OPEN ACCESS

Development of a core outcome set for lateral elbow tendinopathy (COS-LET) using best available evidence and an international consensus process

Marcus Bateman ,[1] Jonathan P Evans ,[2,3] Viana Vuvan,[4] Val Jones,[5] Adam C Watts,[6] Joideep Phadnis,[7,8] Leanne M Bisset,[9,10] Bill Vicenzino ,[4] COS-LET Authorship Group

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/bjsports-2021-105044).

For numbered affiliations see end of article.

Correspondence to
Marcus Bateman, Orthopaedic Department, University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK; marcus.bateman@nhs.net

MB and JPE are joint first authors.

Accepted 14 January 2022
Published Online First 8 February 2022

ABSTRACT

Objectives To develop a core outcome set for lateral elbow tendinopathy (COS-LET) and to provide guidance for outcome evaluation in future studies.

Methods We implemented a multi-stage mixed-methods design combining two systematic reviews, domain mapping of outcome measurement instruments to the core domains of tendinopathy, psychometric analysis of instruments, two patient focus groups and a Delphi study incorporating two surveys and an international consensus meeting. Following the OMERACT guidelines, we used a 70% threshold for consensus.

Results 38 clinicians/researchers and 9 patients participated. 60 instruments were assessed for inclusion. The only instrument that was recommended for the COS-LET was Patient Rated Tennis Elbow Evaluation (PRTEE) for the disability domain. Interim recommendations were made to use: the PRTEE function subscale for the function domain; PRTEE pain subscale items 1, 4 and 5 for the pain over a specified time domain; pain-free grip strength for the physical function capacity domain; a Numerical Rating Scale measuring pain on gripping for the pain on activity/loading domain; and time off work for the participation in life activities domain. No recommendations could be made for the quality-of-life, patient rating of condition and psychological factors domains.

Conclusions The COS-LET comprises the PRTEE for the disability domain. Interim-use recommendations included PRTEE subscales, time off work, pain-free grip strength and a Numerical Rating Scale measuring pain on gripping. Further work is required to validate these interim measures and develop suitable measures to capture the other domains.

INTRODUCTION

Background and objectives

Pain arising from the tendons on the lateral side of the elbow is common in adults, particularly in middle age.[1] Historically, it has been known by various names such as lateral epicondylitis or tennis elbow, but the current accepted description is lateral elbow tendinopathy (LET).[2] It is acknowledged that there is substantial heterogeneity of outcome measure instrument use in elbow research and specifically for LET.[3] With no clear consensus on which instruments most accurately represent a patient's LET-related health status, comparison of effectiveness research and evidence synthesis/meta-analysis has been hampered.

In 2019, an international group of experts in the field of tendinopathy (International Scientific Tendinopathy Symposium Consensus (ICON) Group), comprising researchers, healthcare professionals and patients, published a consensus document defining the nine health-related core domains of tendinopathy. That group recommended researchers and clinicians measure outcomes for specific regional tendinopathies against these domains.[4]

The aim of this project was to develop a core outcome set (COS) for LET that maps to the nine domains. A COS is a minimal set of outcome measures to be used in future research and clinical practice involving people with LET. It enables meta-analysis of findings from different studies using a consistent set of measures. To be included in a COS, measures need to be both practical to perform (based on cost, patient burden and availability) and of high quality (valid, responsive, reliable, interpretable and of acceptable burden for patients and investigators).[5] The result will be a minimum set of outcome measurement instruments to be used in future LET research that allows direct comparison between different studies across the nine domains.

Scope

This COS relates to all adults diagnosed with LET and applies to interventional research (including surgical and non-surgical) and longitudinal assessment. The COS will only apply to the English language.

METHOD

We designed the project following the COSMIN–COMET guideline.[6] We developed a COS that was based on a consensus of perspectives gained from healthcare professionals with expertise in LET and patients with the condition. This involved a multi-stage stepwise process, which started with identifying the instruments used in studies of LET by updating a previous systematic evaluation of patient-rated outcomes for LET.[7] These instruments were then mapped by the steering committee to the nine core tendinopathy domains.[4] The mapped outcome measurement instruments were then subjected to the OMERACT truth (part a) and feasibility filters[8] by participants in the first round of a Delphi survey. We (MB and JPE) then systematically evaluated the

Check for updates

To cite: Bateman M, Evans JP, Vuvan V, *et al*. *Br J Sports Med* 2022;**56**:657–666.

psychometric properties of the included instruments—applying the OMERACT truth (part b) and discrimination filters,[8] using the EMPRO tool.[9] This information then formed the basis of the second Delphi survey, which was conducted to make recommendations for a COS-LET. Focus groups were conducted with patients to review findings after Delphi survey 2. The results of the surveys and focus groups were then reviewed and discussed by participants at an international consensus online meeting (Delphi stage 3), before voting to determine the final COS-LET. The study was led by an international eight-person steering committee with expertise in LET—comprising a mix of junior and senior researchers and clinicians from surgical and physiotherapy backgrounds.

Protocol/ registry entry

We registered the project with the Core Outcome Measures in Effectiveness Trials Initiative (http://www.comet-initiative.org/Studies/Details/1497) and published the protocol in an *open access journal*. (https://trialsjournal.biomedcentral.com/articles/10.1186/s13063-021-05291-9). This report follows the COS Standards for Reporting checklist.[10]

Participants

The Delphi study population comprised of experienced clinicians and researchers nominated by the steering committee, identified by their reputation as elbow clinical specialists or prior publications related to LET. Additionally, a search of the Expertscape[11] and SCOPUS[12] databases by author and filtered by the terms 'tennis elbow' and 'trial' identified a list of other researchers to approach. Representation from a range of nationalities, with a spread of ethnicity and sex was ensured.

Patient representatives were invited by the clinicians on the steering committee.

Information sources

In order to comprehensively evaluate all outcome measurement instruments used in research of LET, we systematically reviewed the literature. To do this, we updated the 2019, Evans *et al*,[7] systematic review of English language instruments used in surgical and non-surgical trials for LET (census date: 1 May 2017). The search results were screened initially by title and abstract by two reviewers (MB and JPE) independently of each other using the online Rayyan tool[13]—any disagreements were discussed and reconciled. We included all study designs except research protocols, case studies and small case series of less than five patients. One hundred and ninety nine full texts from the original search and 93 from the updated searches (to February 2020) were screened down to 256 papers for data extraction—providing a comprehensive list of instruments used in LET research (figure 1). Extracted data included all outcome instruments used, number of patients included in the study and full details of any novel instruments.

Consensus process and outcome scoring

The retrieved instruments were then submitted to a stepwise consensus process that mapped them to the core tendinopathy domains. The mapped instruments where then used to construct the first survey. The instruments agreed to in that survey were then evaluated for their psychometrics—the results of which were included in a second survey. Results of the surveys were discussed in two patient focus groups. Finally, a consensus meeting reviewed and discussed findings before voting on the final COS-LET.

Instruments mapped to domains: the steering committee members mapped each instrument to the nine core tendinopathy domains[4]: patient rating of condition; participation in life activities (day to day, work and sport); pain on activity/loading;

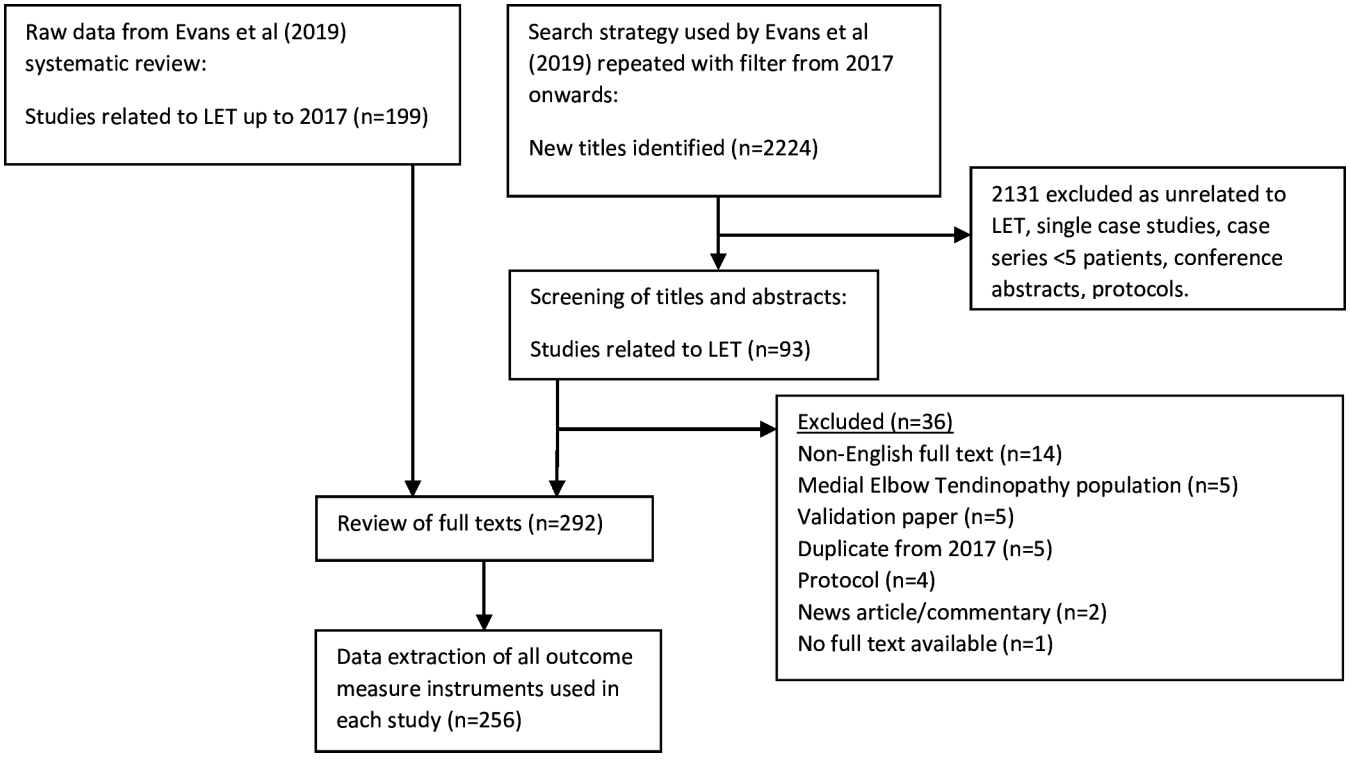


Figure 1 Adapted PRISMA flowchart: to review the outcome measure instruments used in all LET studies. LET, lateral elbow tendinopathy; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

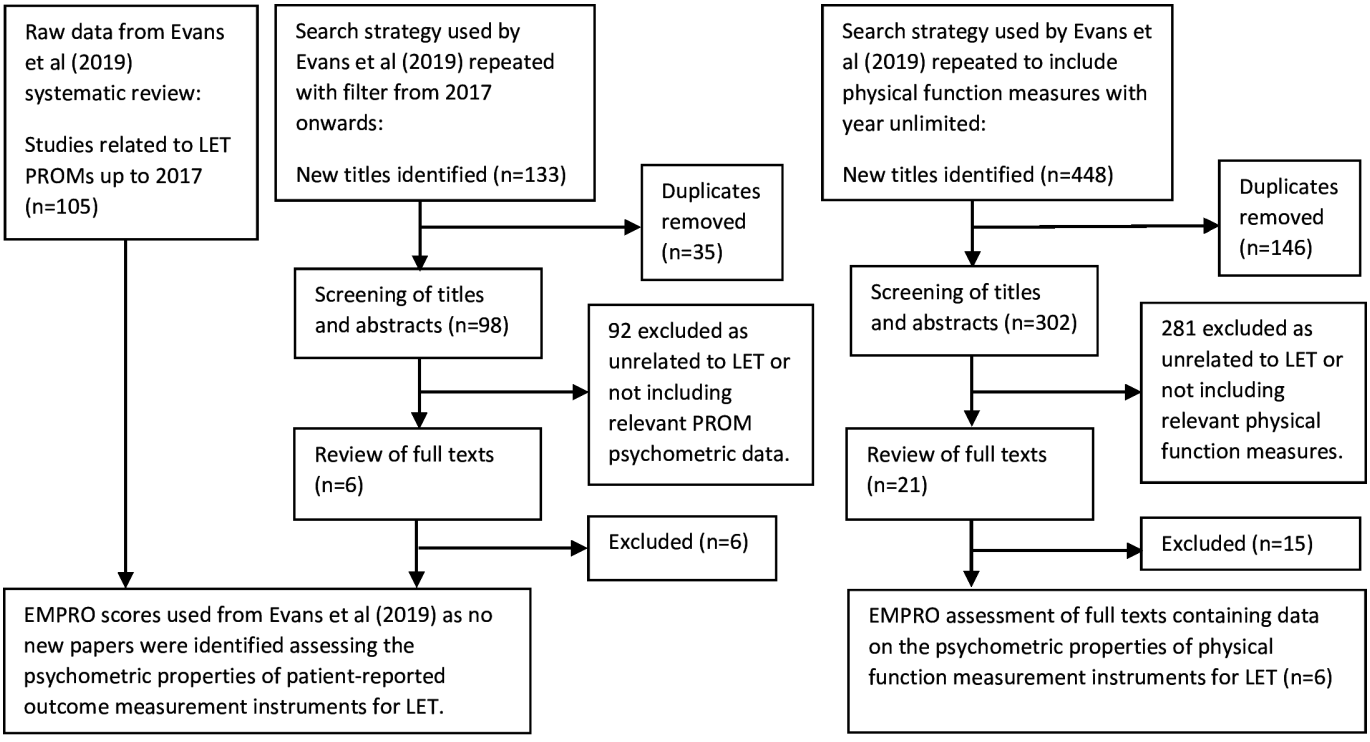


Figure 2 Adapted PRISMA flowchart: to review the psychometric properties of the instruments included after Delphi round 1. LET, lateral elbow tendinopathy; PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; PROMs, Patient-Reported Outcome Measures.

function; psychological factors; physical function capacity; disability; quality of life; and pain over a specified time. Each instrument was mapped by two steering committee members independently, then compared and reconciled if needed—using the instrument's published development study or manual.

