *Med Care*. 1992;30(5 suppl):MS15-MS22.

27. Lansky D, Butler JB, Waller FT. Using health status measures in the hospital setting: from acute care to "outcomes management." *Med Care*. 1992;30(5 suppl):MS57-MS73.

28. Cano SJ, Hobart JC. Watch out, watch out, the FDA are about. *Dev Med Child Neurol*. 2008;50(6):408-409.

29. Fulk G, Field-Fote EC. Measures of evidence in evidence-based practice. *J Neurol Phys Ther*. 2011;35(2):55-56.

30. Horn SD, Gassaway J, Pentz L, James R. Practice-based evidence for clinical practice improvement: an alternative study design for evidence-based medicine. *Stud Health Technol Inform*. 2010;151:446-460.

31. Salter KL, Teasell RW, Foley NC, Jutai JW. Outcome assessment in randomized controlled trials of stroke rehabilitation. *Am J Phys Med Rehabil*. 2007;86(12):1007-1012.

32. Van Peppen RP, Maissan FJ, Van Genderen FR, Van Dolder R, Van Meeteren NL. Outcome measures in physiotherapy management of patients with stroke: a survey into self-reported use, and barriers to and facilitators for use. *Physiother Res Int*. 2008;13(4):255-270.

33. Wedge FM, Braswell-Christy J, Brown CJ, Foley KT, Graham C, Shaw S. Factors influencing the use of outcome measures in physical therapy practice. *Physiother Theory Pract*. 2012;28(2):119-133.

34. Stevens JG, Beurskens AJ. Implementation of measurement instruments in physical therapist practice: development of a tailored strategy. *Phys Ther*. 2010;90(6):953-961.

35. Abrams D, Davidson M, Harrick J, Harcourt P, Zylinski M, Clancy J. Monitoring the change: current trends in outcome measure usage in physiotherapy. *Man Ther*. 2006;11(1):46-53.

36. Beattie P, Maher C. The role of functional status questionnaires for low back pain. *Aust J Physiother*. 1997;43(1):29-38.

37. Centers for Medicare & Medicaid Services. Functional Reporting: PT, OT, and SLP Services Frequently Asked Questions Document Now Available. https://www.cms.gov/Medicare/Billing/Therapy-Services/Functional-Reporting.html. Published 2014. Accessed June 25, 2017.

38. American Physical Therapy Association. Functional Limitation Reporting (FLR) Under Medicare: Tests and Measures for High-Volume Conditions. http://www.ptnow.org/FunctionalLimitation-Reporting/TestsMeasures/. Accessed June 25, 2017.

39. Moore JL, Raad J, Ehrlich-Jones L, Heinemann AW. Development and use of a knowledge translation tool: the rehabilitation measures database. *Arch Phys Med Rehabil*. 2014;95(1):197-202.

40. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res*. 2009;18(8):1115-1123.

41. COnsensus-based Standards for the selection of health Measurement Instruments home page. http://www.cosmin.nl/. Accessed June 25, 2017.

42. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012;21(4):651-657.

43. SurveyMonkey home page. https://www.surveymonkey.com/. Accessed June 25, 2017.

44. Guideline Elements Model. BridgeWiz for APTA 3.0. http://gem. med.yale.edu/BRIDGE-Wiz/. Published 2011. Accessed June 25, 2017.

45. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010;182(18):E839-E842.

46. Shiffman RN, Dixon J, Brandt C, et al. The GuideLine Implementability Appraisal (GLIA): development of an instrument to identify obstacles to guideline implementation. *BMC Med Inform Decis Mak*. 2005;5:23.

47. Gustavsen M, Aamodt G, Mengshoel AM. Measuring balance in sub-acute stroke rehabilitation. *Adv Physiother*. 2006;8(1):15-22.

48. Lemay JF, Nadeau S. Standing balance assessment in ASIA D paraplegic and tetraplegic participants: concurrent validity of the Berg Balance Scale. *Spinal Cord*. 2010;48(3):245-250.

