

# **Recommendations from the ISPO lower-limb COMPASS: Patient-reported and performance-based outcome measures**

Jade M. Tan<sup>1,2</sup>, Gregory R. J. Halford<sup>2,3</sup>, Martina Lukin and Friedbert Kohler<sup>4,5,6</sup>

## Abstract

**Background:** Outcome measures (patient-reported and performance-based) are used widely but not uniformly within the clinical setting for individuals with lower-limb absence (LLA). The need for more detailed information by funding bodies, service planners and providers, and researchers requires the systematic and routine use of outcome measures. Currently, there is no consensus on which outcome measure(s) should be used for individuals with LLA. The aim of the International Society of Prosthetics and Orthotics (ISPO) lower-limb Consensus Outcome Measures for Prosthetic and Amputation Services (COMPASS) was to produce a recommended list of outcome measures to be actively promoted for routine use within clinical practice before and after an episode of care.

**Methods:** Between May and June 2021, 46 users, clinicians, researchers, managers, and policymakers working in the field of LLA and prosthetic users met virtually. Consensus participants were first asked to complete an online survey with questions based on the results from a systematic review and the outcomes from an expert panel. A modified Delphi technique was used to determine outcome measures for use in routine clinical practice. This paper discusses the ISPO lower-limb COMPASS process from which recommendations were made.

**Results:** The ISPO lower-limb COMPASS resulted in the following 6 recommendations: (1) Amputee Mobility Predictor, Timed Up and Go, Two-Minute Walk Test, Prosthetic Evaluation Questionnaire — Residual Limb Health, Prosthetic Evaluation Questionnaire — Utility, and Trinity Amputation and Prosthesis Experience Scales-Revised, which make up the ISPO lower-limb *COMPASS*; (2) Comprehensive High-Level Activity Mobility Predictor and Six-Minute Walk Test are 2 additional outcome measures recommended for higher-activity-level individuals with LLA, which make up the *COMPASS*+; (3) Patient-Specific Function Scale makes up the *COMPASS Adjunct*; (4) a generic health-related quality of life outcome measure such as the European Quality of Life-5D-5L or Patient-Reported Outcomes Measurement Information System—29 item can be used to supplement the COMPASS; (5) outcome measures suited to low- and middle-income countries need to be developed with a focus on activities such as sitting cross-legged, kneeling, squatting, and other culturally important mobility-related activities; and (6) translation, validation, and open sharing of translated outcome measures included in the COMPASS, COMPASS+, and COMPASS Adjunct occurs.

**Conclusion:** The above recommendations represent the current status of knowledge on outcome measures for LLA based on research and international consensus and hence, will change over time. This work has been developed for clinicians and researchers to improve knowledge on outcome measures to guide clinical decision-making and future research initiatives.

## Keywords

amputation, consensus development conference, patient outcome assessment, patient reported outcome measures

Date received: 5 January 2022; accepted 8 September 2022.

<sup>1</sup>La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia

<sup>2</sup>School of Physiotherapy, Podiatry, Prosthetics and Orthotics, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Australia

<sup>3</sup>International Society of Prosthetics and Orthotics (ISPO), Copenhagen, Denmark

<sup>4</sup>School of Clinical Medicine, Medicine & Health, University of New South Wales, Sydney, Australia

<sup>5</sup>Clinical Directorate Aged Care and Rehabilitation, South Western Sydney Local Health District, Liverpool, Australia

 $^6{\rm Rehabilitation}$  Medicine, Braeside Hospital, Braeside Hospital, HammondCare Health, Sydney, Australia

Corresponding author:

Gregory R. J. Halford. Email: G.Halford@latrobe.edu.au.

## Associate Editor: Reza Safari

Copyright © 2023 The Authors. Published by Wolters Kluwer incorporated on behalf of The International Society for Prosthetics and Orthotics. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal. DOI: 10.1097/PXR.00000000000197

## Introduction

Outcome measures to systematically measure and evaluate change and improvement are essential in contemporary clinical practice.<sup>1</sup> At the simplest level, outcome measures are used to describe and measure the level of function or activity of a patient at a point in time, allow objective comparisons to be made between different patients, or measure the progress of a patient over time. This is fundamental not just to good clinical practice, but also to allow the evaluation of therapeutic interventions and componentry. Routine use of outcome measures is a practical application of an important principle, which Lord Kelvin spoke about nearly 140 years ago: "When you can measure what you are speaking about, and express it in numbers, you know something about it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind; it may be the beginning of knowledge, but you have scarcely in your thoughts advanced to the stage of science."<sup>2</sup>

To date, outcome measures are used widely by clinicians and researchers to assess individuals with lower-limb absence (LLA). They are used to inform decisions regarding prosthetic and therapeutic interventions, and to evaluate intervention outcomes at both an individual<sup>3</sup> and a service<sup>4</sup> level. Broadly, outcome measures can be categorized as patient-reported outcome measures (PROMs) or performance-based outcome measures (PerfOMs). Typically administered as questionnaires, PROMs measure the individual's own perspective of their clinical presentation and treatment, without interpretation and analysis of their response by another individual. This minimizes observer bias and captures aspects that are likely to be important to the individual.<sup>5</sup> Conversely, in PerfOMs, clinicians or researchers collect information (data) in a predefined standardized manner while a patient (i.e., an individual with LLA) undertakes a particular task, for example, timing how long it takes for a patient to walk up and down a hallway.<sup>6</sup>

There are a large number of outcome measures available for use in individuals with LLA, with the choice of outcome measure often based upon the individual's functional status, needs, goals, and expected improvement with clinical care.<sup>7-11</sup> However, there are a number of barriers that often limit the use of appropriate outcome measure(s) in clinical practice and research settings.<sup>4</sup> Some of these barriers include limited confidence with the use of outcome measures; limited understanding of how to use outcome measures; quality of outcome measures; administrative and response burden; an imbalance between clinician and patient goals; concerns regarding the financial benefit to both the clinician and patient; limited knowledge surrounding the statistical results of outcome measures; <sup>4,12</sup>

Lack of routine use of outcome measures in LLA populations has resulted in clinicians not always selecting the most appropriate outcome measures to guide clinical decisions, in addition to limiting quality-improvement initiatives. Lack of clear guidance on a standardized set of outcome measures has resulted in difficulty, or in more extreme cases, an inability to make fully informed policylevel decisions. In high-income countries (HICs), funding bodies increasingly demand objective measurements of a patient's functional status and improvements to be measured using standardized outcome measures,<sup>13-15</sup> to facilitate funding of rehabilitation services such as prosthetic components. The use of outcome measures to help meet the need for an evidence-based business case for investment and recurrent funding in HICs has been articulated<sup>4</sup>; this need extends to low- and middle-income countries (LMICs).

Use of a list of a standard set of outcome measures, administered in a consistent manner before and after an episode of care and used in settings throughout the world, will provide information about effectiveness of treatments and ensure comparability of results. Information gained can guide improved decision-making at the micro (clinician and patient), meso (facilities and services), and macro (health policy) levels.

Currently, there is no gold standard outcome measure and no consensus on which outcome measures should be used for individuals with LLA.<sup>16-18</sup> Promotion of widespread minimal standards of data collection and assessment of outcomes requires a process with broad involvement and agreement on the methodology, the data, and the outcome measurements to be used. Therefore, the aim of this paper is to present the outcomes from an international consensus process, which resulted in the development of the International Society of Prosthetics and Orthotics (ISPO) lower-limb Consensus Outcome Measures for Prosthetic and Amputation Services (COMPASS). This

recommended list of outcome measures will be actively promoted and used routinely within clinical practice for individuals with LLA.

## Background

To reach consensus, a three-step process was implemented: (1) systematic review to identify outcome measures used with individuals with LLA and to identify their psychometric properties<sup>10,11,19</sup>; (2) implementation of an expert panel to evaluate the psychometric properties of the outcome measures and determine whether the psychometric properties are of sufficient standard to allow these outcome measures to be recommended for widespread use<sup>20</sup>; and (3) a modified Delphi and consensus process consisting of a preconsensus survey, followed by a consensus conference. Virtual, rather than inperson, meetings were held because of travel restrictions associated with the COVID-19 pandemic. The virtual meetings (4 × 2.5 hours) were used to discuss the results from the preconsensus survey, and an iterative process was used to reach consensus on decisions regarding the final list of outcome measures and consensus recommendations.