Survey 1: each instrument and its reference document, including original development study and/or manual, were presented per mapped domain to participants in the first online survey using Qualtrics (Provo, Utah, USA). Participants were asked to respond yes/no/unsure whether each instrument was a truthful measure of the domain (valid), feasible to use clinically and whether it should be included in the COS-LET.

Outcomes were scored using the OMERACT traffic light system,[8] whereby responses achieving <30% agreement were rated red and excluded; those achieving ≥70% were rated green and included; and those achieving 30%–69% were rated amber, inconclusive but not excluded.

Psychometric evaluation of instruments: following exclusion of instruments (<30% agreement in survey 1) and inclusion of any new instruments proposed by respondents, the OMERACT truth (part b) and discrimination filters were applied by two members of the steering committee (MB and JPE). This involved an update of Evans *et al*'s[7] systematic review to identify instrument development or validation studies (figure 2). These members used the EMPRO tool, separately to each other, to assess the psychometric properties (construct validity, reliability, repeatability, responsiveness and interpretability) of each instrument, then meeting to discuss points of contention.[9] These were resolved without the need of a third assessor. The steering committee then voted anonymously, using the OMERACT traffic light system, on whether each instrument should be considered 'Good to go', 'A concern/more work needed' or 'Stop, do not continue'.[8] This voting stage was included to ratify the psychometric evaluation process and was guided by an EMPRO score threshold of 40% for inclusion.

Survey 2: during the second Delphi survey, participants were presented with the findings of Delphi survey 1 (online supplemental file 1), showing the traffic light rating of each instrument within their associated matched domain, and subsequent outcome of the truth (part b) and discrimination filters (online supplemental file 2). Participants were asked to rate instruments that achieved a nominal EMPRO score of ≥40% for inclusion in the final COS-LET (yes/no/unsure). Those instruments that were no or unsure for the final COS-LET, had no psychometric data or had an EMPRO score of <40% were rated for interim use (yes/no). The responses were analysed and those instruments achieving <30% of votes were excluded.

Patient focus groups: results of the Delphi stages, inclusive of survey 2, were then discussed at an online patient focus group for UK patients and another for Australian patients. Patients were asked to provide their insights/perspectives on the decisions to date and to ratify any instruments voted ≥70%.

Final consensus meeting: participants attended an online consensus meeting to discuss the findings of the Delphi process to date (including patient focus group outcomes) and to vote on outcome measures in the COS-LET and for interim use. A report of the results of previous surveys and patient focus groups (online supplemental file 3) was provided to the participants 2 weeks prior.

Consensus definition

For each domain, instruments voted for by ≥70% of participants in both surveys and at the meeting were included in the COS-LET. For domains where no instruments were agreed, interim suggestions were proposed based on a green light from Delphi

Table 1 Participant characteristics (n (%) unless otherwise stated)

Characteristics	Clinicians/researchers			Patients
	Survey 1 (n=37)	Survey 2* (n=37)	Meeting (n=31)	Survey 1/focus group (n=9)
Sex: male	25 (67.6)	25 (67.6)	22 (71.0)	4 (44.4)
Age: median (IQR; minimum–maximum), years	51 (43–57; 34–68)	51 (43–55; 34–68)	51 (43–53; 34–68)	48 (37–53; 26–59)
Role				
Clinician	2 (5.4)	2 (5.4)	2 (6.4)	NA
Researcher	5 (13.5)	5 (13.5)	4 (12.9)	NA
Clinician researcher	30 (81.1)	30 (81.1)	25 (80.7)	NA
Highest academic qualification				
PhD	21 (56.8)	21 (56.8)	17 (54.8)	–
Master	6 (16.2)	6 (16.2)	5 (16.1)	2 (22.2)
Doctor of Medicine	7 (18.9)	7 (18.9)	7 (22.6)	–
Postgraduate diploma/certificate	–	–	–	2 (22.2)
Bachelor	3 (8.1)	3 (8.1)	2 (6.5)	4 (44.4)
No university qualification	–	–	–	1 (11.1)
Profession				
Physiotherapy	16 (43.2)	16 (43.2)	14 (45.2)	NA
Orthopaedic surgeon	14 (37.8)	14 (37.8)	12 (38.7)	NA
Sports and exercise medicine physician	3 (8.1)	3 (8.1)	2 (6.4)	NA
Not specified	3 (8.1)	3 (8.1)	2 (6.4)	NA
Rheumatologist	1 (2.7)	1 (2.7)	1 (3.2)	NA
Therapy radiographer	–	–	–	1 (11.1)
Health information technology	–	–	–	1 (11.1)
Non-healthcare professional	–	–	–	7 (77.8)
Lateral elbow tendinopathy				
Current case	1 (2.7)	1 (2.7)	1 (3.2)	5 (55.6)
History	10 (27.0)	11 (29.7)	9 (29.0)	6 (66.7)
Country where work†				
Europe	20 (54.1)	20 (54.1)	16 (51.2)	5 (55.6)
Australia	11 (29.7)	10 (27.0)	8 (25.8)	4 (44.4)
North America	5 (13.5)	5 (13.5)	5 (16.1)	–
Asia	1 (2.7)	2 (5.4)	2 (6.5)	–

*1 person from Australia did survey 1 but not 2; another did survey 2 not 1 (technical issues).
†Countries grouped per continent as follows: Europe=Belgium, Finland, Greece, Italy, the Netherlands, Norway, Spain, Sweden, Turkey (Istanbul) and UK; North America=Canada and USA; and Asia=India and Israel.

survey 1 (≥70%), and amber light from Delphi survey 2 (30%–69%) and ≥70% agreement from the consensus meeting vote.

RESULTS

We commenced this study in January 2020, with regular steering committee working meetings to plan and design data collection. Data collection was completed at the consensus meeting on 5 May 2021.

Protocol deviations

The only deviation from the published protocol was that patient focus groups were conducted, rather than one-to-one interviews. This decision was taken to allow for patient interaction and group discussion, the impact of which on our findings is likely low to negligible.

Participants

We invited 58 healthcare professionals of which 40 agreed to participate, 7 did not agree (retired (2) and no reason given (5)) and 11 did not respond (maternity leave out of office message (1) and unknown reason (10)). Thirty eight engaged with the process and 2 withdrew. Thirty six (90%) of the clinicians/researchers who agreed to participate fully completed both surveys (table 1), 2 (5%) completed one survey and 31 (84%) of those completing surveys attended the online meeting. The clinician/researcher cohort consisted mainly of physiotherapists or orthopaedic surgeons, located in Europe or Australia, and had research higher degree training.

Nine patients from the UK and Australia participated in the study—7 completed the first survey and 5 participated in the focus groups.

Outcomes

Sixty unique instruments were identified from the first systematic review and included in survey 1 (table 2). From survey 1, three measures: pain on gripping, Patient Rated Tennis Elbow Evaluation (PRTEE)[14–16] and Quick Disability of Arm, Shoulder and Hand Questionnaire (DASH),[17] were the only ones reaching ≥70% for both the patient and the clinician/researcher groups. The Patient Specific Function Scale[18] (function domain) and Tennis Elbow Functional Scale[19] (both pain domains) were voted to be in the COS-LET by patients who commented favourably on the scores' item level face validity—however, the lack of robust psychometric evaluation in LET populations precluded

Table 2 Showing the prevalence of use of individual instruments (shown as the number of previous studies and %) and responses to Delphi survey 1 (shown as the number of responses and %)

		Clinicians and researchers				Patients (n=7)		
Domain: measure	**Prevalence**	***n**	**a**	**b**	**c**	**a**	**b**	**c**
Patient rating of condition								
Global Perceived Effect Score	1 (<1)	39	22 (56)	28 (72)	36 (92)	5 (71)	6 (86)	6 (86)
GROC	9 (4)	39	25 (64)	29 (74)	34 (87)	4 (57)	6 (86)	6 (86)
Patient Satisfaction Scale	29 (11)	39	20 (51)	25 (64)	35 (90)	5 (71)	6 (86)	6 (86)
Roles and Maudsley†	15 (6)							
Participation (daily activities, work and sport)								
Return to sport	5 (2)	39	15 (38)	25 (64)	32 (82)	5 (71)	5 (71)	6 (86)
Time off work	16 (6)	39	21 (54)	31 (79)	30 (77)	4 (57)	4 (57)	6 (86)
Total Elbow Scoring System	1 (<1)	39	12 (31)	14 (36)	21 (54)	4 (57)	4 (57)	5 (71)
OSTRC question 1†	0							
Pain on activity/loading								
Tennis Elbow Functional Scale	2 (1)	39	11 (28)	24 (62)	22 (56)	6 (86)	6 (86)	6 (86)
Thomsen Test (VAS pain resisted wrist extension)	7 (3)	39	14 (36)	28 (72)	28 (72)	5 (71)	5 (71)	6 (86)
VAS chair pick-up	2 (1)	39	9 (23)	20 (51)	23 (59)	2 (29)	6 (86)	4 (57)
VAS pain during activity	49 (19)	39	31 (79)	35 (90)	34 (87)	4 (57)	5 (71)	6 (86)
VAS pain during elbow movement	5 (2)	39	7 (18)	13 (33)	27 (69)	2 (29)	3 (43)	6 (86)
VAS pain on gripping	16 (6)	39	28 (72)	37 (95)	35 (90)	5 (71)	5 (71)	7 (100)
VAS pain at work	11 (4)	39	15 (38)	23 (59)	29 (74)	3 (43)	3 (43)	4 (57)
Pain Free Functional Index	2 (1)	39	7 (18)	20 (51)	23 (59)	3 (43)	3 (43)	6 (86)
PRTEE†‡	78 (30)	38	24 (63)	30 (79)	27 (71)	7 (100)	7 (100)	7 (100)
Function								
Patient Specific Functional Scale	1 (<1)	38	11 (29)	24 (63)	20 (53)	5 (71)	6 (86)	6 (86)
Upper Extremity Functional Scale	5 (2)	38	17 (45)	26 (68)	26 (68)	4 (57)	4 (57)	7 (100)
VAS function	6 (2)	38	17 (45)	24 (63)	29 (76)	2 (29)	3 (43)	4 (57)
PRTEE†‡	78 (30)							
Psychological factors								
Hospital Anxiety and Depression Scale	3 (1)	38	14 (37)	20 (53)	19 (50)	2 (29)	2 (29)	6 (86)
Tampa Scale of Kinesophobia	4 (2)	38	15 (39)	19 (50)	21 (55)	5 (71)	6 (86)	5 (71)
STAI trait†	0							
Nottingham Health Profile†	3 (1)							
Physical function capacity (eg, strength)								
Grip strength (maximum)	93 (36)	38	18 (47)	31 (82)	26 (68)	6 (86)	6 (86)	7 (100)
Pain-free grip strength	42 (16)	38	25 (66)	30 (79)	27 (71)	6 (86)	6 (86)	7 (100)
Elbow ROM	6 (2)	38	4 (11)	12 (32)	25 (66)	3 (43)	2 (29)	7 (100)
Disability								
Andrews-Carson	2 (1)	38	0 (0)	3 (8)	10 (26)	1 (14)	1 (14)	3 (43)
ASES score	5 (2)	38	4 (11)	17 (45)	13 (34)	3 (43)	3 (43)	5 (71)
Broberg and Morrey Rating System	1 (<1)	38	1 (3)	8 (21)	10 (26)	2 (29)	3 (43)	5 (71)
DASH	77 (30)	38	15 (39)	26 (68)	17 (45)	4 (57)	4 (57)	5 (71)