49. Mao HF, Hsueh IP, Tang PF, Sheu CF, Hsieh CL. Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke*. 2002;33(4):1022-1027.

50. Knorr S, Brouwer B, Garland SJ. Validity of the Community Balance and Mobility Scale in community-dwelling persons after stroke. *Arch Phys Med Rehabil*. 2010;91(6):890-896.

51. Tsang CS, Liao LR, Chung RC, Pang MY. Psychometric properties of the Mini-Balance Evaluation Systems Test (Mini-BESTest) in community-dwelling individuals with chronic stroke. *Phys Ther*. 2013;93(8):1102-1115.

52. Schlenstedt C, Brombacher S, Hartwigsen G, Weisser B, Moller B, Deuschl G. Comparing the Fullerton Advanced Balance Scale with the Mini-BESTest and Berg Balance Scale to assess postural control in patients with Parkinson disease. *Arch Phys Med Rehabil*. 2015;96(2):218-225.

53. Donoghue D, Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *J Rehabil Med*. 2009;41(5):343-346.

54. Downs S, Marquez J, Chiarelli P. The Berg Balance Scale has high intra- and inter-rater reliability but absolute reliability varies across the scale: a systematic review. *J Physiother*. 2013;59(2):93-99.

55. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. *Physiother Can*. 1989;41(6):304-311.

56. Blum L, Korner-Bitensky N. Usefulness of the Berg Balance Scale in stroke rehabilitation: a systematic review. *Phys Ther*. 2008;88(5):559-566.

57. Pickenbrock HM, Diel A, Zapf A. A comparison between the Static Balance Test and the Berg Balance Scale: validity, reliability, and comparative resource use. *Clin Rehabil*. 2016;30(3):288-293.

58. Stevenson TJ. Detecting change in patients with stroke using the Berg Balance Scale. *Aust J Physiother*. 2001;47(1):29-38.

59. Salter K, Jutai J, Foley N, Teasell R. Clinical Outcome Variables Scale: a retrospective validation study in patients after stroke. *J Rehabil Med*. 2010;42(7):609-613.

60. Chinsongkram B, Chaikeeree N, Saengsirisuwan V, Viriyatharakij N, Horak FB, Boonsinsukh R. Reliability and validity of the Balance Evaluation Systems Test (BESTest) in people with subacute stroke. *Phys Ther*. 2014;94(11):1632-1643.

61. Delbaere K, Close JC, Mikolaizak AS, Sachdev PS, Brodaty H, Lord SR. The Falls Efficacy Scale International (FES-I). A comprehensive longitudinal validation study. *Age Ageing*. 2010;39(2):210-216.

62. Leddy AL, Crowner BE, Earhart GM. Functional gait assessment and balance evaluation system test: reliability, validity, sensitivity, and specificity for identifying individuals with Parkinson disease who fall. *Phys Ther*. 2011;91(1):102-113.

63. Hiengkaew V, Jitaree K, Chaiyawat P. Minimal detectable changes of the Berg Balance Scale, Fugl-Meyer Assessment Scale, Timed

"Up & Go" Test, gait speeds, and 2-minute walk test in individuals with chronic stroke with different degrees of ankle plantarflexor tone. *Arch Phys Med Rehabil*. 2012;93(7):1201-1208.

64. Liaw LJ, Hsieh CL, Lo SK, Chen HM, Lee S, Lin JH. The relative and absolute reliability of two balance performance measures in chronic stroke patients. *Disabil Rehabil*. 2008;30(9):656-661.

65. Wirz M, Muller R, Bastiaenen C. Falls in persons with spinal cord injury: validity and reliability of the Berg Balance Scale. *Neurorehabil Neural Repair*. 2010;24(1):70-77.

66. Paltamaa J, Sarasoja T, Leskinen E, Wikstrom J, Malkia E. Measuring deterioration in international classification of functioning domains of people with multiple sclerosis who are ambulatory. *Phys Ther*. 2008;88(2):176-190.