## Step 1: Systematic review of outcome measures for individuals with LLAs

A systematic review of the psychometric properties of outcome measures (PROM and PerfOM) in individuals with LLA was conducted by the project team with input from the ISPO industry advisory group (IAG). After duplicate removals, 7920 articles were screened from titles and abstracts. One hundred eight (108) full texts (spanning 60 individual outcome measures) were selected and included. 19,21-128 A flow diagram of the study selection process according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guideline<sup>129</sup> can be viewed in Figure 1.10,11 The quality of included studies investigating the psychometric properties of outcome measures according to COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) risk of bias checklists ranged from "inadequate" through to "very good,"10,11 with the kappa coefficient for quality assessment ranging from moderate to almost perfect agreement (k = 0.566-0.859).<sup>130</sup> The systematic review including the methodology and results have been described in detail and previously published.<sup>10,11,19</sup>

## Step 2: Formulation of an expert panel

As this consensus process aimed to include a variety of participants ranging from clinicians, health policymakers, health clinic managers, researchers, and prosthetic users (i.e., individuals with LLA who have experience using prosthetic devices and/or services), it was anticipated that many participants would not have the sound knowledge or background understanding required to adequately interpret psychometric properties or the COSMIN results on the 108 articles identified in the systematic review, and incorporate these findings into their consensus votes. Classification of the 60 individual outcome measures (identified in the systematic review) into 3 categories (A: recommended; B: recommended with qualification; and C: unable to recommend) would make it feasible or at least facilitate the incorporation of the psychometric strengths and weaknesses into the consensus participant's votes. Expert panel members were selected based on their expertise in the development and/or validation of outcome measures commonly



Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow chart for the systematic review of outcome measures for individuals with LLA. LLA, lower-limb absence; P&O, prosthetist/orthotist; PerfOM, performance-based outcome measure; PROM, patient-reported outcome measure; ULA, upper-limb absence.

used in LLA populations. Although a mix of sex and geographical location was sought, this was difficult because the experts were all based in HICs. To assist in reducing bias during the classification process, where multiple researchers worked in the same research center or were involved in multiple collaborations together, only one researcher from the research center or collaborative team was invited to partake. The expert panel were able to reduce the 60 individual outcome measures to 23 (12 PROMs and 11 PerfOMs) plus 11 subscales, which contained adequate psychometric properties. The process and outcomes of the expert panel have been described in detail and published previously.<sup>20</sup>

## Methods

#### Step 3: A modified Delphi and consensus process

## Selection of consensus participants

The project team met in December 2020 with input from the ISPO IAG to establish selection criteria for consensus participants, as well

as to identify specific potential participants. It was important to have broad representation in the consensus group so as to increase the diversity of opinions in the process and increase the buy-in from a broad range of stakeholders. Organizations who are significant contributors to prosthetic service provision around the world were invited to nominate consensus conference participants. This included professional organizations (World Federation of Occupation Therapists, World Confederation for Physical Therapy, International Society of Physical and Rehabilitation Medicine, and the International Pedorthic Association); international organizations (World Health Organization, International Committee of the Red Cross, and Humanity and Inclusion), industry (including members of the ISPO IAG), scientific experts (Cochrane Community, members of the expert panel), funders (AtScale and United Stated Agency for International Development), regional organizations (La Fédération Africaine des Techniciens Orthoprothésistes, ISPO Asia, United Arab Society for Prosthetics and Orthotics, ISPO Amesur, Range of Motion Project, and Exceed Worldwide), and consumer representatives (International Confederation of Amputee

Demographics	N = 46							
Sex, n (%)	Female = 20 (43)							
	Male = 26 (57)							
Nationality, <sup>a</sup> n	Afghani = 1 American (United States) = 5; Argentinian = 2; Australian = 1; Austrian = 2; Brazilian = 1;British (United Kingdom) = 8; Cambodian = 1; Columbian = 1; Danish = 1; Dutch (the Netherlands) = 2;El Salvadorans = 1; German = 2; Guatemalan = 1; Indian = 2; Iraqi = 1; Irish = 3; Madagascan = 1;Malaysian = 1; South African = 2; Spanish = 1; Swedish = 1 Swiss = 3; Tanzanian = 1; Thai = 2							
Profession, <sup>b</sup> n	Diabetologist/endocrinologist = 1; end-user = 6; occupational therapist = 2; pedorthotist = 1; physical rehabilitation medicine = 7; physiotherapist = 7; policymaker = 3; prosthetist/orthotist = 16; public health = 1; industry = 2; researcher = 3							

Associations). Participants were also selected based on their clinical experience with individuals with LLA (i.e., prosthetists and orthotists, physiotherapists, occupational therapists, psychologists, and medical practitioners), being a health clinic manager or health policymaker, and active researchers in the fields of LLA outcome measures and psychometrics. Prosthetic users were invited and included in the process to capture their perspectives and ensure their needs were represented. Their viewpoints were deemed essential for this process and thus representation from these prosthetic users was imperative. All expert panel members were also invited to take part in the consensus process. Potential consensus participants were invited via email between December 2020 and April 2021. Participants were required to be available to complete both the preconsensus survey and attend the 4 virtual consensus meetings. Potential consensus participants were also able to nominate other participants for invitation that the project team and IAG may not have identified. Where possible, participants and nominations from underrepresented regions (e.g., South America) were actively sought.

#### Preconsensus survey and voting

Based on the results of the systematic review and outcomes of the expert panel, 23 individual outcome measures (12 PROMs and 11 PerfOMs) and 11 subscales were introduced into the consensus process. During the expert panel process, it was identified that 2 PROMs (Prosthesis Evaluation Questionnaire [PEQ]<sup>62</sup> and Trinity Amputation and Prosthesis Experience Scales-Revised [TAPES-R]<sup>83,86</sup>) should be presented in their entirety as well as by individual subscales, because the subscales were developed and tested to assess different constructs, which results in varying psychometric properties. Therefore, it was deemed more precise to evaluate the subscales individually; hence, both the entire questionnaire and the subscales were recommended to be voted on individually within the consensus process. This resulted in a total of 34 individual questions regarding 23 outcome measures and 11 outcome measure subscales.

The project team then developed the following set of 4 questions to be asked for each individual outcome measure and/or subscale: (1) "Based on your experience and the information presented, is this outcome measure one you would recommend for use in routine *clinical practice*?" with a "yes"/"no" response option; (2) "Based on your experience and the information presented, is this outcome measure one you would recommend for use in *research*?" with a "yes"/"no" response option; (3) "Please rate your experience using this outcome measure" with response options "I have no experience and am not familiar with it," "I have heard of it and/or read about it, but have not used it," "I have some experience using it," or "I consider myself an expert and have used it extensively in clinical practice or research"; and (4) "Do you have any other comments you would like to add about your recommendation or experiences with this outcome measure?" The preconsensus survey was accompanied by a manual, which contained detailed information on each outcome measure (i.e., outcome measure description, instructions, video resources where applicable, chapter-level International Classification of Function [ICF] categorization, and references). For efficiency and focus, the survey was divided into 2—PerfOMs (survey 1) and PROMs (survey 2).

Each consensus participant was sent a link to an anonymized survey (Survey Monkey, San Mateo, CA), at least 1 week before the virtual meeting. Two authors compiled the results of the survey to present to the consensus participants at the start of the first and second virtual consensus meetings, demonstrating which outcome measures had achieved consensus (70% or higher agreed or disagreed<sup>131,132</sup>—Supplemental Digital Content 1, http://links.lww.com/POI/A117 and Supplemental Digital Content 2, http:// links.lww.com/POI/A118).

## Virtual meetings

Four, 2.5-hour virtual meetings were conducted to accommodate the multiple time zones where consensus participants were located. Virtual meetings were conducted via Zoom teleconference software (San Jose, CA) and facilitated by CollaborateUp (an 8(a) SBA-Certified Small Business, Washington, DC) with assistance from 4 authors (G.R.J.H., F.K., M.L., and J.M.T.). Throughout the virtual meetings, discussions on themes which emerged and were considered relevant to the development of a global list of outcome measures were explored and noted for further discussion in meeting 4.

At the first virtual meeting, PerfOM survey results were presented, discussed, and revoted (as "include" or "exclude") in a blinded process using the Zoom poll function, with the results announced after each vote. No option was provided for "Don't know/Not sure" to reach a final resolution on outcome measures. Given the diversity of consensus attendees, the purpose of the discussion and live revote was to allow individuals the chance to share experiences from a variety of contexts, raise questions, or obtain clarification about specific outcome measures. Discussion about the appropriateness of outcome measures and how realistic

	Name of Time outcome (min) measure		) currently	Chapter level ICF categories in order as ranked by consensus participants (left considered most important)													Not in ICF		
		available in (other than English)	Mobility (d4)	Sensory functions and pain (b2)	Self- care (d5)	General tasks and demands (d2)	Functions of the skin and related structures (b8)	Domestic life (d6)	Community, social, and civil life (d9)	Major life areas (d8)	Products and technology (e1)	Mental functions (b1)	Interpersonal interactions and relationships (d7)	Support and relationships (e3)	Socket comfort	Satisfaction with prosthesis	Donning and doffing		
PerfOMs	AMP	10–15	1	×	×							×							
	TUG	1–2	-	×															
	2MWT	<5	-	×															
	Total time	16 min																	
PROMs	PEQ-residual limb health	<5	5			×		×											
	TAPES-R	15	3	×	×		×			х	×	×		×	×	×	х		
	PEQ-utility	<5	5			×						×				×	Х	×	
	Total time	<25 min																	
High- activity PerfOMs	CHAMP	15	-	×															
	6MWT	<10	-	×															
	Total time	25 min																	
Generic	EQ-5D-5L	<5	130	×	×	×			×	х	×		×						
HRQoL	PROMIS-29®	<5	48	×	×		×		×	×			×	×	×				
	Total time	<10 min																	
Generic	PSFS	5-10	5																

17

it would be in various contexts to perform each outcome measure (e.g., in terms of equipment and space) were encouraged from a variety of contexts. Outcome measures achieved consensus when 70% of the attendees agreed to include the measure in the live revote. Those that did not reach consensus ( $\leq 69\%$ ) from the revote but achieved a vote between 31 to 69% were placed into a "yet to reach consensus" list, on the premise that they may contain elements (e.g., ICF categorization) required to achieve a wellrounded list of measures for the ISPO lower-limb COMPASS. Outcome measures that achieved a vote  $\leq 30\%$  were removed entirely.