Continued

Table 2 Continued

Domain: measure	Prevalence	Clinicians and researchers				Patients (n=7)		
		*n	a	b	c	a	b	c
HAND10	1 (<1)	38	3 (8)	17 (45)	23 (61)	4 (57)	4 (57)	6 (86)
Japanese Orthopaedic Association Elbow Score	2 (1)	38	4 (11)	17 (45)	16 (42)	0 (0)	0 (0)	4 (57)
Laitinen Questionnaire	1 (<1)	38	5 (13)	14 (37)	15 (39)	1 (14)	1 (14)	4 (57)
Liverpool Elbow Score	1 (<1)	38	1 (3)	13 (34)	13 (34)	2 (29)	3 (43)	3 (43)
Mayo Elbow Performance Score	30 (12)	38	3 (8)	12 (32)	15 (39)	0 (0)	0 (0)	2 (29)
Nirschl	18 (7)	38	7 (18)	22 (58)	21 (55)	4 (57)	4 (57)	4 (57)
Nottingham Health Profile	3 (1)	38	1 (3)	6 (16)	12 (32)	0 (0)	0 (0)	0 (0)
Oxford Elbow Score	4 (2)	38	9 (24)	20 (53)	21 (55)	4 (57)	5 (71)	6 (86)
Patient-Rated Wrist Evaluation Questionnaire	1 (<1)	38	5 (13)	17 (45)	19 (50)	4 (57)	4 (57)	6 (86)
PRTEE†‡	78 (30)	37	27 (73)	34 (92)	32 (86)	6 (86)	6 (86)	6 (86)
Quick DASH	37 (14)	38	27 (71)	29 (76)	31 (82)	7 (100)	7 (100)	7 (100)
Total Elbow Scoring System	1 (<1)	38	4 (11)	14 (37)	20 (53)	2 (29)	2 (29)	4 (57)
Roles and Maudsley	15 (6)	38	2 (5)	11 (29)	22 (58)	1 (14)	2 (29)	4 (57)
Quality of life								
EQ5D	9 (4)	37	19 (51)	22 (59)	25 (68)	4 (57)	4 (57)	6 (86)
SF-36	13 (5)	37	6 (16)	23 (62)	9 (24)	1 (14)	5 (71)	2 (29)
SF-12	4 (2)	37	15 (41)	25 (68)	25 (68)	3 (43)	5 (71)	4 (57)
Nottingham Health Profile†	3 (1)							
WHOQol BREF†	0							
SMFA†	0							
Pain over a specified timeframe								
VAS night pain	9 (4)	37	14 (38)	26 (70)	32 (86)	2 (29)	4 (57)	6 (86)
VAS pain defined time period	22 (9)	37	25 (68)	30 (81)	29 (78)	2 (29)	3 (43)	5 (71)
VAS pain at rest	52 (20)	37	21 (57)	25 (68)	31 (84)	1 (14)	2 (29)	5 (71)
Tennis Elbow Functional Scale	2 (1)	37	8 (22)	22 (59)	21 (57)	6 (86)	6 (86)	6 (86)
Measures from review not mapped								
Analgesic use	2 (1)	37	16 (43)			1 (14)		
Canadian Occupational Performance Measure	1 (<1)	37	2 (5)			1 (14)		
Cold Pain Threshold	5 (2)	37	3 (8)			1 (14)		
EMG	5 (2)	37	0 (0)			2 (29)		
Gothenburg QoL Instrument	1 (<1)	37	0 (0)			0 (0)		
MRI appearance	5 (2)	37	3 (8)			1 (14)		
ORI-TETS	3 (1)	37	1 (3)			1 (14)		
Placzek Score	1 (<1)	37	4 (11)			2 (29)		
Pressure Pain Threshold	24 (9)	37	7 (19)			1 (14)		
US appearance	13 (5)	37	4 (11)			1 (14)		
UoP-PFQ	2 (1)	37	1 (3)			1 (14)		
VAS pain on palpation	12 (5)	37	4 (11)			2 (29)		

Continued

Table 2 Continued

Domain: measure	Prevalence	Clinicians and researchers				Patients (n=7)		
	*n	a	b	c		a	b	c
VAS pain overall	101 (39)	37	9 (24)			2 (29)		
WL-26	1 (<1)	37	2 (5)			4 (57)		

a: this measure should be in the core outcome set for lateral elbow tendinopathy.
b: this measure has content and face validity.
c: the measure is clinically feasible.
Green (yes responses: ≥70%) and amber (unsure responses: 30%–69%) were progressed to subsequent stages of the Delphi, whereas red (<30%) were removed.
*2 participants partially completed the first survey and their data was included herein.
†Outcome measures suggested by participants, noting PRTEE was not new to the list.
‡Removal of one participant's response who was designer of the measure.
ASES, American Shoulder and Elbow Surgeons; DASH, Disability of Arm, Shoulder and Hand Questionnaire; EMG, electromyography; GROC, Global Rating of Change; ORI-TETS, Orthopaedic Research Institute-Tennis Elbow Testing System; OSTRC, Oslo Sport Trauma Research Centre Questionnaire; PRTEE, Patient Rated Tennis Elbow Evaluation; QoL, quality of life; ROM, range of movement; SF-12, 12-item Short Form Health Survey; SF-36, 36-item Short Form Health Survey; SMFA, Short Musculoskeletal Function Assessment; STAI, State–Trait Anxiety Inventory; UoP-PFQ, University of Peloponnese Pain, Functionality and Quality of Life Questionnaire; US, ultrasound; VAS, Visual Analogue Scale; WHOQol BREF, WHO Quality of Life Abbreviated Questionnaire; WL-26, Work Limitation 26-item Questionnaire.

their inclusion following clinician/researcher evaluation (table 2). Twenty four instruments were excluded after this survey as they received <30% of votes for inclusion in the COS-LET by both patients and clinicians/researchers. Participants proposed an additional seven instruments (table 2).

Core outcome set

A search for studies of the measures with ≥30% responses from survey 1 revealed that only 8/21 (38%) instruments had been evaluated for their psychometric properties in specific LET populations. These measures were submitted to analysis with the EMPRO tool—scoring between 25% and 73% (table 3). The additional instruments proposed by responders in survey 1 had no psychometric data for the LET population and were not considered further in the development of the COS-LET. Seven instruments had an EMPRO score of ≥40 and received ≥70% of votes in survey 1 from either patients or clinicians/researchers—they were: DASH,[17 20–22] Quick DASH,[17 23 24] Oxford Elbow Score,[25 26] PRTEE,[14–16] Tennis Elbow Functional Scale,[19] as well as grip strength.[27–29] These were then independently assessed by the steering committee using the OMERACT truth (part b) and discrimination filters (results given in online supplemental file 2), which along with results from survey 1, were available to inform clinician's/researcher's decisions in survey 2. The results of survey 2 are presented in table 3—only the PRTEE met the threshold for inclusion in the COS-LET for the disability domain, which was ratified at the consensus meeting.

Interim recommendations

Where there was no measure in a domain that met the criteria for the COS-LET, we considered measures that reached ≥70% hurdle in survey 1—aiming to recommend one per domain to be used in the interim and as a direction for future research (table 3, Part B).

Of the measures that had psychometric data—that is, Tennis Elbow Functional Scale, maximum grip strength, pain-free grip strength and the PRTEE pain and function subscales—only the latter was voted as an interim measure in survey 2 (table 3). The patients (focus groups) agreed that items/subscales of the PRTEE (for pain on loading/activity, function and pain over specified time domains) and pain-free grip strength (for physical function capacity domain) were relevant to their condition. The consensus meeting decided that relevant subscale items from the PRTEE would be recommended as interim measures for function and pain over specified time domains. Pain-free grip strength was selected over maximum grip strength as it was thought to be more clinically and patient relevant. The meeting decided that pain on gripping, which had support in survey 1 but with no psychometric data, would be the preferred measure for pain on loading/activity domain, instead of the relevant subscale items from the PRTEE (table 3).

Of the other measures that did not have psychometric data (see table 3), only time off work was voted as an interim measure in survey 2—it was ratified at the consensus meeting.

Notably, after the consensus meeting, there were three domains for which no interim measures were agreed. These were: quality of life, participant overall rating of condition and psychological.

DISCUSSION

This is the first attempt to determine the minimum COS-LET. We reached agreement for an outcome measure for one of the nine tendinopathy domains—PRTEE for disability. Although the PRTEE has been found to be psychometrically robust, we

Table 3 Results of the clinimetric evaluation (EMPRO), second survey, patient focus group and final consensus meeting decision arranged in reverse chronology across the table—for each measure per domain. Data are frequency count (%) unless otherwise specified. The COS-LET is in 'Part A' and the interim suggestions of measures that might be used and studied are in 'Part B'

Part A: core outcome to be used in clinical trials and cohort studies									
Domain	**Measure**	**Decision**	**Votes**	**Patient**	**Survey 2**				**EMPRO**
			Yes*	**Agreed**	**In**	**Out**	**Unsure**	**Interim**	**Score %**
Disability	PRTEE	√	29 (100)	√	26 (70.3)	5 (13.5)	6 (16)	n/a	57.0
	DASH	x			3 (8.1)	25 (67.6)	9 (24.3)	n/a	66.9
	Quick DASH	x			22 (59.5)	9 (24.3)	6 (16.2)	n/a	72.5
	Oxford Elbow Score	x			6 (16.2)	19 (51.4)	12 (32.4)	n/a	66.6
Part B: interim suggestion for use in clinical trials and as a focus of future clinimetric research									
Domain	**Measure**	**Decision**	**Votes**	**Patient**	**Survey 2**				**EMPRO**
			Yes*	**Agreed**	**In**	**Out**	**Unsure**	**Interim**	**Score %**
Function	PRTEE—relevant items	√	30 (96.8)	√	24 (64.9)	5 (13.5)	8 (21.6)	33 (89.2)	n/a
Pain over specified time	PRTEE pain subscale items 1, 4 and 5†	√	30 (96.8)						n/a
	Tennis Elbow Functional Scale	x	n/a		3 (8.1)	27 (73)	7 (18.9)	9 (24.3)	41.7
Physical function capacity	Pain-free grip strength	√	26 (86.7)	√	15 (40.5)	11 (29.7)	11 (29.7)	24 (64.9)	32.9
	Maximum grip strength	x	n/a		6 (16.2)	17 (46)	14 (37.8)	12 (32.4)	25.1
Pain on loading/activity	Pain on gripping†	√	25 (83.3)		n/a	n/a	n/a	n/a	n/a
	Tennis Elbow Functional Scale	x	–		4 (10.8)	27 (73)	6 (16.2)	7 (18.9)	41.7
	PRTEE pain subscale items 2 and 3	x	19 (65.5)	√	24 (64.9)	7 (18.9)	6 (16.2)	31 (83.8)	n/a
Participation	Time off work	√	22 (73.3)		n/a	n/a	n/a	26 (70.3)	n/a
	Time off sport	x	18 (60)		n/a	n/a	n/a	21 (56.8)	n/a
QoL	EQ5D	x	20 (69)		n/a	n/a	n/a	22 (59.5)	n/a
	SF-12	x	6 (20.7)		n/a	n/a	n/a	14 (37.8)	n/a
Participant Rating of Condition	GROC	x	20 (66.7)	√‡	n/a	n/a	n/a	21 (56.8)	n/a
	Global Perceived Effect Score	x	n/a		n/a	n/a	n/a	13 (35.1)	n/a
	Patient Satisfaction Scale	x	n/a		n/a	n/a	n/a	17 (46)	n/a
Psychological	Tampa Scale of Kinesiophobia	x	10 (34.5)	√‡	n/a	n/a	n/a	16 (43.2)	n/a
	Hospital Anxiety and Depression Scale	x	n/a		n/a	n/a	n/a	14 (37.8)	n/a

*Note that not all 31 attendees voted on all items (at least, 29 voted on 4 items)—due to time zone differences. See appendices for full data.
†Pain on gripping was voted in survey 1, had no clinimetric evidence but was strongly supported for interim use—noting there were 2 (6.7%) participants unsure.
‡Patients were asked for their opinions on which of the measures most closely measured their condition domain and was feasible clinically.
COS-LET, core outcome set for lateral elbow tendinopathy; DASH, Disabilities of the Arm, Shoulder and Hand Questionnaire; GROC, Global Rating of Change; PRTEE, Patient Rated Tennis Elbow Evaluation; QoL, quality of life; SF-12, 12-item Short Form Health Survey.

note that the total EMPRO score was lower than other measures considered for inclusion. As previously identified, as the PRTEE was developed without patient involvement, its EMPRO score is, therefore, reduced as a consequence.[7] However, the expert and patient groups within this COS development were concordant in their agreement that the PRTEE preferentially aligned with the disability domain. It was not possible to include measures in a COS-LET for the remaining domains, because there was either no instrument or a lack of instrument validation. In the interim, we decided on measures to recommend for validation studies and use in trials of LET.