67. Verheyden G, Vereeck L, Truijen S, et al. Trunk performance after stroke and the relationship with balance, gait and functional ability. *Clin Rehabil*. 2006;20(5):451-458.

68. Verheyden G, Nuyens G, Nieuwboer A, Van Asch P, Ketelaer P, De Weerdt W. Reliability and validity of trunk assessment for people with multiple sclerosis. *Phys Ther*. 2006;86(1):66-76.

69. Marchetti GF, Lin CC, Alghadir A, Whitney SL. Responsiveness and minimal detectable change of the dynamic gait index and functional gait index in persons with balance and vestibular disorders. *J Neurol Phys Ther*. 2014;38(2):119-124.

70. Wrisley DM, Marchetti GF, Kuharsky DK, Whitney SL. Reliability, internal consistency, and validity of data obtained with the functional gait assessment. *Phys Ther*. 2004;84(10):906-918.

71. Lin JH, Hsu MJ, Hsu HW, Wu HC, Hsieh CL. Psychometric comparisons of 3 functional ambulation measures for patients with stroke. *Stroke*. 2010;41(9):2021-2025.

72. Nilsagård Y, Kollén L, Axelsson H, Bjerlemo B, Forsberg A. Functional Gait Assessment: reliability and validity in people with peripheral vestibular disorders. *Int J Ther Rehabil*. 2014;21(8):367-373.

73. Marchetti GF, Whitney SL. Construction and validation of the 4-item dynamic gait index. *Phys Ther*. 2006;86(12):1651-1660.

74. Wrisley DM, Walker ML, Echternach JL, Strasnick B. Reliability of the dynamic gait index in people with vestibular disorders. *Arch Phys Med Rehabil*. 2003;84(10):1528-1533.

75. Huang SL, Hsieh CL, Wu RM, Tai CH, Lin CH, Lu WS. Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. *Phys Ther*. 2011;91(1):114-121.

76. Matsuda PN, Taylor CS, Shumway-Cook A. Evidence for the validity of the modified dynamic gait index across diagnostic groups. *Phys Ther*. 2014;94(7):996-1004.

77. Leddy AL, Crowner BE, Earhart GM. Utility of the Mini-BESTest, BESTest, and BESTest sections for balance assessments in individuals with Parkinson disease. *J Neurol Phys Ther*. 2011;35(2):90-97.

78. Nilsagård Y, Lundholm C, Gunnarsson LG, Dcnison E. Clinical relevance using timed walk tests and "timed up and go" testing in persons with multiple sclerosis. *Physiother Res Int*. 2007;12(2):105-114.

79. Paul SS, Canning CG, Sherrington C, Fung VS. Reproducibility of measures of leg muscle power, leg muscle strength, postural sway and mobility in people with Parkinson's disease. *Gait Posture*. 2012;36(3):639-642.

80. Stolwijk-Swuste JM, Beelen A, Lankhorst GJ, Nollet F; CARPA Study Group. SF36 physical functioning scale and 2-minute walk test advocated as core qualifiers to evaluate physical functioning in patients with late-onset sequelae of poliomyelitis. *J Rehabil Med*. 2008;40(5):387-394.

81. Ng SS, Hui-Chan CW. The timed up & go test: its reliability and association with lower-limb impairments and locomotor capacities in people with chronic stroke. *Arch Phys Med Rehabil*. 2005;86(8):1641-1647.

82. Powell LE, Myers AM. The Activities-specific Balance Confidence (ABC) Scale. *J Gerontol A Biol Sci Med Sci*. 1995;50A(1):M28-M34.

83. Jonasson SB, Nilsson MH, Lexell J. Psychometric properties of four fear of falling rating scales in people with Parkinson's disease. *BMC Geriatr*. 2014;14:66.

84. Salbach NM, Mayo NE, Hanley JA, Richards CL, Wood-Dauphinee S. Psychometric evaluation of the original and Canadian French version of the activities-specific balance confidence scale among people with stroke. *Arch Phys Med Rehabil*. 2006;87(12):1597-1604.