The second virtual meeting focused on PROMs, using the same process as the PerfOMs. After the second virtual meeting, consensus attendees were asked to (1) recommend outcome measures that had not yet been included in the consensus process and that they felt were important; (2) suggest ideas for consolidating the outcome measures into the ISPO lower-limb COMPASS; and (3) in order of importance, rank the chapter-level ICF categories that the remaining outcome measures covered.

The third virtual meeting was used as an opportunity for open discussion among the consensus attendees regarding the outcome measures that had reached consensus ( $\geq 70\%$ ) and those that achieved a vote of 31%-69%, as per the revotes in the first and second virtual meetings. The outcome measures were presented by type (PROM or PerfOM), chapter-level ICF category, and importance (as voted by consensus attendees after the second virtual meeting). Additional outcome measures that were considered important and nominated by consensus attendees were also presented to the consensus group for further discussion and voting. Attendees were asked to consider the micro (clinician and patient), meso (clinic manager and research), and macro (research and policy) utility of these outcome measures during the open discussion. Discussion was held about amputation and prosthetic-specific constructs, not detailed by the ICF, which outcome measures examine. At the conclusion of the third virtual meeting, attendees were asked to provide a recommended list of outcome measures (based on the remaining lists of "included" and "yet to reach consensus") with a justification for their choices. It was not a requirement in forming individual recommended lists of outcome measures that all mapped ICF categories or additional identified constructs were included, rather it was discussed that some coverage of multiple categories should be considered. The lists were consolidated and the outcome measures that recurred the most were placed into a final ISPO lower-limb COMPASS.

During the fourth and final virtual meeting, the final ISPO lower-limb COMPASS and the associated chapter-level ICF categories were presented to the consensus attendees for a concluding group discussion. Furthermore, relevant themes that had emerged within the previous meetings were discussed and formed into additional recommendations.

## Results

This consensus process (May and June 2021) involved 46 participants (20 female and 26 male; 17 from low- and middle-income countries, and 29 from HICs). A full demographic breakdown can be viewed in Table 1.

After the preconsensus survey and live revoting, 14 outcome measures achieved the predetermined threshold for consensus of  $\geq$ 70% agreement for use in clinical practice and 20 for use in research, with only one outcome measure reaching  $\geq$ 70% agreement to be excluded from use within clinical practice (see Supplemental Digital Content 1, http://links.lww.com/POI/A117 and Supplemental Digital Content 2, http://links.lww.com/POI/A118).

After an in-depth discussion and revoting during the first 2 virtual meetings, 7 outcome measures and 3 subscales achieved consensus to be included, 11 outcome measures and 4 subscales achieved a vote of 31%–69%, and 5 outcome measures and 4 subscales achieved consensus to be excluded (see Supplemental Digital Content 3, http://links.lww.com/POI/A119). Two additional outcome measures that were not included in the initial survey were identified by the consensus participants as being important to include—European Quality of Life (EQ)-5D-5L<sup>133,134</sup> and Patient-Specific Function Scale (PSFS).<sup>135,136</sup> The outcome measures that reoccurred most frequently when individuals compiled their recommended list after the third virtual meeting can be viewed in Supplemental Digital Content 4 (http://links.lww.com/POI/A120).

At the conclusion of the consensus process, the *COMPASS* included 3 PerfOMs (Amputee Mobility Predictor [AMP],<sup>91</sup> Timed Up and Go [TUG],<sup>137</sup> and Two-Minute Walk Test [2MWT]<sup>138</sup>) and 3 PROMs (PEQ—Residual Limb Health,<sup>62</sup> PEQ—Utility,<sup>62</sup> and TAPES-R<sup>83,86</sup>), which collectively span 10 chapter-level ICF categories (Table 2). They are recommended for use routinely for episodes of rehabilitation care and should be administered before and after to measure change in functioning.

In addition, the COMPASS+ included 2 PerfOMs (Comprehensive High-Level Activity Mobility Predictor [CHAMP]<sup>98</sup> and Six-Minute Walk Test [6MWT]<sup>138</sup>). The COMPASS+ for higher-activity-level individuals with LLA was recommended to overcome the ceiling effects of the COMPASS and thus their use is optional. The COMPASS + is intended to be used in addition to the COMPASS for highly active individuals.

Furthermore, the COMPASS Adjunct contains one generic PROM (PSFS<sup>135,136</sup>) chosen because of its high clinical relevance.

It was recommended that a locally relevant health-related quality of life (HRQoL) measure be used such as the EQ-5D- $5L^{133,134,139,140}$  or the Patient-Reported Outcomes Measurement Information System—29 item (PROMIS-29©).<sup>141</sup>

Six key recommendations regarding the use of outcome measures for individuals with LLA arose from the consensus process. These recommendations are as follows:

- 1. Recommendation 1: AMP, TUG, 2MWT, PEQ—Residual Limb Health, PEQ—Utility, and TAPES-R make up the ISPO lower-limb COMPASS;
- Recommendation 2: CHAMP and 6MWT are 2 additional PerfOMs recommended for higher-activity-level individuals with LLA, which make up the COMPASS+;
- Recommendation 3: PSFS makes up the COMPASS Adjunct because of the clinical utility and focus on patients defining their own goals;
- Recommendation 4: A generic HRQoL outcome measure such as the EQ-5D-5L or PROMIS-29<sup>©</sup> can be used to supplement the COMPASS;
- Recommendation 5: Outcome measures suited to LMICs need to be developed with a focus on activities such as sitting cross-legged, kneeling, squatting, and other culturally important mobility-related activities;

6. Recommendation 6: Translation, validation, and open sharing of translated outcome measures included in the COMPASS, COMPASS+, and COMPASS Adjunct occur.

## Discussion

#### **Recommendation 1: COMPASS**

It was recommended that the AMP, TUG, 2MWT, PEQ—Residual Limb Health, PEQ—Utility, and TAPES-R make up the ISPO COMPASS.

Collectively, these 6 outcome measures take approximately 45 minutes to complete-~20 minutes for PerfOMs, which require inperson administration, and ~25 minutes for PROMs, which can be done in a waiting area or online. The consensus process participants agreed these times are feasible within the time that majority of clinicians spend on their first patient-clinician interaction, while noting there might be variability about what is feasible in various contexts. This combination of outcome measures span 10 different chapter-level ICF categories plus important non-ICF categories such as socket comfort, satisfaction with prosthesis, and prosthesis donning. The physical space required for the PerfOMs are a 4.5-m walkway for the AMP, a 3.5-m walkway for the TUG, and an adequate walking track to walk unimpeded for 2 minutes for the 2MWT. These outcome measures will enable detailed assessments and description of functional status at various time points throughout the rehabilitation process, noting that as devices need to be replaced and renewed, there will be multiple episodes of care.

#### **Recommendation 2: COMPASS+**

It was recommended that the CHAMP and 6MWT, 2 additional PerfOMs recommended for higher-activity-level individuals with LLA, make up the COMPASS+.

The CHAMP<sup>98</sup> and 6MWT<sup>138</sup> are 2 PerfOMs that are recommended for higher activity individuals with LLA and help to mitigate the ceiling effect of some of the outcome measures included in the COMPASS. These 2 outcome measures take approximately 25 minutes to complete and cover the mobility (d4) ICF category. Many prosthetic users will not be highly active individuals; so, it is not warranted that the COMPASS+ be routinely administered to all people with LLA. However, mobility is extremely important for highly active individuals with LLA and is linked with prosthetic prescription and modifications<sup>142,143</sup> because of the differing components to those required in an everyday prosthetic device, thus highlighting the importance of accurately capturing this outcome. It is important to note that the COMPASS+ requires significant physical space, with a minimum area of 13 m by 10 m needed for the CHAMP and an adequate walking track to walk unimpeded for 6 minutes for the 6MWT. Although some rehabilitation facilities have the space for these outcome measures, it is acknowledged that many do not, with space being severely limited especially where rehabilitation services are colocated within a hospital.