The PRTEE function subscale was selected as an interim measure for the function domain, but concerns were raised that it mainly queries basic tasks and not higher-level tasks required in sports. This requires further investigation, along with the psychometric properties of the subscale. Patient responses from survey 1 favoured the Patient Specific Functional Scale for the function domain, but this was not supported by clinicians/researchers. This scale can be tailored to athletic/high-level tasks so may be an area for future investigation.

The pain over a specified time domain had no suitable instruments following survey 2. We resolved at the consensus meeting to recommend three questions from the PRTEE (pain at rest, least and worst over the last week) in the interim. The PRTEE had already been accepted for the COS-LET, thereby minimising patient burden—a key priority identified in the patient focus groups; however, the psychometric properties of these PRTEE items require further assessment.

Pain-free grip strength was recommended as an interim measure for physical function capacity domain, but it was not included in the final COS-LET due to limited validation in LET populations. With clear stakeholder approval, further validation work should be prioritised.

Measuring the pain on activity/loading domain was the source of a lengthy discussion, because the two options with sufficient psychometric evidence failed to reach consensus. Discussion then moved to rating pain on gripping as an interim measure. Gripping was raised as a common pain provoking activity in the patient focus groups. It had been voted in survey 1 to be in the COS-LET by patients and clinicians/researchers, but due to a lack of research it was not voted in survey 2. In the meeting, concerns were raised about standardising the gripping task—as this would be difficult across sexes and different work/sport contexts. We resolved to recommend using a Numerical Rating Scale to record pain during gripping in the interim and to prioritise its validation.

Time off work was recommended as an interim measure of the participation domain, but there were concerns regarding the definition of work and whether this was applicable to patients who were retired, unemployed, students or full-time parents/

carers. Due to the value of using time off work as part of health economic evaluation, however, it was agreed that it should still be used in the interim. It was recommended that future research should consider how this measure may be individualised to a patient's context.

No recommendations could be made for the domains of quality of life, patient rating of condition and psychology, primarily because there were no measures that had been validated for use in LET. For quality of life, the EQ5D narrowly missed interim selection and provoked mixed feelings from the patient focus groups. It was considered useful from a health economics perspective, for the calculation of quality adjusted life years.[30 31] Of the three instruments considered for the patient rating of condition domain, the Global Rating of Change (GROC) was preferred by the patient focus groups and is also regarded as an appropriate anchor for responsiveness analysis of other outcome measures.[32 33] The Single Assessment Numerical Evaluation[34] was proposed as an alternative option with the advantage that it can be used pretreatment and posttreatment, rather than relying on symptom recall, like the GROC. For the psychology domain, patients indicated the Tampa Scale of Kinesiophobia[35] was more representative of their condition than anxiety and depression scales. Future studies should investigate the psychometric properties of these instruments for the LET population.

Strengths and Limitations

A strength of this study is that we included experts in LET (ie, patients and clinicians/researchers) from across the globe and followed a robust methodology that was published in advance.

There are several limitations that need to be considered in implementing the findings of this study. First, the COS-LET was developed on the basis of previously agreed on core domains for tendinopathy and are dependent on that work—any changes to those core domains will require revision of the COS-LET. Second, we were unable to recommend outcome measures for all of the core domains of tendinopathy—in which case we made interim recommendations. These interim recommendations should not be misconstrued as being part of the COS-LET, because they were made on the basis of opinions of participants without appropriate instrument validation. Third, we restricted the study to English outcome measures—using the COS-LET in non-English language situations requires validation. Fourth, we did not include a patient in our steering committee.

Areas for future research

Future research is required to establish valid and feasible measures across all health-related tendinopathy domains in patients who have LET. We have identified some targets herein: PRTEE subscales/items, pain on gripping, GROC, EQ5D and Tampa Scale of Kinesiophobia.

CONCLUSION

The PRTEE should be used in all future studies related to LET—especially for the disability domain. Time off work, pain-free grip strength and a Numerical Rating Scale measuring pain on gripping should also be used until future studies recommend alternative, more robust, measures of participation in life activities, physical function capacity and pain on activity/loading. A COS-LET Tool, containing these recommended measures, has been composed (see online supplemental file 4). Further work is required to (a) validate the interim measures for use in research involving persons with LET and (b) develop/validate suitable measures of the patient rating of condition, quality of life and psychological factors domains.

Key messages

What is already known?

⇒ Core outcome sets (COSs) are recommended for research and clinical practice to facilitate comparison and meta-analysis of results.
⇒ COSs for tendinopathies should map to the established list of nine core health-related domains.
⇒ Lateral elbow tendinopathy (LET) is a common clinical problem that has received considerable research attention.
⇒ There is no agreed COS-LET—limiting meta-analysis.

What are the findings?

⇒ The COS-LET consists of the Patient Rated Tennis Elbow Evaluation—it should be used to capture the disability domain in clinical settings and in all research of this condition.
⇒ The Patient Rated Tennis Elbow Evaluation and its subscales offer insights into the domains of pain and function in addition to disability.
⇒ When measuring participation, physical function capacity and pain on loading, we recommend as interim measures, respectively, time off work, pain-free grip strength and a numerical rating scale for pain on gripping.
⇒ Further work is required to validate many of the measures used in clinical practice.

How might it impact on clinical practice in the future?

⇒ Systematic use of the COS-LET will allow for meta-analysis of research studies and comparison of results between different clinical practices during service evaluation.
⇒ Meta-analysis and network meta-analysis across multiple research trials that apply the COS-LET will increase understanding of treatment effectiveness for people with LET.

Author affiliations
[1]Orthopaedic Department, University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK
[2]Princess Elizabeth Orthopaedic Centre, Royal Devon and Exeter Hospital, Exeter, UK
[3]Health Services and Policy Research Group, University of Exeter Medical School, Exeter, UK
[4]School of Health & Rehabilitation Sciences: Physiotherapy, University of Queensland, Brisbane, Queensland, Australia
[5]Orthopaedic Department, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK
[6]Upper Limb Unit, Wrightington Hospital, Wrightington, UK
[7]Department of Trauma and Orthopaedics, University Hospitals Sussex NHS Foundation Trust, Brighton, UK
[8]Brighton and Sussex Medical School, Brighton, UK
[9]Menzies Health Institute Queensland, Griffith University, Gold Coast, Queensland, Australia
[10]School of Health Sciences & Social Work, Griffith University, Gold Coast, Queensland, Australia

Twitter Marcus Bateman @MarcusBatemanPT, Jonathan P Evans @jon_evans_uk, Viana Vuvan @vianavuvan, Val Jones @Valjones2305val, Adam C Watts @myelbowdoc, Leanne M Bisset @leannebisset and Bill Vicenzino @Bill_Vicenzino

Acknowledgements The steering committee would like to acknowledge the contributions of the core outcome set for lateral elbow tendinopathy participant group, who donated their time voluntarily. Without them, this project would not have been possible.

Collaborators COS-LET consensus group: Isabel Andia (BioCruces Health Research Institute, Spain), Paolo Arrigoni (Università degli Studi di Milano, Italy), Canan Aydin (Metin Sabanci Baltalimani Bone Diseases Training and Research Hospital, Turkey), Paul Barratt (Salford Royal Foundation Trust, UK), Ram Chidambaram (MGM Healthcare, India), Joshua Cleland (School of Medicine, Tufts University, Medford,

Massachusetts, USA), Brooke Coombes (Griffith University, Australia), Uri Farkash (Assuta Ashdod University Hospital, Israel), Taco Gosens (Elisabeth Tweesteden Hospital, the Netherlands), Luke Heales (Central Queensland University, Australia), Martin Kroslak (University of New South Wales, Australia), Chris Littlewood (Manchester Metropolitan University, UK), Joy MacDermid (University of Western Ontario, Canada), George Murrell (University of New South Wales, Australia), Morten Olaussen (University of Oslo, Oslo, Norway), John Orchard (University of Sydney, Australia), Tuomo Pienimaki (Oulu Regional Institute of Occupational Health, Finland), Amar Rangan (James Cook University Hospital, UK), Eirik Solheim (University of Bergen, Norway), Dimitrios Stasinopoulos (University of West Attica, Greece), Jill Thomas (Wrightington, Wigan and Leigh NHS Foundation Trust, UK), Timothy Tyler (Lenox Hill Hospital, New York City, New York, USA), Danielle Van Der Windt (Keele University, UK), Roger Van Riet (Monica Hospital Antwerp, Belgium), Allan Wang (St John of God Subiaco Hospital, Australia), Esther Waugh (University of Toronto, Canada), Jennifer Moriatis Wolf (University of Chicago, Chicago, Illinois, USA), Tony Wright (Curtin University, Australia), Eva Zeisig (Umeå University, Sweden), Ming Hao Zheng (University of Western Australia, Australia), CH (patient representative, UK), Priyesh Mistry (patient representative, UK), RN (patient representative, UK), Ruth Newsome (patient representative, UK), MW (patient representative, UK), Mark Ford (patient representative, Australia), Karen Jackson (patient representative, Australia), Phuong Le (patient representative, Australia) and Joo Yeun Song (patient representative, Australia).

Contributors MB and JPE share joint first authorship and were responsible for study concept, design and the writing of this manuscript. BV was responsible for study concept, design, oversight and the writing of this manuscript. AW, JP, VJ, VV and LMB were responsible for study concept, design and the editing of this manuscript. All authors contributed to the conduct of the study and have approved the final version of the manuscript. MB is guarantor.

Funding This work is not funded and relies on the individuals involved donating their time. Open access publication fees were funded by the University Hospitals of Derby and Burton NHS Foundation Trust Charity.

Competing interests MB and JPE contributed equally and share first authorship. MB, JPE, AW, JP, VJ, VV, LMB and BV declare that they have no competing interests. The core outcome set for lateral elbow tendinopathy (COS-LET) consensus group has no competing interests except for Joy MacDermid, who is the author of the Patient Rated Tennis Elbow Evaluation (PRTEE). Her responses were removed from survey and consensus voting related to the use of the PRTEE in the COS-LET.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by the University of Queensland's research ethics committee (reference number: 2020001340). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

ORCID iDs
Marcus Bateman http://orcid.org/0000-0002-3203-506X
Jonathan P Evans http://orcid.org/0000-0003-0090-1012
Bill Vicenzino http://orcid.org/0000-0003-0253-5933