85. Morgan MT, Friscia LA, Whitney SL, Furman JM, Sparto PJ. Reliability and validity of the Falls Efficacy Scale-International (FES-I) in individuals with dizziness and imbalance. *Otol Neurotol*. 2013;34(6):1104-1108.

86. van Vliet R, Hoang P, Lord S, Gandevia S, Delbaere K. Falls efficacy scale-international: a cross-sectional validation in people with multiple sclerosis. *Arch Phys Med Rehabil*. 2013;94(5):883-889.

87. Scrivener K, Schurr K, Sherrington C. Responsiveness of the ten-metre walk test, Step Test and Motor Assessment Scale in inpatient care after stroke. *BMC Neurol*. 2014;14:129.

88. Perera S, Mody SH, Woodman RC, Studenski SA. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc*. 2006;54(5):743-749.

89. Scivoletto G, Tamburella F, Laurenza L, Foti C, Ditunno JF, Molinari M. Validity and reliability of the 10-m walk test and the 6-min walk test in spinal cord injury patients. *Spinal Cord*. 2011;49(6):736-740.

90. Rossier P, Wade DT. Validity and reliability comparison of 4 mobility measures in patients presenting with neurologic impairment. *Arch Phys Med Rehabil*. 2001;82(1):9-13.

91. Franchignoni F, Tesio L, Benevolo E, Ottonello M. Psychometric properties of the Rivermead Mobility Index in Italian stroke rehabilitation inpatients. *Clin Rehabil*. 2003;17(3):273-282.

92. Hsueh IP, Wang CH, Sheu CF, Hsieh CL. Comparison of psychometric properties of three mobility measures for patients with stroke. *Stroke*. 2003;34(7):1741-1745.

93. Chen HM, Hsieh CL, Sing Kai L, Liaw LJ, Chen SM, Lin JH. The test-retest reliability of 2 mobility performance tests in patients with chronic stroke. *Neurorehabil Neural Repair*. 2007;21(4):347-352.

94. Freeman J, Walters R, Ingram W, Slade A, Hobart J, Zajicek J. Evaluating change in mobility in people with multiple sclerosis: relative responsiveness of four clinical measures. *Mult Scler*. 2013;19(12):1632-1639.

95. Solari A, Radice D, Manneschi L, Motti L, Montanari E. The multiple sclerosis functional composite: different practice effects in the three test components. *J Neurol Sci*. 2005;228(1):71-74.

96. Larson RD, Larson DJ, Baumgartner TB, White LJ. Repeatability of the timed 25-foot walk test for individuals with multiple sclerosis. *Clin Rehabil*. 2013;27(8):719-723.

97. Learmonth YC, Dlugonski DD, Pilutti LA, Sandroff BM, Motl RW. The reliability, precision and clinically meaningful change of walking assessments in multiple sclerosis. *Mult Scler*. 2013;19(13):1784-1791.

98. Motl RW, Learmonth YC, Pilutti LA, Dlugonski D, Klaren R. Validity of minimal clinically important difference values for the multiple sclerosis walking scale-12? *Eur Neurol*. 2014;71(3/4):196-202.

99. Nieuwenhuis MM, Van Tongeren H, Sorensen PS, Ravnborg M. The six spot step test: a new measurement for walking ability in multiple sclerosis. *Mult Scler*. 2006;12(4):495-500.

100. Phan-Ba R, Pace A, Calay P, et al. Comparison of the timed 25-foot and the 100-meter walk as performance measures in multiple sclerosis. *Neurorehabil Neural Repair*. 2011;25(7):672-679.

101. Stellmann JP, Vettorazzi E, Poettgen J, Heesen C. A 3meter Timed Tandem Walk is an early marker of motor and cerebellar impairment in fully ambulatory MS patients. *J Neurol Sci*. 2014;346(1/2):99-106.

102. Baert I, Freeman J, Smedal T, et al. Responsiveness and clinically meaningful improvement, according to disability level,

of five walking measures after rehabilitation in multiple sclerosis: a European multicenter study. *Neurorehabil Neural Repair*. 2014;28(7):621-631.