## **Recommendation 3: COMPASS Adjunct**

It was recommended that the PSFS (a generic PROM) make up the COMPASS Adjunct due to the clinical utility and focus on patients defining their own goals.

Despite the PSFS not being deemed psychometrically adequate by the expert panel process because of poor reliability and measurement error,<sup>18,20</sup> the high clinical utility of this outcome measure resulted in it being reintroduced and commended by the consensus participants because of the focus on users being able to define their own goals. The PSFS takes between 5 and 10 minutes to complete (up to 5 activities can be selected), demonstrates very good internal consistency,<sup>10</sup> but is unable to be categorized using the ICF because of the personalized nature of the activities chosen by the respondent. This outcome measure is especially valuable within a clinical setting because of the ability of individuals to nominate any activities they feel are important and thus it focuses clinical discussions on achievement of individually identified goals that may not be adequately captured using standardized outcome measures. Furthermore, the ability of individuals to change their nominated activities over time accurately reflects individual changes that occur throughout rehabilitation. The highly individualized nature of the PSFS should also enable increased specificity of rehabilitation plans that hopefully translate to further improved outcomes for individuals with LLA.

#### **Recommendation 4: Use of a generic HRQoL instrument**

It was recommended that a generic HRQoL outcome measure such as the EQ-5D-5L or PROMIS-29<sup>©</sup> be used to supplement the COMPASS.

The EQ-5D-5L<sup>139,140</sup> and PROMIS-29©<sup>141</sup> are 2 generic QoL PROMs that take less than 5 minutes each to complete and were deemed important to include as optional additional outcome measures, given their ability to facilitate health economic evaluations. It was recognized during the consensus process that both outcome measures were deemed equally important because of differing government and organizational requirements, which may already have certain health economic outcome measures in place. For example, within the United Kingdom, the EQ-5D has been widely used throughout the National Health Service since 2009<sup>144</sup> compared with the United States, where the PROMIS-29© is often used for health economic evaluation.<sup>145,146</sup> Of note are the potential costs associated with the use of these 2 outcome measures, with the PROMIS-29© being free to use, while the EQ-5D-5L incurs some costs, which may affect their use within LMICs. Furthermore, despite its widespread use, the EQ-5D-5L has not been psychometrically tested within an LLA population. Other HRQoL measures commonly used in a particular jurisdiction should also be considered for ease of comparability.

## Recommendation 5: Outcome measure development for LMICs

It was recommended that outcomes measures suited to LMICs are developed with a focus on activities such as sitting cross-legged, kneeling, squatting, and other culturally important mobilityrelated activities.

Although the outcome measures included within the COMPASS span 10 different chapter-level ICF categories, it was highlighted that no psychometrically sound outcome measure currently exists to accommodate culturally important mobility activities such as sitting cross-legged, kneeling, or squatting. These activities are an important part of everyday living in many LMICs, and if they are unable to be effectively executed, it may prevent an individual from being able to perform basic living functions such as going to the bathroom or sharing a meal with family. Furthermore, it was clear from the systematic reviews<sup>10,11</sup> and emphasized during the consensus meeting discussions that not many outcome measures have been translated into languages commonly used in LMICs. As such, it is recommended that future studies focus on the development of outcome measures that are able to capture a greater diversity of culturally important mobility activities, are psychometrically tested for people with LLA, and are translated into languages appropriate for LMICs.

# Recommendation 6: Translation and open sharing of outcome measures

It was recommended that translation, validation, and open sharing of translated outcome measures included in the COMPASS, COMPASS+, and COMPASS Adjunct occur.

It was apparent during the consensus process that outcome measures are often adapted to suit different languages or settings. However, without undergoing adequate psychometric testing, modified outcome measure data cannot be confidently compared worldwide. Therefore, outcome measures included in the COM-PASS should be translated into multiple languages (particularly those commonly used in LMICs) and their psychometric properties evaluated. This will assist to capture the multiple cultural contexts of people living with LLA and not limit outcome measure use to HICs. These outcome measures should then be made openly available at no charge so they are able to be used routinely. Psychometric testing of translated outcome measures is important to ensure that the validity and reliability of the translated measures is adequate, and thus enable comparability between outcome measures administered in different languages. Outcome measures that have not been psychometrically evaluated in the target language cannot be assumed to have the same validity and reliability as the instrument presented in the original language. Particular attention should be given to PROMs due to the ease of usability worldwide, as PerfOMs use are often limited by lack of experienced clinical staff or access to equipment required to complete the PerfOM.

#### Impact on funding bodies

Individuals and organisations involved in the rehabilitation of people with LLA should consider the COMPASS to improve routine data collection and reporting, to assist in maximising the potential of investments.

#### Strengths

A key strength of this consensus process was the diversity of individuals involved. Consensus attendees included various experts and professionals who had extensive experience in the field of LLA, as well as people with lived experience of LLA. Further, there was representation from multiple high-, middle-, and low-income countries, as well a diversity in the gender of those present at the consensus meetings (Table 1). The ability to conduct this consensus meeting virtually enabled a larger number and greater diversity of individuals to be involved, thus increasing the worldwide representation and rigor of this process. Consideration of outcome measures appropriateness in a variety of settings and cultures has produced a standardized list of outcome measures that can be implemented in a diverse range of settings.

Moreover, as this was an iterative consensus process, participants had the ability to recommend outcome measures based on their personal experience. This helped to strengthen the worldwide applicability of the final ISPO lower-limb COMPASS. Three outcome measures (Basic Amputee Mobility Score,<sup>147</sup> Reintegration to Normal Living Index,<sup>148</sup> and Handling of the Prosthesis) identified by the consensus participants were not included in the final list due to not having undergone psychometric testing within an LLA population, containing domains already covered in other psychometrically sound outcome measures, not currently available in the English language, not yet published within a peer-reviewed journal, and only capturing the immediate postsurgical stage of rehabilitation, thus limiting the patient scope (e.g., those born with congenital limb absence) and ability to track the entire rehabilitation journey.

Mapping of outcome measures to chapter level ICF categories and other identified constructs (socket comfort, satisfaction with prosthesis and prosthesis donning) allowed consensus participants to consider diverse outcomes and ensure that a single construct was not measured multiple times while unconsciously excluding constructs that are meaningful to users.

Lastly, benefits of a list of agreed-upon list of outcome measures not only include improved comparability and ability to aggregate outcome measure data internationally, but also the potential to inform clinical care at the patient level (microlevel impact), decision-making at a centre level (mesolevel impact), and policymaking at the national level (macrolevel impact).

## Limitations

The COSMIN risk of bias tool, used within the preparatory phases of this consensus process, was deemed the most appropriate tool to use for this process, however the interpretation of the criteria and hence the scoring are open to individual interpretation. The COSMIN tool has very rigorous standards with many outcome measures often not meeting these high standards in an LLA population.

Mapping outcome measures to the chapter level ICF categories and identification of additional constructs was done to highlight to participants when forming their recommended list if there was repeated measurement, with multiple outcome measures, of the same construct within the COMPASS. Whilst consensus participants were not required to nominate an individual list of outcome measures that spanned all identified ICF categories (or other non ICF identified constructs) many did so. This method may have led to the exclusion of high-quality outcome measures which span fewer ICF categories.

Despite significant research and development of mobility-specific outcome measures in recent years, no PROM exclusively measuring mobility was included in the COMPASS. This potentially reflects the lack of global uptake of PROMs compared with PerfOMs, with some likely finding it a challenge to contextualise an individual's mobility via the use of a PROM. Any update of the COMPASS should carefully examine the inclusion of a mobility-specific PROM.

Psychometric properties are critical to ensure that outcome measures accurately measure the construct within the population of interest. However, they may be difficult for some stakeholders, particularly those without scientific or clinical training, to understand. Thus, an expert panel before this consensus process determined which outcome measures would be recommended and hence considered by consensus participants. The expert panel did not rank or make further differential recommendations about individual outcome measures. Hence, consideration of psychometric properties of outcome measures was likely not considered by most consensus participants.

The systematic review that helped to inform the consensus process was restricted to the English language. This was a limitation from the historical as well as the implementation perspectives. Non-English outcome measures would have been missed, and adoption of these outcome measures in routine practice is restricted to the English language, with limited validated translated outcome measures currently available. This may impede the implementation of outcome measures in LMICs.

There is a vast array of outcome measures that have been developed and tested in other clinical populations but may also be appropriate for use in LLA. However, further psychometric testing would need to be undertaken to ensure these outcome measures function appropriately in the LLA population.

Although it is unlikely that significant patient harm can come from the implementation of these recommendations, in areas where there is significant unmet need and clinical time is the major constraint to meeting that need, time spent on outcome measure data collection may result in further limiting clinical service provision.