REFERENCES

1 Walker-Bone K, Palmer KT, Reading I, *et al*. Prevalence and impact of musculoskeletal disorders of the upper limb in the general population. *Arthritis Rheum* 2004;51:642–51.
2 Scott A, Squier K, Alfredson H, *et al*. Icon 2019: international scientific tendinopathy symposium consensus: clinical terminology. *Br J Sports Med* 2020;54:260–2.
3 Evans JP, Smith CD, Fine NF, *et al*. Clinical rating systems in elbow research-a systematic review exploring trends and distributions of use. *J Shoulder Elbow Surg* 2018;27:e98–106.
4 Vicenzino B, de Vos R-J, Alfredson H. Icon 2019—International scientific tendinopathy symposium consensus: there are nine core health-related domains for tendinopathy (core domains): Delphi study of healthcare professionals and patients. *Br J Sports Med* 2019;54.
5 Reeve BB, Wyrwich KW, Wu AW, *et al*. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905.
6 Prinsen CAC, Vohra S, Rose MR, *et al*. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. *Trials* 2016;17:449.
7 Evans JP, Porter I, Gangannagaripalli JB, *et al*. Assessing patient-centred outcomes in lateral elbow tendinopathy: a systematic review and standardised comparison of English language clinical rating systems. *Sports Med Open* 2019;5:10.
8 Boers M, Kirwan J, Tugwell P. *OMERACT Handbook*, 2018.
9 Valderas JM, Ferrer M, Mendívil J, *et al*. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health* 2008;11:700–8.
10 Kirkham JJ, Gorst S, Altman DG, *et al*. Core outcome Set-STAndards for reporting: the COS-STAR statement. *PLoS Med* 2016;13:e1002148.
11 Expertscape Inc. *Expertscape. USA 2000-2020*, 2000.
12 Elsevier BV. Scopus® Expertly curated abstract & citation database, 2020. Available: https://www.elsevier.com/en-in/solutions/scopus
13 Ouzzani M, Hammady H, Fedorowicz Z, *et al*. Rayyan-a web and mobile APP for systematic reviews. *Syst Rev* 2016;5:210.
14 Macdermid J. *The Patient-Rated Tennis Elbow Evaluation (PRTEE)© User Manual*, 2008.
15 Vincent J, MacDermid JC. Patient-rated tennis elbow evaluation questionnaire. *J Physiother* 2014;60:240.
16 Shafiee E, MacDermid JC, Walton D, *et al*. Psychometric properties and cross-cultural adaptation of the Patient-Rated tennis elbow evaluation (PRTEE); a systematic review and meta-analysis. *Disabil Rehabil* 2021;24:1–16.
17 Kennedy C, Beaton D, Solway S. *The DASH and Quick DASH Outcome Measure User's Manual*. 3 edn. Toronto, Ontario: Institute for Work & Health, 2011.
18 Stratford Pet *al*. Assessing disability and change on individual patients: a report of a patient specific measure. *Physiotherapy Canada* 1995;47:258–63.
19 Lowe KA. *The test retest reliability, construct validity, and responsiveness of the tennis elbow function scale*. University of Alberta, 1999.
20 Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996;29:602–8.
21 Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskelet Disord* 2003;4:11.
22 Janssen S, De Smet L. Responsiveness of the DASH questionnaire for surgically treated tennis elbow. *Acta Chir Belg* 2008;108:583–5.
23 Beaton DE, Wright JG, Katz JN, *et al*. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 2005;87:1038–46.
24 Polson K, Reid D, McNair PJ, *et al*. Responsiveness, minimal importance difference and minimal detectable change scores of the shortened disability arm shoulder hand (QuickDASH) questionnaire. *Man Ther* 2010;15:404–7.
25 Dawson J, Doll H, Boller I, *et al*. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br* 2008;90:466–73.
26 Dawson J, Doll H, Boller I, *et al*. Comparative responsiveness and minimal change for the Oxford elbow score following surgery. *Qual Life Res* 2008;17:1257–67.
27 Stratford PW, Levy DR. Assessing valid change over time in patients with lateral epicondylitis at the elbow. *Clinical Journal of Sport Medicine* 1994;4:88–91.
28 Smidt N, van der Windt DA, Assendelft WJ, *et al*. Interobserver reproducibility of the assessment of severity of complaints, grip strength, and pressure pain threshold in patients with lateral epicondylitis. *Arch Phys Med Rehabil* 2002;83:1145–50.
29 Lim ECW. Pain free grip strength test. *J Physiother* 2013;59:59.
30 EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy* 1990;16:199–208.
31 Connelly LB. Economic Evaluations with Pre-Scored Health Status Instruments. In: Oakland T, Mpofu E, eds. *Rehabilitation and health assessment: applying ICF guidelines*. New York: Springer, 2010: 163–88.
32 Kamper S. Global rating of change scales. *Aust J Physiother* 2009;55:289.
33 Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009;17:163–70.
34 Furtado R, MacDermid J. Clinimetrics: single assessment numeric evaluation. *J Physiother* 2019;65:111.
35 Swinkels-Meewisse EJCM, Swinkels RAHM, Verbeek ALM, *et al*. Psychometric properties of the Tampa scale for kinesiophobia and the fear-avoidance beliefs questionnaire in acute low back pain. *Man Ther* 2003;8:29–36.

Delphi Process for Core Outcome Set for Lateral Elbow Tendinopathy (COS-LET): Round 1 Survey Data

Context:

1. There is a high level of heterogeneity in outcome measures used in trials of lateral elbow tendinopathy (LET), which makes evidence synthesis across studies difficult.
2. Previous work in the field of tendinopathy has established through a consensus exercise nine core health-related domains that should be measured in tendinopathy research.
3. The aim of this study is to develop a Core Outcome Set for Lateral Elbow Tendinopathy (COS-LET) mapping to these core domains.

Methods:

The development of the COS-LET is being developed as per the following process:

1. Systematic review of studies investigating LET has revealed a comprehensive list of all instruments that have previously been used to quantify treatment effect or outcome.
2. These instruments were matched to the list of nine core tendinopathy outcome domains by a Steering Committee of clinicians and researchers with a specialist interest in LET resulting in a set of candidate instruments.
3. A 3-stage international consensus process involving experienced clinicians, researchers and patients will be conducted to determine agreement on what should be the COS-LET.
4. We, including you, have completed the 1st stage, which was to respond to a survey.
    - The committee has now collated your responses – reported herein – and then reviewed the psychometric/clinimetric literature to find and rate the available data on the measures you the responders considered should be in a COS-LET.
    - The second stage is where we are at now – the following survey will seek out your responses to a series of questions about including or not measures in a COS-LET.

Results:

The results of the first survey of healthcare professionals and patients are shown herein in Table 1 and 2, and Figures 1-10. **Table 1** shows the characteristics of the healthcare professionals and patients responding to the survey. **Table 2** is a summary of the results with data representing % agreement for each instrument. **Figures 1-10** illustrates the absolute number and % agreement, disagreement and unsure for each instrument.

In summary, there were four instruments that were above the 70% agreement threshold for inclusion in the COS-LET based on responses from healthcare professionals. Of these instruments, three were also above the 70% agreement threshold based on responses from patients:

- VAS pain on gripping (in the Pain on Activity or Loading Domain)
- Patient-Rated Tennis Elbow Evaluation (in the Disability Domain)
- Quick DASH (in the Disability Domain)

There were a number of instruments that were above the 70% disagreement threshold for inclusion in the COS-LET based on responses from both healthcare professionals and patients: 2 in the Pain on Activity or Loading Domain; 9 in the Disability Domain; 1 in the Quality of Life Domain; and 12 that were not mapped to any of the 9 core health-related domains for tendinopathy. These instruments were subsequently excluded.

There were eight instruments that were not previously included that were listed as important or critical for inclusion. These are highlighted in yellow in Table 1.

Table 1: Participant Characteristics (n (%) unless otherwise stated) of those who completed the full survey and provided these details (39 participants commenced, but did not complete).

Characteristics	Healthcare Professionals (N=37)	Patients (N=7)
Sex: Male	25 (67.6)	2 (28.6)
Age: median (IQR; min-max) years	51 (43-57; 34-68)	48 (47.5-54.5; 26-59)
Role:		
Clinician	2 (5.4)	
Researcher	5 (13.5)	
Clinician Researcher	30 (81.1)	
Not a Clinician or Researcher		7 (100)
Highest academic qualification:		
PhD	21 (56.8)	
Master	6 (16.2)	2 (28.6)
Doctor of Medicine	6 (16.2)	
Bachelor	3 (8.1)	3 (42.9)
Undergraduate Diploma/Certificate		1 (14.3)
Not specified	1 (2.7)	
No university qualification		1 (14.3)
Profession:		
Physiotherapist	16 (43.2)	
Orthopaedic surgeon	14 (37.8)	
Sports & Exercise Medicine Physician	3 (8.1)	
Not specified	3 (8.1)	
Rheumatologist	1 (2.7)	
Patient		7 (100%)
Lateral elbow tendinopathy:		
Current history	1 (2.7)	5 (71.4)
Past history	10 (27.0)	4 (57.1)
Country where work:		
Australia	11 (29.7)	2 (28.6)
United Kingdom	10 (27.0)	5 (71.4)
USA	3 (8.1)	
Canada and Norway each:	2 (5.4)	
Belgium, Finland, Greece, Israel, Italy, Netherlands, Spain, Sweden, and Turkey each:	1 (2.7)	

Table 1: Summary of Round 1 Results: data are % responses, with green representing >70% agree, red >70% disagree, and amber neither green or red.

	Health Care Professionals (N = 39^)			Patients (N = 7)		
	In COS-LET?	**Truth (a)**	**Feasibility**	**In COS-LET?**	**Truth (a)**	**Feasibility**
Patient Rating of Condition Domain						
Global Perceived Effect score	56.41	71.79	92.31	71.43	85.71	85.71
Global Rating of Change	64.1	74.36	87.18	57.14	85.71	85.71
Patient Satisfaction Scale	51.28	64.1	89.74	71.43	85.71	85.71
Roles & Maudsley Score	Proposed in Survey 1 comments					
Participation in Live Activities Domain						
Return to sport	38.46	64.1	82.05	71.43	71.42	85.71
Time off work	53.85	79.49	76.92	57.14	57.14	85.71
Total Elbow Scoring System	30.77	35.9	53.85	57.14	57.14	71.43
OSTRC Question 1	Proposed in Survey 1 comments					
Pain on Activity or Loading Domain						
Tennis Elbow Functional Scale*	28.21	61.54	56.41	85.71	85.71	85.71
Thomsen Test	35.9	71.79	71.79	71.43	71.43	85.71
VAS chair pick-up	23.08	51.28	58.97	28.57	85.71	57.14
VAS pain during activity	79.49	89.74	87.18	57.14	71.43	85.71
VAS pain during elbow movement	17.95	33.33	69.23	28.57	42.86	85.71
VAS pain on gripping	71.79	94.87	89.74	71.43	71.43	100
VAS pain at work	38.46	58.97	74.36	42.86	42.86	57.14
Pain-Free Functional Index	17.95	51.28	58.97	42.86	42.86	85.71
Patient-Rated Tennis Elbow Evaluation*	61.54	79.49	71.79	100	100	100
Function Domain						
Patient Specific Functional Scale	28.95	63.16	52.63	71.43	85.71	85.71
Upper Extremity Functional Scale	44.74	68.42	68.42	57.14	57.14	100
VAS function	44.74	63.16	76.32	28.57	42.86	57.14
Patient-Rated Tennis Elbow Evaluation*	Proposed in Survey 1 comments					
Psychological Factors Domain						
Hospital Anxiety and Depression Scale	36.84	52.63	50	28.57	28.57	85.71
Tampa Scale of Kinesophobia	39.47	50	55.26	71.43	85.71	71.43
State-Trait Anxiety Inventory	Proposed in Survey 1 comments					
Nottingham Health Profile	Proposed in Survey 1 comments					
Physical Function Capacity Domain						
Grip strength (maximum)*	47.37	81.58	68.42	85.71	85.71	100
Pain free grip strength*	65.79	78.95	71.05	85.71	85.71	100
Elbow ROM	10.53	31.58	65.79	42.86	28.57	100

Disability Domain						
Andrews-Carson Score	0	7.89	26.32	14.29	14.29	42.86
American Shoulder & Elbow Score	10.53	44.74	34.21	42.86	42.86	71.43
Broberg & Morrey Rating System	2.63	21.05	26.32	28.57	42.86	71.43
Disabilities of the Arm Shoulder and Hand*	39.47	68.42	44.74	57.14	57.14	71.43
HAND10	7.89	44.74	60.53	57.14	57.14	85.71
Japanese Orthopaedic Association Elbow Score	10.53	44.74	42.11	0	0	57.14
Laitinen Questionnaire	13.16	36.84	39.47	14.29	14.29	57.14
Liverpool Elbow Score	2.63	34.21	34.21	28.57	42.86	42.86
Mayo Elbow Performance Score	7.89	31.58	39.47	0	0	28.57
Nirschl Tennis Elbow Score	18.42	57.89	55.26	57.14	57.14	57.14
Nottingham Health Profile	2.63	15.79	31.58	0	0	0
Oxford Elbow Score*	23.68	52.63	55.26	57.14	71.43	85.71
Patient-Rated Wrist Evaluation Questionnaire	13.16	44.74	50.00	57.14	57.14	85.71
Patient-Rated Tennis Elbow Evaluation*	73.68	92.11	86.84	85.71	85.71	85.71
Quick DASH*	71.05	76.32	81.58	100	100	100
Total Elbow Scoring System	10.53	36.84	52.63	28.57	28.57	57.14
Roles & Maudsley Score	5.26	28.95	57.89	14.29	28.57	57.14
Quality of Life Domain						
EuroQoL (EQ5D)	51.35	59.46	67.57	57.14	57.14	85.71
Short Form Survey (SF-36)	16.22	62.16	24.32	14.29	71.43	28.57
SF-12 Health Survey (SF-12)	40.54	67.57	67.57	42.86	71.43	57.14
Nottingham Health Profile	Proposed in Survey 1 comments					
World Health Organization Quality of Life Instruments (WHOQOL-BREF)	Proposed in Survey 1 comments					
Short Musculoskeletal Function Assessment	Proposed in Survey 1 comments					
Pain Over a Specified Timeframe Domain						
VAS night pain	37.84	70.27	86.49	28.57	57.14	85.71
VAS pain defined time period	67.57	81.08	78.38	28.57	42.86	71.43
VAS pain at rest	56.76	67.57	83.78	14.29	28.57	71.43
Tennis Elbow Functional Scale*	21.62	59.46	56.76	85.71	85.71	85.71
Others						
Analgesic use	43.24			14.29		
Canadian Occupational Performance Measure	5.41			14.29		
Cold Pain Threshold	8.11			14.29		
Electromyography	0			28.57		
Gothenburg Quality of Life Instrument	0			0		
MRI evaluation	8.11			14.29		