103. Holland A, O'Connor RJ, Thompson AJ, Playford ED, Hobart JC. Talking the talk on walking the walk: a 12-item generic walking scale suitable for neurological conditions? *J Neurol*. 2006;253(12):1594-1602.

104. Liu J, Drutz C, Kumar R, et al. Use of the six-minute walk test poststroke: is there a practice effect? *Arch Phys Med Rehabil*. 2008;89(9):1686-1692.

105. ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med*. 2002;166(1):111-117.

106. Salbach NM, O'Brien KK, Brooks D, et al. Considerations for the selection of time-limited walk tests poststroke: a systematic review of test protocols and measurement properties. *J Neurol Phys Ther*. 2017;41(1):3-17.

107. Lennon S, Johnson L. The modified Rivermead mobility index: validity and reliability. *Disabil Rehabil*. 2000;22(18):833-839.

108. Walsh JM, Barrett A, Murray D, Ryan J, Moroney J, Shannon M. The Modified Rivermead Mobility Index: reliability and convergent validity in a mixed neurological population. *Disabil Rehabil*. 2010;32(14):1133-1139.

109. Radman L, Forsberg A, Nilsagard Y. Modified Rivermead Mobility Index: a reliable measure in people within 14 days post-stroke. *Physiother Theory Pract*. 2015;31(2):126-129.

110. Duncan PW, Lai SM, Tyler D, Perera S, Reker DM, Studenski S. Evaluation of proxy responses to the Stroke Impact Scale. *Stroke*. 2002;33(11):2593-2599.

111. Hilari K, Owen S, Farrelly SJ. Proxy and self-report agreement on the Stroke and Aphasia Quality of Life Scale-39. *J Neurol Neurosurg Psychiatry*. 2007;78(10):1072-1075.

112. Oczkowski C, O'Donnell M. Reliability of proxy respondents for patients with stroke: a systematic review. *J Stroke Cerebrovasc Dis*. 2010;19(5):410-416.

113. Elder NC, Imhoff R, Chubinski J, et al. Congruence of patient self-rating of health with family physician ratings. *J Am Board Fam Med*. 2017;30(2):196-204.

114. Ruble L, McGrew JH, Toland MD. Goal attainment scaling as an outcome measure in randomized controlled trials of psychosocial interventions in autism. *J Autism Dev Disord*. 2012;42(9):1974-1983.

115. Stolee P, Rockwood K, Fox RA, Streiner DL. The use of goal attainment scaling in a geriatric care setting. *J Am Geriatr Soc*. 1992;40(6):574-578.

116. Turner-Stokes L, Williams H, Johnson J. Goal attainment scaling: does it provide added value as a person-centred measure for evaluation of outcome in neurorehabilitation following acquired brain injury? *J Rehabil Med*. 2009;41(7):528-535.

117. Khan F, Pallant JF, Turner-Stokes L. Use of goal attainment scaling in inpatient rehabilitation for persons with multiple sclerosis. *Arch Phys Med Rehabil*. 2008;89(4):652-659.

118. Cusick A, McIntyre S, Novak I, Lannin N, Lowe K. A comparison of goal attainment scaling and the Canadian Occupational Performance Measure for paediatric rehabilitation research. *Pediatr Rehabil*. 2006;9(2):149-157.

119. Hurn J, Kneebone I, Cropley M. Goal setting as an outcome measure: a systematic review. *Clin Rehabil*. 2006;20(9):756-772.

120. Harlos K, Tetroe J, Graham ID, Bird M, Robinson N. Mining the management literature for insights into implementing evidence-based change in healthcare. *Healthc Policy*. 2012;8(1):33-48.

121. Field-Fote E. Towards Optimal Practice. What can we gain from assessment of patient progress with standardized outcome measures? http://www.ptresearch.org/article/104/resources/researchers/edge-task-force-evaluation-database-to-guide-effectiveness. Accessed July 26, 2017.