#### **Future directions**

Regular review of the outcome measures available and the associated literature to support their use will be needed as new research evidence emerges. The ISPO lower-limb COMPASS should be reviewed and revised as new measures and evidence become available.

Normative outcome measure data and responsiveness statistics (e.g., minimal clinical important difference or minimal detectable change) for LLA populations were not included in this consensus process. Known normative data (although somewhat limited in this population) and responsiveness statistics would be highly valuable for clinicians and researchers with regards to benchmarking patients, formulating rehabilitation and management plans, and tracking improvements over time.

The use of the ISPO lower-limb COMPASS will facilitate future comparisons and pooling of data, which is highly useful with regards to identifying areas for efficient allocation of funding and establishing real-world effectiveness of interventions based on outcomes. Organizational (e.g., government, regulatory bodies, hospitals, etc.) requirements for routine automated collection of outcome measures should base their data requirements on the ISPO lower-limb COMPASS to ensure comparability.

Further linkage of the included outcome measures to cover more ICF categories at the second level and the third level is required. Additional development of outcome measures with diverse consumer involvement to enable broader coverage of the ICF categories relevant to people with LLA is needed.

Refinement of outcome measures to allow better evaluation of technical and/or environmental components that contribute to patient outcomes would allow better evaluation of the benefits of environmental/technological contributions to the functioning of the individual with LLA.

Infographics, instructional videos, outcome measure instructions, and access to data collection sheets are available within the COMPASS User Guide<sup>149</sup> available through the ISPO website to assist with worldwide dissemination and implementation.

## Conclusion

The consensus process successfully developed a short list of outcome measures recommended for routine clinical use worldwide, which has been formulated into the COMPASS, the COMPASS+, and the COMPASS Adjunct. Participants of the ISPO lower-limb COMPASS process recommend the use of 6 core outcome measures in individuals with LLA (AMP, TUG, 2MWT, PEQ-Residual Limb Health, PEQ-Utility, and TAPES-R) as part of the COMPASS. Collectively, these outcome measures span 10 chapter-level ICF categories. Additional outcome measures have been recommended to supplement the core ISPO lower-limb COMPASS and include the CHAMP and 6MWT, which make up the COMPASS+, the PSFS included in the COMPASS Adjunct, and the use of a generic HRQoL outcome measure (e.g., EQ-5D-5L and PROMIS-29<sup>©</sup>). Outcome measures suited to LMICs need to be developed with a focus on culturally important mobility activities such as sitting cross-legged, kneeling, and squatting. To ensure global uptake, outcome measures included in the ISPO COMPASS should be translated into multiple languages, tested for psychometric performance in the target language, and shared free of charge openly to facilitate widespread use.

## Author contributions

JMT, GRJH, ML and FK designed the consensus process. JMT drafted the manuscript with input from GRJH, ML and FK. All authors have read and approved the final manuscript.

#### Funding

International Society of Prosthetics and Orthotics (ISPO) (Copenhagen, Denmark) coordinated and supervised this work, which was funded by United States Agency for International Development (USAID; Washington, DC), in support of AtScale (Geneva, Switzerland), and managed through the United Nations Office for Project Services (UNOPS; Copenhagen, Denmark). Grant ID: UNOPS/CFP-2020/006/ATSCALE.

## **Declaration of conflicting interest**

F.K. on behalf of International Society of Prosthetics and Orthotics (ISPO) obtained the funding but was not funded for his contribution to the project. F.K. helped to run the consensus process.

#### Acknowledgments

The authors would like to acknowledge the following consensus participants: Abir Mitra, Andreas Hahn, Barbara Goedde, Bernd Schwien, Breda Clancy, Brendan McCormack, Carolina Schiappacasse, Carson Harte, Chantel Ostler, Christopher Hutchinson, Christian Rafamatanantosa, Christina Garces, David Boone, David Henson, Gavin Jenkins, Gerfried Peternell, Ghassan Ali, Grazielle Oliveira Da silva, Gulapar Srisawasdi, Julia Patrick Engkasan, Kajsa Lindberg, Katherine Cook, Katie O'Halloran, Lisa Gramnäs, Luis Aragón, Luphiwo Mduzaa, Mahpekai Sediqi, Maria Paleo, Matthew Major, Michael McGrath, Nicolás Muñoz Buderus, Niles Jonkergouw, Oinam Sanjoy Singh, Oscar Bermudez, Phillip Stevens, Rob Verwaard, Robert Gailey, Roxanne Maritz, Sarah Hughes, Sirirat Seng-iad, Sisary Kheng, Sourona Visagie, Yudith Chivatá, and Zulfiqarali Abbas. ISPO and the project team would like to thank all consensus attendees for their contribution to the ISPO lower-limb COMPASS.

#### **ORCID** iDs

J.M. Tan: D https://orcid.org/0000-0003-3677-7666 F. Kohler: D https://orcid.org/0000-0001-6226-3314

#### Supplemental material

Supplemental material for this article is available in this article. Direct URL citation appears in the text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.POIjournal.org).

#### References

- Sackett DL, Rosenberg WM, Gray JM, et al. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312: 71–72.
- 2. Thomson W. Electrical units of measurement. *Popular Lect Address* 1883; 1: 73.
- 3. Agrawal V. Clinical outcome measures for rehabilitation of amputees: a review. *Phys Med Rehabil Int* 2016; 3: 1080–1084.
- 4. Gaunaurd I, Spaulding SE, Amtmann D, et al. Use of and confidence in administering outcome measures among clinical prosthetists: results from a national survey and mixed-methods training program. *Prosthet Orthot Int* 2015; 39: 314–321.
- Kingsley C and Patel S. Patient-reported outcome measures and patientreported experience measures. BJA Educ 2017; 17: 137–144.
- Hawkins EJ and Riddick W. Reliability, validity, and responsiveness of clinical performance–based outcome measures of walking for individuals with lower limb amputations: a systematic review. *Phys Ther* 2018; 98: 1037–1045.
- Condie E, Scott H and Treweek S. Lower limb prosthetic outcome measures: a review of the literature 1995 to 2005. *J Prosthet Orthot* 2006; 18: P13–P45.
- Heinemann AW, Connelly L, Ehrlich-Jones L, et al. Outcome instruments for prosthetics: clinical applications. *Phys Med Rehabil Clin* 2014; 25: 179–198.
- Hawkins AT, Henry AJ, Crandell DM, et al. A systematic review of functional and quality of life assessment after major lower extremity amputation. *Ann Vasc Surg* 2014; 28: 763–780.
- Tan JM, Halford GRJ, McBain BK, et al. A systematic review of psychometric properties of outcome measures for lower-limb absence: part I—patient-reported. *Prosthet Orthot Int* 2021 [epub ahead of print].
- Tan JM, Halford GRJ, McBain BK, et al. A systematic review of psychometric properties of outcome measures for lower-limb absence: part II—performance-based. *Prosthet Orthot Int* 2021 [epub ahead of print].
- Miller LA and McCay JA. Summary and conclusions from the Academy's Sixth State-of-the-Science Conference on lower limb prosthetic outcome measures. J Prosthet Orthot 2006; 18: 2–7.
- Hough A. Regulating disability services: the case of Australia's National Disability Insurance Scheme quality and safeguarding system. *Res Pract Intellect Develop Disabil* 2021; 8: 99–110.
- Transport Accident Commission (TAC). Clinical Framework for the Delivery of Health Services. Principle 1: Measure and Demonstrate the Effectivenss of Treatment. 2012. Available at: https://www.tac.vic.gov. au/providers/working-with-the-tac/clinical-framework/principle-1. Accessed October 26, 2021.
- 15. National Disability Independent Scheme (NDIS). Assessment Tools to Improve Consistency of NDIS Access and Plans. Media Release. 2020.

Available at: https://www.ndis.gov.au/news/5355-assessment-tools-improve-consistency-ndis-access-and-plans. Accessed 26 October, 2021.