Orthopaedic Research Institute-Tennis Elbow Testing System	2.7				14.29		
Placzek Score	10.81				28.57		
Pressure Pain Threshold	18.92				14.29		
Ultrasound Appearance	10.81				14.29		
University of Pelopponnese Pain, Functionality and Quality of Life Questionnaire	2.7				14.29		
VAS pain on palpation	10.81				28.57		
VAS pain overall	24.32				28.57		
Work Limitations Questionnaire	5.41				57.14		

^ 39 responders completed the first 3 domains, 38 continued on to the next 4 domains and 37 completed it all

**indicates the measures that had clinimetric data/study(ies) with a score of at least 40%.*

Figure 1: Participant response for **Patient Rating of Condition** Domain

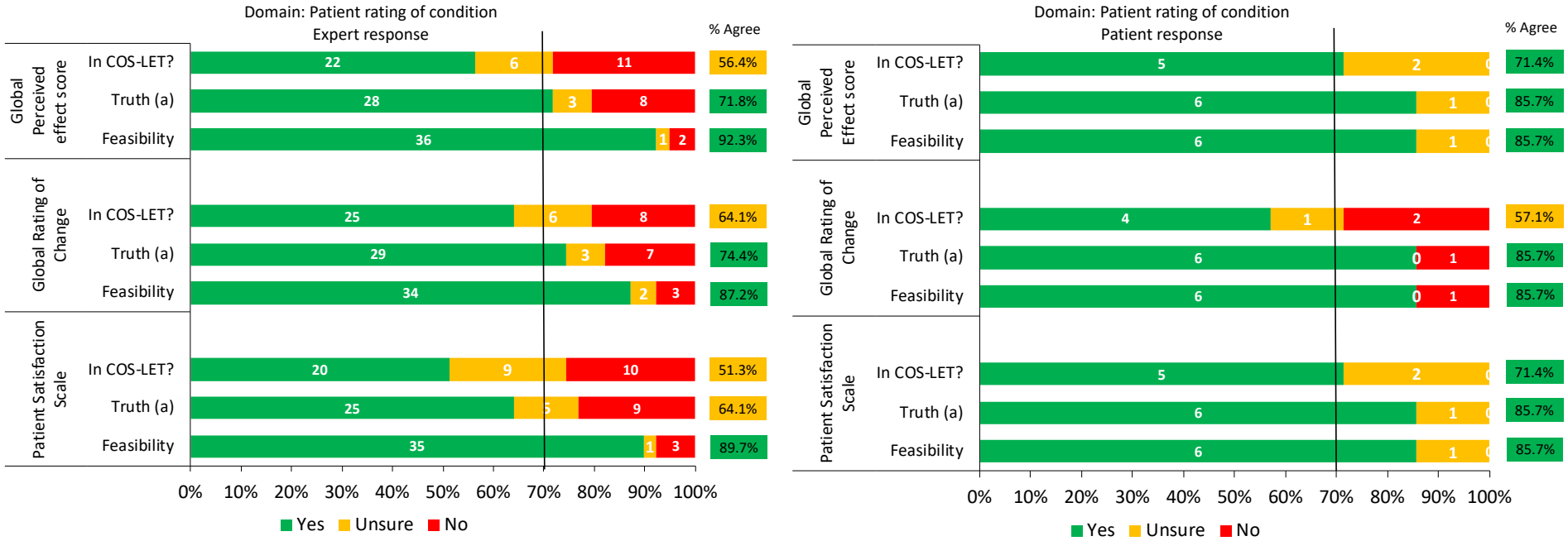
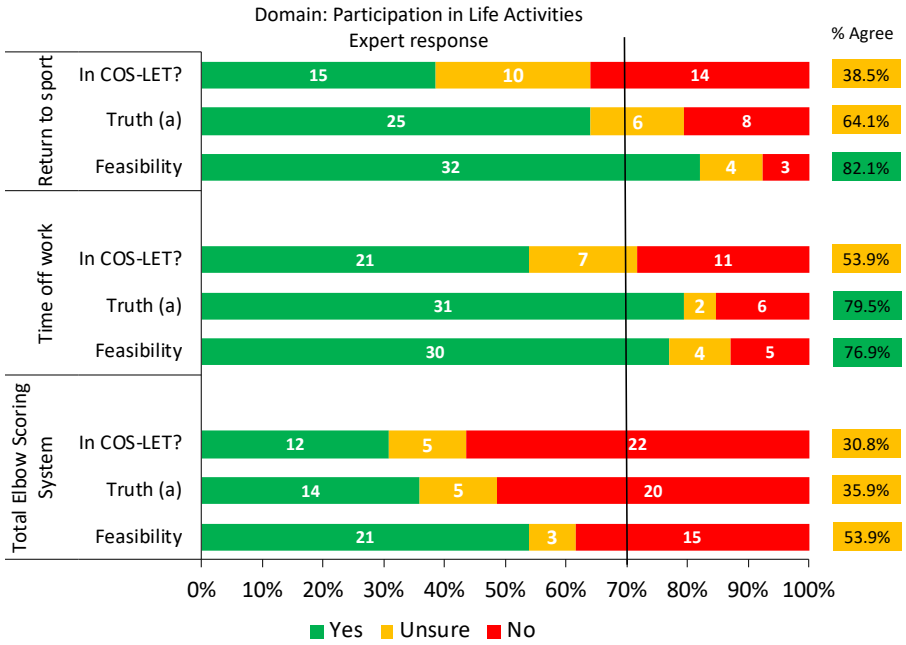


Figure 2: Participant response for **Participation in Life Activities** Domain



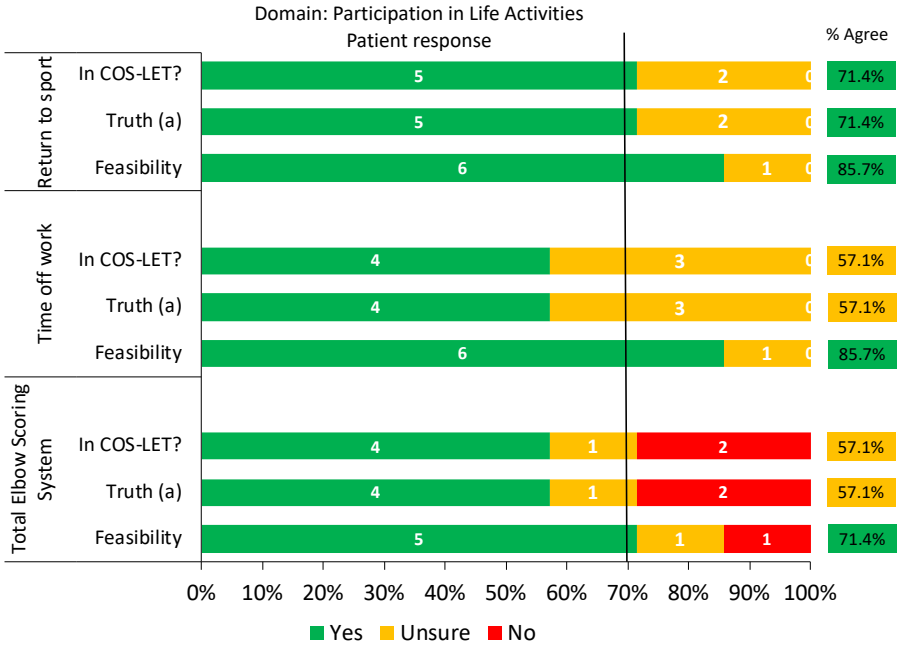
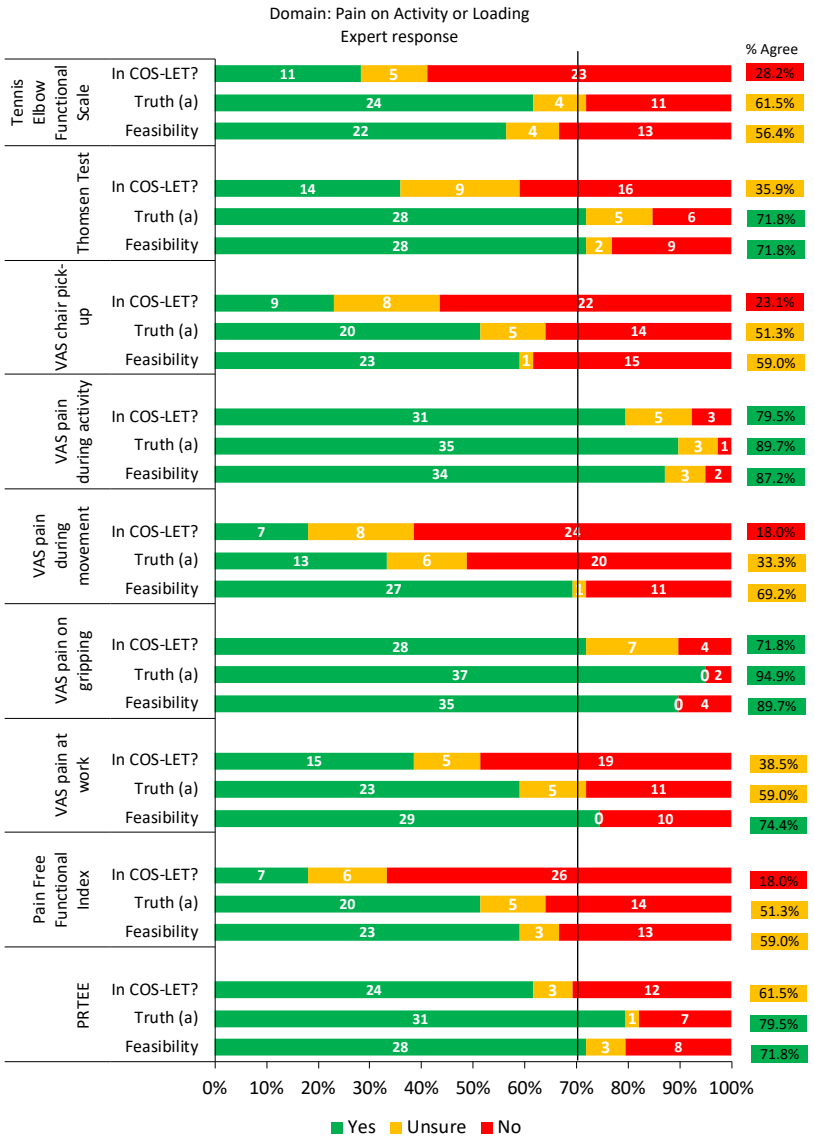