122. Paul L, Coote S, Crosbie J, et al. Core outcome measures for exercise studies in people with multiple sclerosis: recommendations from a multidisciplinary consensus meeting. *Mult Scler*. 2014;20(12):1641-1650.

123. Verschuren O, Ketelaar M, Keefer D, et al. Identification of a core set of exercise tests for children and adolescents with cerebral palsy: a Delphi survey of researchers and clinicians. *Dev Med Child Neurol*. 2011;53(5):449-456.

124. Grill E, Bronstein A, Furman J, Zee DS, Muller M. International Classification of Functioning, Disability and Health (ICF) Core Set for patients with vertigo, dizziness and balance disorders. *J Vestib Res*. 2012;22(5/6):261-271.

125. Winser SJ, Smith C, Hale LA, Claydon LS, Whitney SL. Balance outcome measures in cerebellar ataxia: a Delphi survey. *Disabil Rehabil*. 2015;37(2):165-170.

126. Sibley KM, Howe T, Lamb SE, et al. Recommendations for a core outcome set for measuring standing balance in adult populations: a consensus-based approach. *PLoS One*. 2015;10(3):e0120568.

127. Herrmann KH, Kirchberger I, Stucki G, Cieza A. The comprehensive ICF core sets for spinal cord injury from the perspective of physical therapists: a worldwide validation study using the Delphi technique. *Spinal Cord*. 2011;49(4):502-514.

128. Lennon S. Physiotherapy practice in stroke rehabilitation: a survey. *Disabil Rehabil*. 2003;25(9):455-461.

129. Stacey D, Legare F, Lewis K, et al. Decision aids for people facing health treatment or screening decisions. *Cochrane Database Syst Rev*. 2017;4:CD001431.

130. Elwyn G, Frosch D, Thomson R, et al. Shared decision making: a model for clinical practice. *J Gen Intern Med*. 2012;27(10):1361-1367.

131. Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. *BMJ*. 2010;341:c5146.

132. Government of Canada. Canadian Institutes of Health Research Act. http://laws.justice.gc.ca/eng/acts/c-18.1/page-1.html. Accessed July 27, 2017.

133. Nilsen P. Making sense of implementation theories, models and frameworks. *Implement Sci*. 2015;10:53.

134. Hudon A, Gervais MJ, Hunt M. The contribution of conceptual frameworks to knowledge translation interventions in physical therapy. *Phys Ther*. 2015;95(4):630-639.

135. Levac D, Glegg SMN, Camden C, Rivard LM, Missiuna C. Best practice recommendations for the development, implementation, and evaluation of online knowledge translation resources in rehabilitation. *Phys Ther*. 2015;95(4):648-662.

136. Zidarov D, Thomas A, Poissant L. Knowledge translation in physical therapy: from theory to practice. *Disabil Rehabil*. 2013;35(18):1571-1577.

137. Field B, Booth A, Ilott I, Gerrish K. Using the Knowledge to Action Framework in practice: a citation analysis and systematic review. *Implement Sci*. 2014;9:172.

138. Moulding NT, Silagy CA, Weller DP. A framework for effective management of change in clinical practice: dissemination and implementation of clinical practice guidelines. *Qual Health Care*. 1999;8(3):177-183.

139. Straus SE, Tetroe J, Graham ID. *Knowledge Translation in Health Care*. West Sussex, UK: John Wiley & Sons; 2013.

140. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil*. 2000;81(12, suppl 2):S15-S20.

141. Post MW. What to do with "moderate" reliability and validity coefficients? *Arch Phys Med Rehabil*. 2016;97(7):1051-1052.

142. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess*. 1998;2(14):i-iv, 1-74.

143. Colquhoun HL, Lamontagne ME, Duncan EA, Fiander M, Champagne C, Grimshaw JM. A systematic review of interventions to increase the use of standardized outcome measures by rehabilitation professionals. *Clin Rehabil*. 2017;31(3):299-309.

144. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. 2012;(6):CD000259.