- Deathe AB, Wolfe DL, Devlin M, et al. Selection of outcome measures in lower extremity amputation rehabilitation: ICF activities. *Disabil Rehabil* 2009; 31: 1455–1473.
- Kohler F, Cieza A, Stucki G, et al. Developing core sets for persons following amputation based on the International Classification of Functioning, Disability and Health as a way to specify functioning. *Prosthet Orthot Int* 2009; 33: 117–129.
- Resnik L and Borgia M. Reliability of outcome measures for people with lower-limb amputations: distinguishing true change from statistical error. *Phys Ther* 2011; 91: 555–565.
- Tan JM, Halford GRJ, McBain BK, et al. Psychometric properties of outcome measures for lower-limb absence: a protocol for a systematic review. *Prosthet Orthot Int* 2021 [epub ahead of print].
- Kohler F, Halford GRJ, Lukin M, et al. Expert panel categorisation and recommendations for outcome measures for lower limb absence. *Prosthet Orthot Int* 2021 [epub ahead of print].
- Panesar BS, Morrison P and Hunter J. A comparison of three measures of progress in early lower limb amputee rehabilitation. *Clin Rehabil* 2001; 15: 157–171.
- 22. Day HJ. The assessment and description of amputee activity. *Prosthet Orthot Int* 1981; 5: 23–28.
- Miller WC, Deathe AB and Speechley M. Psychometric properties of the Activities-specific Balance Confidence Scale among individuals with a lower-limb amputation. Arch Phys Med Rehabil 2003; 84: 656–661.
- Sakakibara BM, Miller WC and Backman CL. Rasch analyses of the Activities-specific Balance Confidence Scale with individuals 50 years and older with lower-limb amputations. *Arch Phys Med Rehabil* 2011; 92: 1257–1263.
- Hafner BJ, Morgan SJ, Askew RL, et al. Psychometric evaluation of selfreport outcome measures for prosthetic applications. J Rehabil Res Dev 2016; 53: 797–812.
- Fuller K, Omaña Moreno HA, et al. Reliability, validity, and agreement of the short-form Activities-specific Balance Confidence Scale in people with lower extremity amputations. *Prosthet Orthot Int* 2019; 43: 609–617.
- 27. Breakey JW. Body image: the lower-limb amputee. J Prosthet Orthot 1997; 9: 58–66.
- Lai FH-Y, Wong E, Wong SK-M, et al. Development and validation of a Body Image Assessment for patient after lower limb amputation—the Chinese Amputee Body Image Scale-CABIS. *Asian J Occup Ther* 2005; 4: 1–11.
- Gallagher P, Horgan O, Franchignoni F, et al. Body image in people with lower-limb amputation: a Rasch analysis of the Amputee Body Image Scale. Am J Phys Med Rehabil 2007; 86: 205–215.
- Bumin G, Bayramlar K, Yakut Y, et al. Cross cultural adaptation and reliability of the Turkish version of Amputee Body Image Scale (ABIS). *J Back Musculoskelet Rehabil* 2009; 22: 11–16.
- Safaz I, Yilmaz B, Goktepe AS, et al. Turkish version of the Amputee Body Image Scale and relationship with quality of life. *Bull Clin Psychopharmacol* 2010; 20: 79–83.
- 32. Norvell DC, Williams RM, Turner AP, et al. The development and validation of a novel outcome measure to quantify mobility in the dysvascular lower extremity amputee: The amputee single item mobility measure. *Clin Rehabil* 2016; 30: 878–889.
- Rybarczyk BD, Nyenhuis DL, Nicholas JJ, et al. Social discomfort and depression in a sample of adults with leg amputations. *Arch Phys Med Rehabil* 1992; 73: 1169–1173.
- 34. Gailey R, Kristal A, Lucarevic J, et al. The development and internal consistency of the comprehensive lower limb amputee socket survey in active lower limb amputees. *Prosthet Orthot Int* 2019; 43: 80–87.
- de Laat FA, Rommers GM, Geertzen JH, et al. Construct validity and testretest reliability of the climbing stairs questionnaire in lower-limb amputees. Arch Phys Med Rehabil 2010; 91: 1396–1401.
- Monteiro RP, Pfeifer LI, Soares I, et al. Validation of the functional and social performance—DSF-84 checklist: preliminary study. *Disabil Rehabil* 2013; 35: 1527–1533.
- Donovan-Hall MK, Yardley L and Watts RJ. Engagement in activities revealing the body and psychosocial adjustment in adults with a transtibial prosthesis. *Prosthet Orthot Int* 2002; 26: 15–22.
- Leung EC, Rush PJ and Devlin M. Predicting prosthetic rehabilitation outcome in lower limb amputee patients with the functional independence measure. *Arch Phys Med Rehabil* 1996; 77: 605–608.
- Stineman MG, Shea JA, Jette A, et al. The Functional Independence Measure: tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories. *Arch Phys Med Rehabil* 1996; 77: 1101–1108.

- Callaghan BG and Condie ME. A post-discharge quality of life outcome measure for lower limb amputees: test-retest reliability and construct validity. *Clin Rehabil* 2003; 17: 858–864.
- Miller WC, Deathe AB and Harris J. Measurement properties of the Frenchay Activities Index among individuals with a lower limb amputation. *Clin Rehabil* 2004; 18: 414–422.
- Gardiner MD, Faux S and Jones LE. Inter-observer reliability of clinical outcome measures in a lower limb amputee population. *Disabil Rehabil* 2002; 24: 219–225.
- Devlin M, Pauley T, Head K, et al. Houghton Scale of prosthetic use in people with lower-extremity amputations: reliability, validity, and responsiveness to change. Arch Phys Med Rehabil 2004; 85: 1339–1344.
- 44. Wong CK, Gibbs W and Chen ES. Use of the Houghton Scale to classify community and household walking ability in people with lower-limb amputation: criterion-related validity. *Arch Phys Med Rehabil* 2016; 97: 1130–1136.
- Gauthier-Gagnon C and Grise MC. Prosthetic profile of the amputee questionnaire: validity and reliability. Arch Phys Med Rehabil 1994; 75: 1309–1314.
- Miller WC, Deathe AB and Speechley M. Lower extremity prosthetic mobility: a comparison of 3 self-report scales. *Arch Phys Med Rehabil* 2001; 82: 1432–1440.
- Franchignoni F, Orlandini D, Ferriero G, et al. Reliability, validity, and responsiveness of the locomotor capabilities index in adults with lowerlimb amputation undergoing prosthetic training. *Arch Phys Med Rehabil* 2004; 85: 743–748.
- Ferriero G, Dughi D, Orlandini D, et al. Measuring long-term outcome in people with lower limb amputation: cross-validation of the Italian versions of the Prosthetic Profile of the Amputee and Prosthesis Evaluation Questionnaire. *Eur Medicophys* 2005; 41: 1–6.
- Franchignoni F, Giordano A, Ferriero G, et al. Rasch analysis of the Locomotor Capabilities Index-5 in people with lower limb amputation. *Prosthet Orthot Int* 2007; 31: 394–404.
- Dite W, Connor HJ and Curtis HC. Clinical identification of multiple fall risk early after unilateral transtibial amputation. *Arch Phys Med Rehabil* 2007; 88: 109–114.
- Larsson B, Johannesson A, Andersson IH, et al. The Locomotor Capabilities Index; validity and reliability of the Swedish version in adults with lower limb amputation. *Health Qual Life Outcomes* 2009; 7: 1–9.
- Salavati M, Mazaheri M, Khosrozadeh F, et al. The Persian version of Locomotor Capabilities Index: translation, reliability and validity in individuals with lower limb amputation. *Qual Life Res* 2011; 20: 1–7.
- 53. Becker P, Repo JP, Piitulainen K, et al. Validity and reliability of the Finnish version of the Locomotor Capabilities Index-5 in patients fitted with a prosthesis after major lower extremity amputation. J Prosthet Orthot 2019; 31: 43–50.
- 54. Franchignoni F, Traballesi M, Monticone M, et al. Sensitivity to change and minimal clinically important difference of the Locomotor Capabilities Index-5 in people with lower limb amputation undergoing prosthetic training. Ann Phys Rehabil Med 2019; 62: 137–141.
- Wong CK. Limb laterality recognition score: a reliable clinical measure related to phantom limb pain. *Pain Med* 2018; 19: 753–756.
- Wetterhahn KA, Hanson C and Levy CE. Effect of participation in physical activity on body image of amputees. *Am J Phys Med Rehabil* 2002; 81: 194–201.
- Hart DL. Orthotics and Prosthetics National Outcomes Tool (OPOT): initial reliability and validity assessment for lower extremity prosthetics. *J Prosthet Orthot* 1999; 11: 101–111.
- Jarl G, Heinemann AW, Lindner HY, et al. Cross-cultural validity and differential item functioning of the orthotics and prosthetics users' survey with Swedish and United States users of lower-limb prosthesis. *Arch Phys Med Rehabil* 2015; 96: 1615–1626.
- Abu Osman NA, Eshraghi A, Gholizadeh H, et al. Prosthesis donning and doffing questionnaire: development and validation. *Prosthet Orthot Int* 2017; 41: 571–578.
- Giray E, Eshraghi A and Gündüz OH. Validity and reliability of the Turkish version of the Prosthesis Donning and Doffing Questionnaire for persons with transtibial amputations. J Prosthet Orthot 2022; 34: 122–130.
- Bekrater-Bodmann R. Perceptual correlates of successful body-prosthesis interaction in lower limb amputees: psychometric characterisation and development of the Prosthesis Embodiment Scale. Sci Rep 2020; 10: 1–13.
- Legro MW, Reiber GD, Smith DG, et al. Prosthesis evaluation questionnaire for persons with lower limb amputations: assessing prosthesisrelated quality of life. Arch Phys Med Rehabil 1998; 79: 931–938.
- 63. Franchignoni F, Giordano A, Ferriero G, et al. Measuring mobility in people with lower limb amputation: Rasch analysis of the mobility section

of the prosthesis evaluation questionnaire. J Rehabil Med 2007; 39: 138-144.