Figure 3: Participant response for **Pain on Activity or Loading** Domain



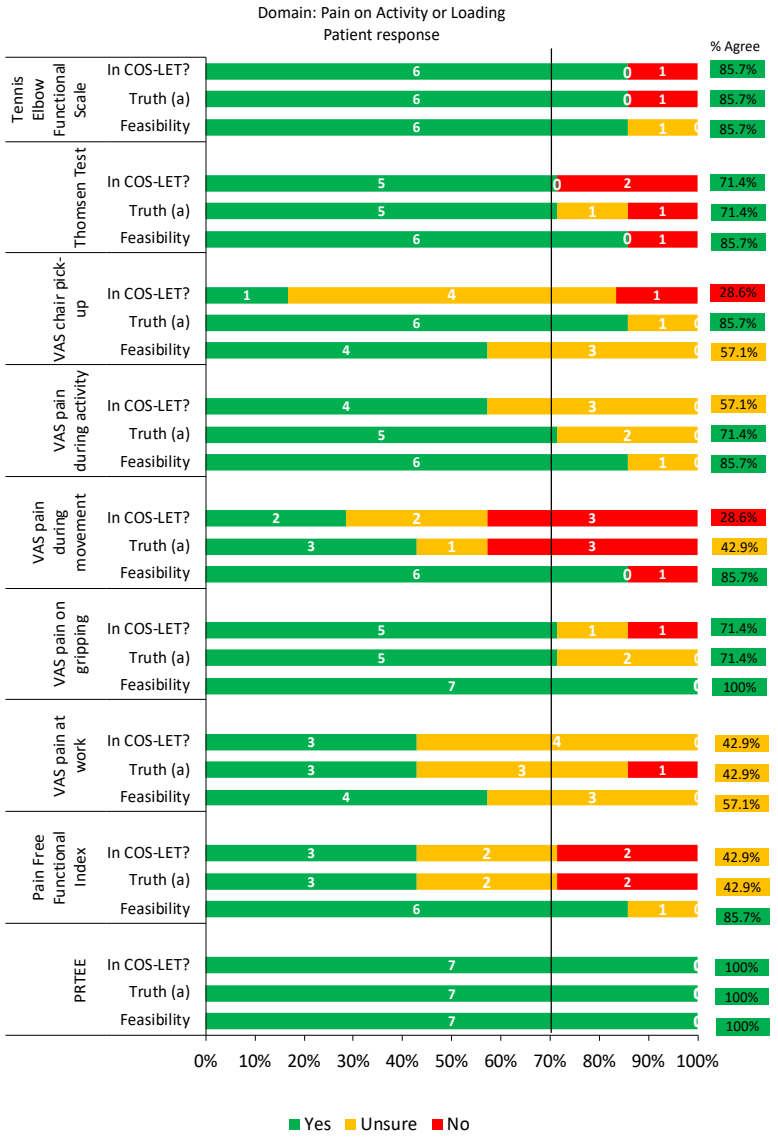


Figure 4: Participant response for **Function** Domain

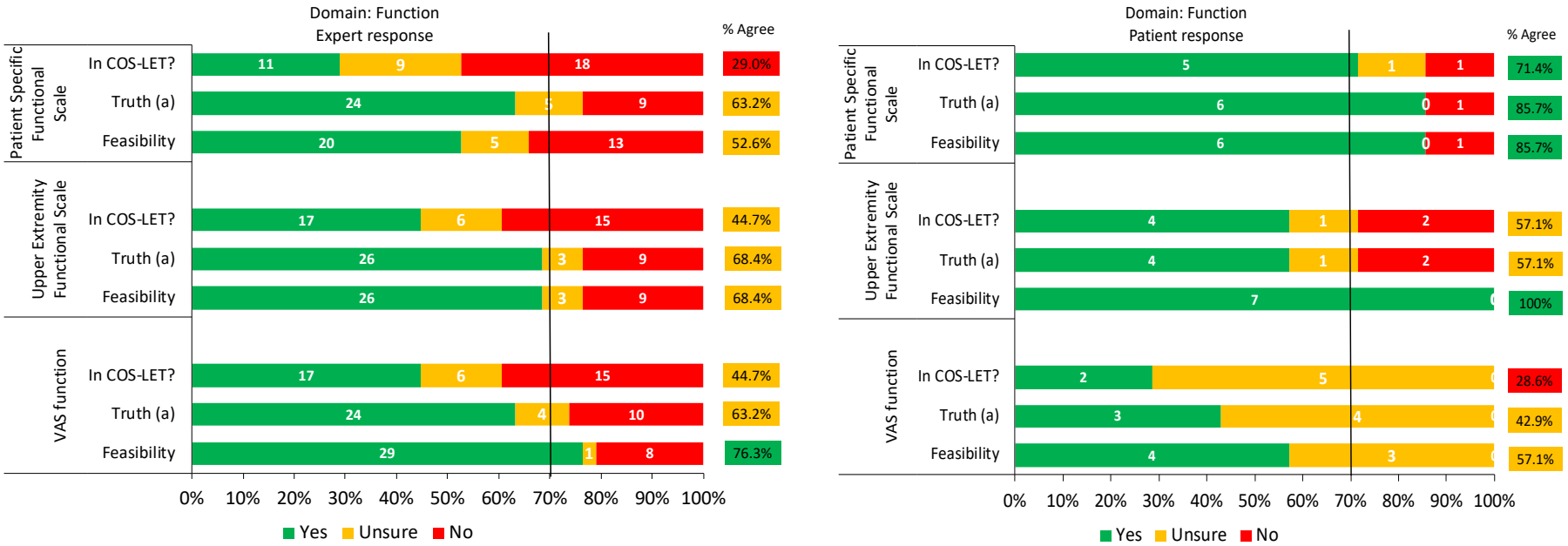
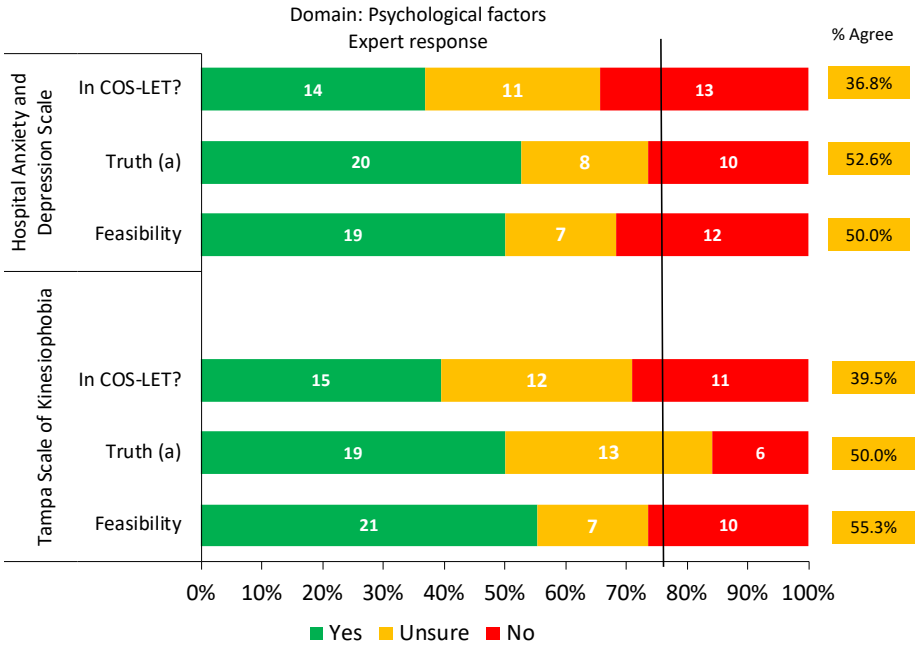


Figure 5: Participant response for **Psychological Factors** Domain



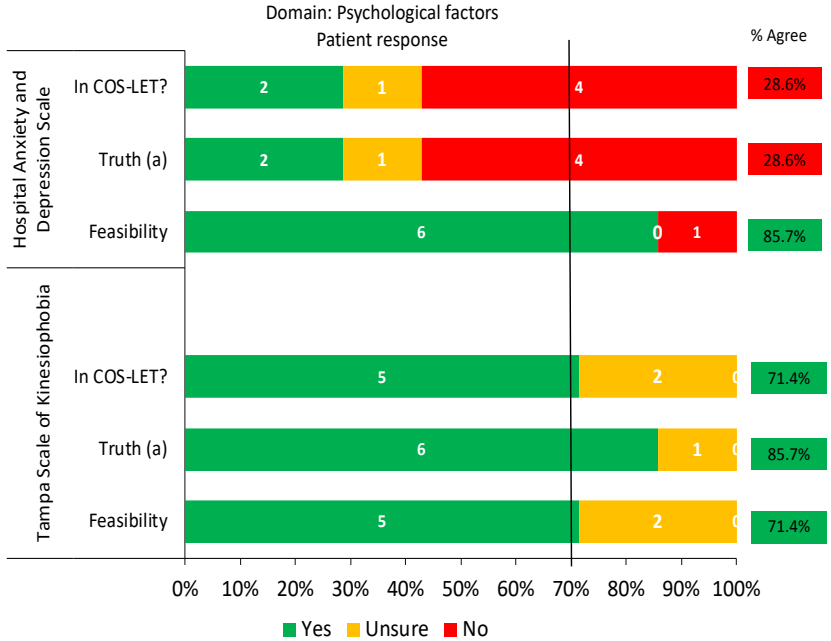
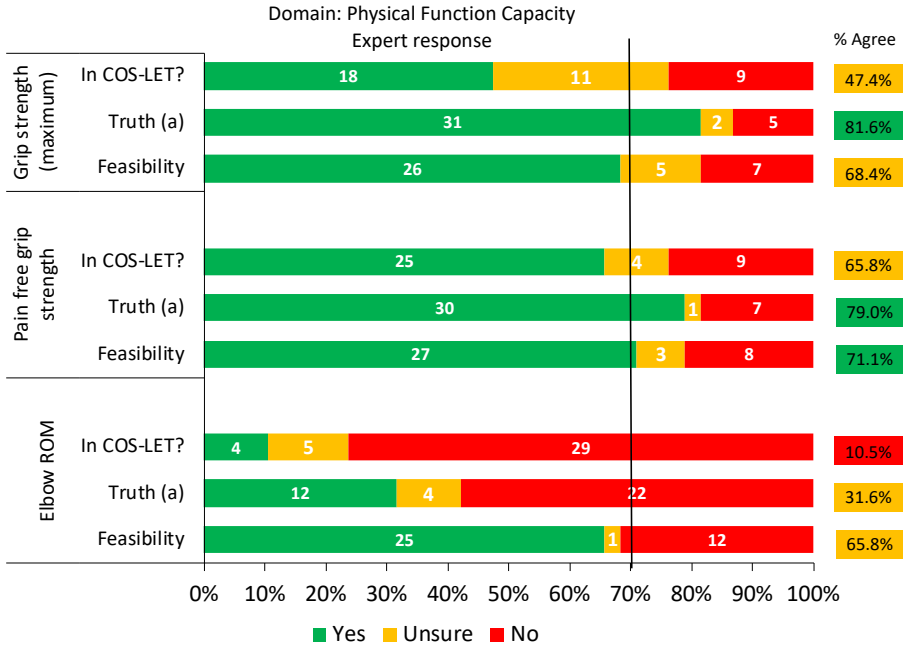