- Christensen J, Doherty P, Bjorner JB, et al. Reliability and construct validity of a new Danish translation of the Prosthesis Evaluation Questionnaire in a population of Danish amputees. *Prosthet Orthot Int* 2017; 41: 469–475.
- Repo JP, Piitulainen K, Hakkinen A, et al. Reliability and validity of the Finnish version of the prosthesis evaluation questionnaire. *Disabil Rehabil* 2018; 40: 2081–2087.
- 66. Franchignoni F, Ferriero G, Giordano A, et al. The minimal clinicallyimportant difference of the Prosthesis Evaluation Questionnaire— Mobility Scale in subjects undergoing lower limb prosthetic rehabilitation training. *Eur J Phys Rehabil Med* 2020; 56: 82–87.
- 67. Repo JP, Häkkinen AH, Barner-Rasmussen I, et al. Comparison of measurement properties of the mobility section of the Prosthesis Evaluation Questionnaire on a visual analogue scale and its modified version with five response categories. J Prosthet Orthot 2020; 32: 188–193.
- Morgan SJ, Amtmann D, Abrahamson DC, et al. Use of cognitive interviews in the development of the PLUS-M item bank. *Qual Life Res* 2014; 23: 1767–1775.
- Hafner BJ, Morgan SJ, Abrahamson DC, et al. Characterizing mobility from the prosthetic limb user's perspective: use of focus groups to guide development of the Prosthetic Limb Users Survey of Mobility. *Prosthet Orthot Int* 2016; 40: 582–590.
- Karatzios C, Loiret I, Luthi F, et al. Transcultural adaptation and validation of a French version of the Prosthetic Limb Users Survey of Mobility 12-item Short-Form (PLUS-M/FC-12) in active amputees. *Ann Phys Rehabil Med* 2019; 62: 142–148.
- Franchignoni F, Monticone M, Giordano A, et al. Rasch validation of the Prosthetic Mobility Questionnaire: a new outcome measure for assessing mobility in people with lower limb amputation. *J Rehabil Med* 2015; 47: 460–465.
- Burger H, Giordano A, Bavec A, et al. The Prosthetic Mobility Questionnaire, a tool for assessing mobility in people with lower-limb amputation: validation of PMQ 2.0 in Slovenia. *Int J Rehabil Res* 2019; 42: 263–269.
- Grise MC, Gauthier-Gagnon C and Martineau GG. Prosthetic profile of people with lower extremity amputation: conception and design of a follow-up questionnaire. *Arch Phys Med Rehabil* 1993; 74: 862–870.
- 74. Amtmann D, Morgan SJ, Kim J, et al. Health-related profiles of people with lower limb loss. *Arch Phys Med Rehabil* 2015; 96: 1474–1483.
- Hagberg K, Branemark R and Hagg O. Questionnaire for Persons with a Transfemoral Amputation (Q-TFA): initial validity and reliability of a new outcome measure. J Rehabil Res Dev 2004; 41: 695–706.
- de Laat FA, Rommers GM, Geertzen JH, et al. Construct validity and testretest reliability of the questionnaire rising and sitting down in lower-limb amputees. Arch Phys Med Rehabil 2011; 92: 1305–1310.
- Herbst J, Polanski-Schrader LMT, Vogt J, et al. Validation and revision of the questionnaire to explore human factors and their technical potential for lower limb prosthetics. *Prosthet Orthot Int* 2021; 45: 6–11.
- Hanspal RS, Fisher K and Nieveen R. Prosthetic socket fit comfort score. Disabil Rehabil 2003; 25: 1278–1280.
- Ryall N, Eyres S, Neumann V, et al. The SIGAM mobility grades: a new population-specific measure for lower limb amputees. *Disabil Rehabil* 2003; 25: 833–844.
- Joussain C, Laroche D, Casillas JM, et al. Transcultural validation of the SIGAM mobility grades in French: the SIGAM-Fr. *Ann Phys Rehabil Med* 2015; 58: 161–166.
- Yilmaz H, Gafuroglu U, Ryall N, et al. Establishing the Turkish version of the SIGAM mobility scale, and determining its validity and reliability in lower extremity amputees. *Disabil Rehabil* 2018; 40: 346–352.
- de Laat FA, Roorda LD, Geertzen JH, et al. Test-retest reliability of the Special Interest Group on Amputation Medicine/Dutch Working Group on Amputations and Prosthetics Mobility Scale, in persons wearing a prosthesis after a lower-limb amputation. *Disabil Rehabil* 2020; 42: 1762–1766.
- Gallagher P and MacLachlan M. Development and psychometric evaluation of the trinity amputation and prosthesis experience scales (TAPES). *Rehabil Psychol* 2000; 45: 130–154.
- Gallagher P and Maclachlan M. The Trinity Amputation and Prosthesis Experience Scales and quality of life in people with lower-limb amputation. Arch Phys Med Rehabil 2004; 85: 730–736.
- Mazaheri M, Fardipour S, Salavati M, et al. The Persian version of Trinity Amputation and Prosthetics Experience Scale: translation, factor structure, reliability and validity. *Disabil Rehabil* 2011; 33: 1737–1745.
- Gallagher P, Franchignoni F, Giordano A, et al. Trinity amputation and prosthesis experience scales: a psychometric assessment using classical test theory and Rasch analysis. *Am J Phys Med Rehabil* 2010; 89: 487–496.

- Massarweh R and Sobuh MMD. The Arabic Version of Trinity Amputation and Prosthetic Experience Scale—Revised (TAPES-R) for lower limb amputees: reliability and validity. *Disabil CBR Inclusive Develop* 2019; 30: 44–56.
- Luthi F, Praz C, Leger B, et al. Cross-cultural adaptation and measurement properties of the French version of the Trinity Amputation and Prosthesis Experience Scales-Revised (TAPES-R). *PLoS One* 2020; 15: e0229084.
- de Laat FA, Rommers GM, Geertzen JH, et al. Construct validity and testretest reliability of the walking questionnaire in people with a lower limb amputation. Arch Phys Med Rehabil 2012; 93: 983–989.
- Theeven P, Hemmen B, Stevens C, et al. Feasibility of a new concept for measuring actual functional performance in daily life of transfermoral amputees. J Rehabil Med 2010; 42: 744–751.
- Gailey RS, Roach KE, Applegate EB, et al. The amputee mobility predictor: an instrument to assess determinants of the lower-limb amputee's ability to ambulate. *Arch Phys Med Rehabil* 2002; 83: 613–627.
- Côté-Martin M-É, Tremblay A, Couture M, et al. Translation, reliability, and validity of the French version of the amputee mobility predictor. *J Prosthet Orthot* 2020; 32: 101–106.
- Wong CK, Chen CC and Welsh J. Preliminary assessment of balance with the Berg Balance Scale in adults who have a leg amputation and dwell in the community: Rasch rating scale analysis. *Phys Ther* 2013; 93: 1520–1529.
- Major MJ, Fatone S and Roth EJ. Validity and reliability of the Berg Balance Scale for community-dwelling persons with lower-limb amputation. Arch Phys Med Rehabil 2013; 94: 2194–2202.
- Wong CK. Interrater reliability of the Berg Balance Scale when used by clinicians of various experience levels to assess people with lower limb amputations. *Phys Ther* 2014; 94: 371–378.
- Azuma Y, Chin T and Miura Y. The relationship between balance ability and walking ability using the Berg Balance Scale in people with transfemoral amputation. *Prosthet Orthot Int* 2019; 43: 396–401.
- Gailey RS. Construct validity of Comprehensive High-Level Activity Mobility Predictor (CHAMP) for male servicemembers with traumatic lower-limb loss. J Rehabil Res Dev 2013; 50: 919–930.
- Gailey RS. Development and reliability testing of the Comprehensive High-Level Activity Mobility Predictor (CHAMP) in male servicemembers with traumatic lower-limb loss. J Rehabil Res Dev 2013; 50: 905–918.
- Anton A, Legault Z and Dudek N. Validity of the Comprehensive High-Level Activity Mobility Predictor in a heterogeneous population with lower extremity amputations. *Prosthet Orthot Int* 2020; 44: 60–65.
- Schack J, Mirtaheri P, Steen H, et al. Assessing mobility for persons with lower limb amputation: the Figure-of-Eight Walk Test with the inclusion of two novel conditions. *Disabil Rehabil* 2019; 43: 1323–1332.
- Sawers A, Kim J, Balkman G, et al. Interrater and test-retest reliability of performance-based clinical tests administered to established users of lower limb prostheses. *Phys Ther* 2020; 100: 1206–1216.
- 102. Gremeaux V, Damak S, Troisgros O, et al. Selecting a test for the clinical assessment of balance and walking capacity at the definitive fitting state after unilateral amputation: a comparative study. *Prosthet Orthot Int* 2012; 36: 415–422.
- 103. Highsmith MJ, Kahle JT, Miro RM, et al. Ramp descent performance with the C-Leg and interrater reliability of the Hill Assessment Index. *Prosthet Orthot Int* 2013; 37: 362–368.
- 104. Highsmith MJ, Kahle JT, Kaluf B, et al. Psychometric evaluation of the Hill Assessment Index (HAI) and Stair Assessment Index (SAI) in highfunctioning transfemoral amputees. *Technol Innovat* 2016; 18: 193.
- 105. Deathe AB and Miller WC. The L test of functional mobility: measurement properties of a modified version of the timed "up & go" test designed for people with lower-limb amputations. *Phys Ther* 2005; 85: 626–635.
- 106. Rushton PW, Miller WC and Deathe AB. Minimal clinically important difference of the L test for individuals with lower limb amputation: a pilot study. *Prosthet Orthot Int* 2015; 39: 470–476.
- 107. Hunter SW, Frengopoulos C, Holmes J, et al. Determining reliability of a dual-task functional mobility protocol for individuals with lower extremity amputation. *Arch Phys Med Rehabil* 2018; 99: 707–712.
- Sawers A and Hafner B. Validation of the narrowing beam walking test in lower limb prosthesis users. *Arch Phys Med Rehabil* 2018; 99: 1491–1498.
- 109. Hagberg K, Tranberg R, Zügner R, et al. Reproducibility of the Physiological Cost Index among individuals with a lower-limb amputation and healthy adults. *Physiother Res Int* 2011; 16: 92–100.
- 110. Lin SJ and Bose NH. Six-minute walk test in persons with transtibial amputation. Arch Phys Med Rehabil 2008; 89: 2354–2359.
- 111. Reid L, Thomson P, Besemann M, et al. Going places: does the two-minute walk test predict the six-minute walk test in lower extremity amputees? *J Rehabil Med* 2015; 47: 256–261.

- 112. Schoppen T, Boonstra A, Groothoff JW, et al. The timed "up and go" test: reliability and validity in persons with unilateral lower limb amputation. *Arch Phys Med Rehabil* 1999; 80: 825–828.
- Clemens SM, Gailey RS, Bennett CL, et al. The component timed-up-and-go test: the utility and psychometric properties of using a mobile application to determine prosthetic mobility in people with lower limb amputations. *Clin Rehabil* 2018, 32: 388–397.
- 114. Brooks D, Parsons J, Hunter JP, et al. The 2-minute walk test as a measure of functional improvement in persons with lower limb amputation. Arch Phys Med Rehabil 2001; 82: 1478–1483.
- 115. Brooks D, Hunter JP, Parsons J, et al. Reliability of the two-minute walk test in individuals with transtibial amputation. Validation study. *Arch Phys Med Rehabil* 2002; 83: 1562–1565.
- 116. Kark L, Odell R, McIntosh AS, et al. Quantifying prosthetic gait deviation using simple outcome measures. *World J Orthop* 2016; 7: 383.
- Boonstra A, Fidler V and Eisma W. Walking speed of normal subjects and amputees: aspects of validity of gait analysis. Prosthet Orthot Int 1993; 17: 78–82.
- 118. Datta D, Ariyaratnam R and Hilton S. Timed walking test—an allembracing outcome measure for lower-limb amputees? *Clin Rehabil* 1996; 10: 227–232.
- 119. Franchignoni F, Brunelli S, Orlandini D, et al. Is the Rivermead Mobility Index a suitable outcome measure in lower limb amputees?—a psychometric validation study. *J Rehabil Med* 2003; 35: 141–144.
- Ryall NH, Eyres SB, Neumann VC, et al. Is the Rivermead Mobility Index appropriate to measure mobility in lower limb amputees? *Disabil Rehabil* 2003; 25: 143–153.
- Gauthier-Gagnon C, Grise M and Lepage Y. The locomotor capabilities index: content validity. J Rehabil Outcomes Meas 1998; 2: 40–46.
- 122. Pruitt SD, Varni JW and Setoguchi Y. Functional status in children with limb deficiency: development and initial validation of an outcome measure. Arch Phys Med Rehabil 1996; 77: 1233–1238.
- Pruitt SD, Varni JW, Seid M, et al. Functional status in limb deficiency: development of an outcome measure for preschool children. *Arch Phys Med Rehabil* 1998; 79: 405–411.
- 124. Pruitt SD, Seid M, Varni JW, et al. Toddlers with limb deficiency: conceptual basis and initial application of a functional status outcome measure. *Arch Phys Med Rehabil* 1999; 80: 819–824.
- 125. Pierce S, Fergus A, Brady B, et al. Examination of the functional mobility assessment tool for children and adolescents with lower extremity amputations. *Pediatr Phys Ther* 2011; 23: 171–177.
- 126. Hafner BJ, Gaunaurd IA, Morgan SJ, et al. Construct validity of the prosthetic limb users survey of mobility (PLUS-M) in adults with lower limb amputation. *Arch Phys Med Rehabil* 2017; 98: 277–285.
- 127. Topuz S, Ulger O, Yakut Y, et al. Reliability and construct validity of the Turkish version of the Trinity Amputation and Prosthetic Experience Scales (TAPES) in lower limb amputees. *Prosthet Orthot Int* 2011; 35: 201–206.
- 128. Callaghan BG, Sockalingam S, Treweek SP, et al. A post-discharge functional outcome measure for lower limb amputees: test-retest reliability with trans-tibial amputees. *Prosthet Orthot Int* 2002; 26: 113–119.
- 129. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Int J Surg* 2010; 8: 336–341.
- 130. Landis JR and Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
- Fink A, Kosecoff J, Chassin M, et al. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984; 74: 979–983.
- 132. Boers M, Kirwan J and Tugwell P. The OMERACT Handook. Chapter 7—Methods for Reaching Consensus. OMERACT. 2017. Available at: http://www.omeract.org/pdf. Updated Novomber 1, 2018. Accessed March 31, 2021.
- 133. Rabin R and Charro F. EQ-5D: a measure of health status from the EuroQol Group. *Ann Med* 2001; 33: 337–343.
- 134. EuroQoL Research Foundation. *EuroQoL*. 2009. Available at: https:// euroqol.org/eq-5d-instruments/eq-5d-5l-about/. Updated April 2021. Accessed June 4, 2021.
- Sterling M and Brentnall D. Patient specific functional scale. Aust J Physiother 2007; 53: 65–66.
- Stratford P, Gill C, Westaway M, et al. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiother Can* 1995; 47: 258–263.
- 137. Podsiadlo D and Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991; 39: 142–148.
- 138. Butland R, Pang J, Gross E, et al. Two-, six-, and 12-minute walking tests in respiratory disease. *Br Med J (Clin Res Ed)* 1982; 284: 1607.
- Brooks F, Paringe V, Tonge A, et al. Patient estimates of healthcare costs in trauma and orthopaedics. *Eur J Orthop Surg Traumatol* 2013; 23: 639–642.

- Group TE. EuroQol—a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199–208.
- 141. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010; 63: 1179–1194.
- Burger H. Can the International Classification of Functioning, Disability and Health (ICF) be used in a prosthetics and orthotics outpatient clinic? *Prosthet Orthot Int* 2011; 35: 302–309.
- 143. Schaffalitzky E, Gallagher P, MacLachlan M, et al. Developing consensus on important factors associated with lower limb prosthetic prescription and use. *Disabil Rehabil* 2012; 34: 2085–2094.
- Devlin NJ, Parkin D and Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. *Health Econ* 2010; 19: 886–905.

- 145. Craig BM, Reeve BB, Brown PM, et al. US valuation of health outcomes measured using the PROMIS-29. *Value Health* 2014; 17: 846–853.
- 146. Craig BM, Reeve BB, Cella D, et al. Demographic differences in health preferences in the United States. *Med Care* 2014; 52: 307.
- 147. Kristensen MT, Nielsen AØ, Topp UM, et al. Development and psychometric properties of the Basic Amputee Mobility Score for use in patients with a major lower extremity amputation. *Geriatr Gerontol Int* 2018; 18: 138–145.
- Wood-Dauphinee S, Opzoomer M, Williams JI, et al. Assessment of global function: the Reintegration to Normal Living Index. Arch Phys Med Rehabil 1988; 69: 583–590.
- 149. International Society of Prosthetics and Orthotics. COMPASS (Consensus Outcome Measures for Prosthetic and Amputation Services) User Guide. (n.d.). Available at: www.ispoint.org/wp-content/uploads/2022/03/ispo\_compass\_user\_guide\_2021.pdf. Accessed July 8, 2022.