Figure 6: Participant response for **Physical Function Capacity** Domain



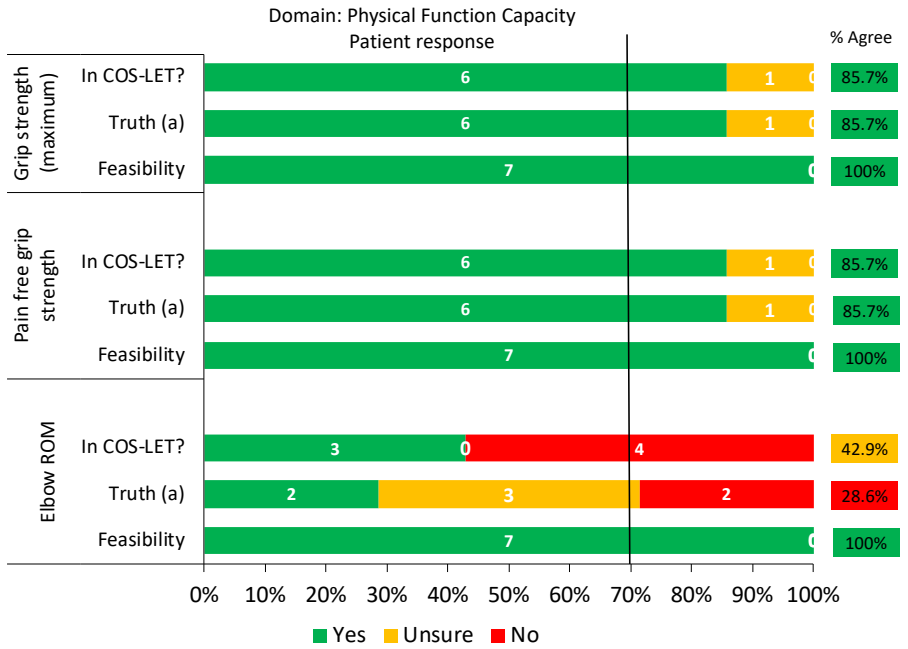
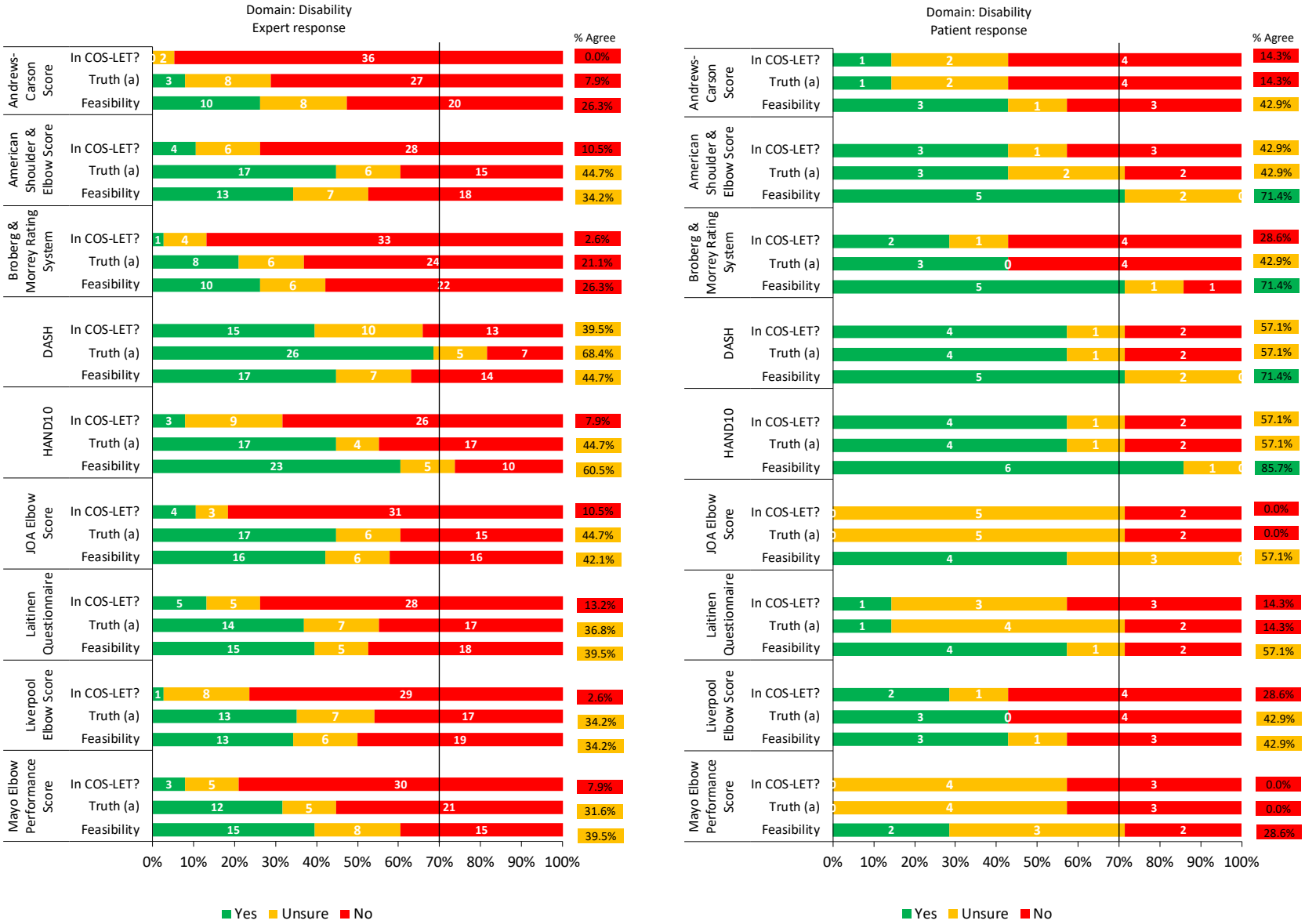
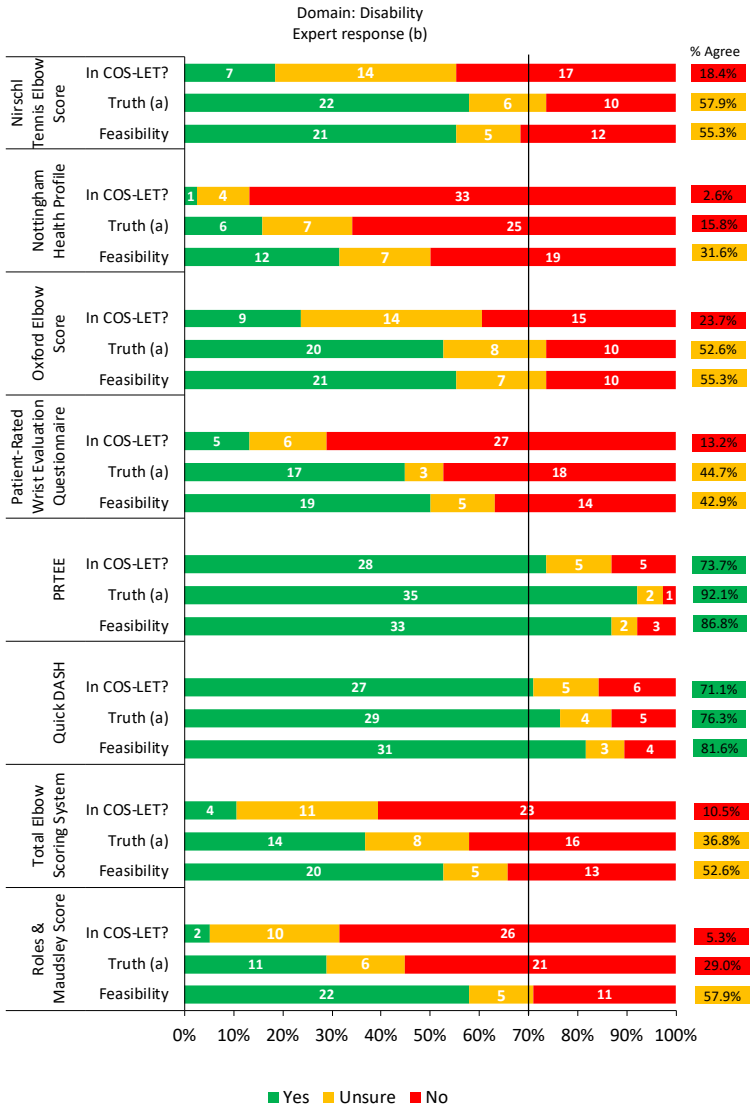


Figure 7: Participant response for **Disability** Domain




Domain: Disability
Expert response (b)
% Agree
Nirschl Tennis Elbow Score
In COS-LET? 7 14 17 18.4%
Truth (a) 22 6 10 57.9%
Feasibility 21 5 12 55.3%
Nottingham Health Profile
In COS-LET? 1 4 33 2.6%
Truth (a) 6 7 25 15.8%
Feasibility 12 7 19 31.6%
Oxford Elbow Score
In COS-LET? 9 14 15 23.7%
Truth (a) 20 8 10 52.6%
Feasibility 21 7 10 55.3%
Patient-Rated Wrist Evaluation Questionnaire
In COS-LET? 5 6 27 13.2%
Truth (a) 17 3 18 44.7%
Feasibility 19 5 14 42.9%
PRTEE
In COS-LET? 28 5 5 73.7%
Truth (a) 35 2 1 92.1%
Feasibility 33 2 3 86.8%
Quick DASH
In COS-LET? 27 5 6 71.1%
Truth (a) 29 4 5 76.3%
Feasibility 31 3 4 81.6%
Total Elbow Scoring System
In COS-LET? 4 11 23 10.5%
Truth (a) 14 8 16 36.8%
Feasibility 20 5 13 52.6%
Roles & Maudsley Score
In COS-LET? 2 10 26 5.3%
Truth (a) 11 6 21 29.0%
Feasibility 22 5 11 57.9%
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Yes Unsure No

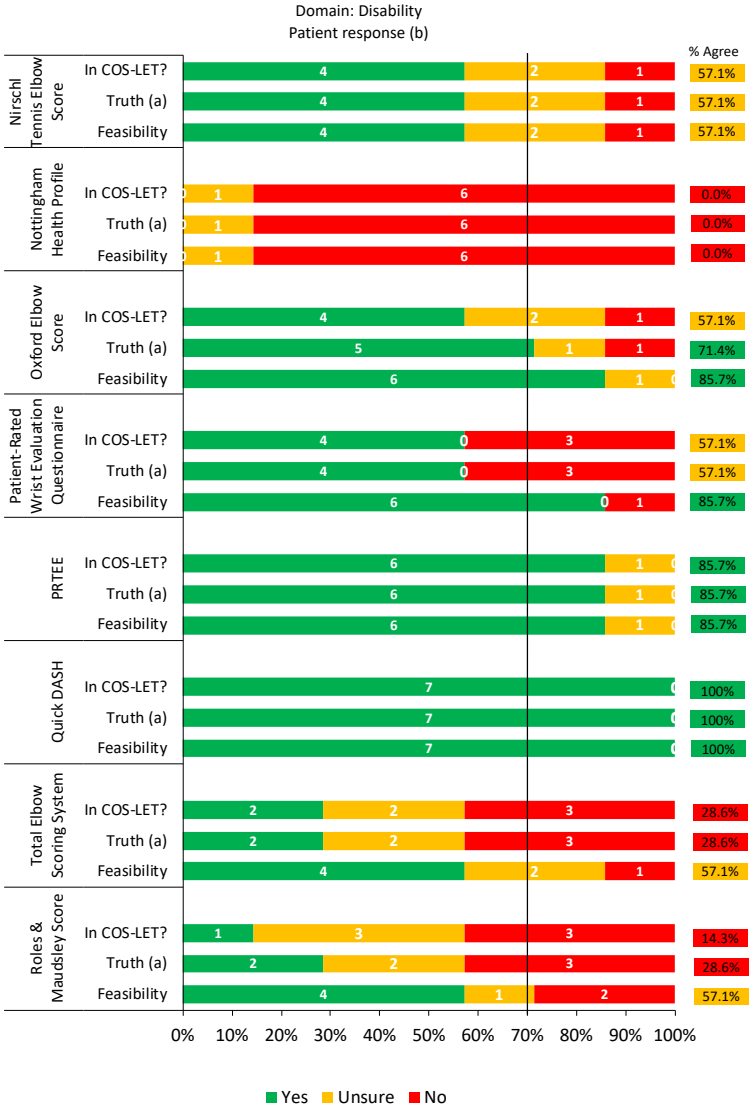
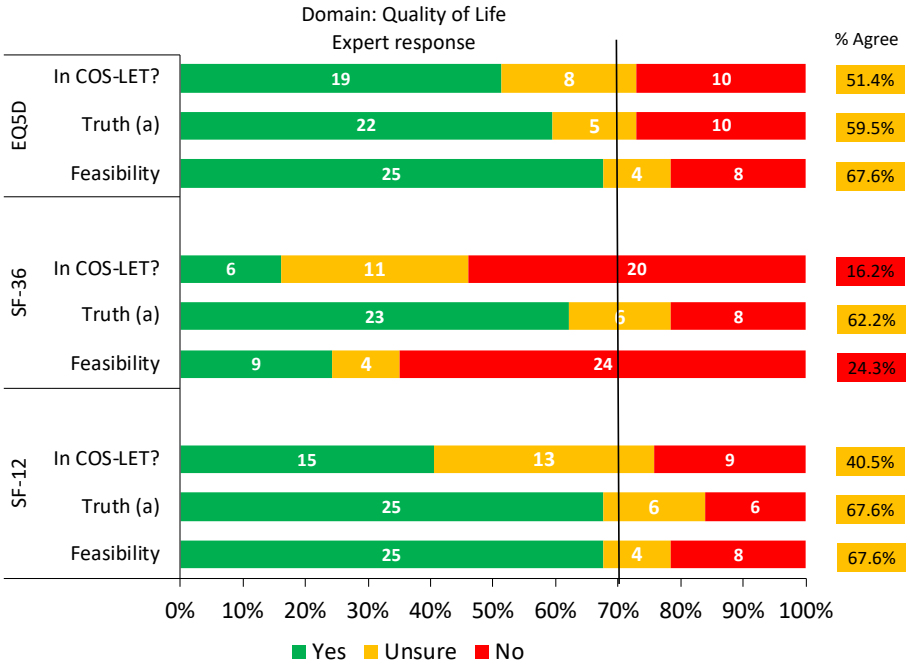

Domain: Disability
Patient response (b)
% Agree
Nirschl Tennis Elbow Score
In COS-LET? 4 2 1 57.1%
Truth (a) 4 2 1 57.1%
Feasibility 4 2 1 57.1%
Nottingham Health Profile
In COS-LET? 1 6 0.0%
Truth (a) 1 6 0.0%
Feasibility 1 6 0.0%
Oxford Elbow Score
In COS-LET? 4 2 1 57.1%
Truth (a) 5 1 1 71.4%
Feasibility 6 1 0 85.7%
Patient-Rated Wrist Evaluation Questionnaire
In COS-LET? 4 0 3 57.1%
Truth (a) 4 0 3 57.1%
Feasibility 6 0 1 85.7%
PRTEE
In COS-LET? 6 1 0 85.7%
Truth (a) 6 1 0 85.7%
Feasibility 6 1 0 85.7%
Quick DASH
In COS-LET? 7 0 100%
Truth (a) 7 0 100%
Feasibility 7 0 100%
Total Elbow Scoring System
In COS-LET? 2 2 3 28.6%
Truth (a) 2 2 3 28.6%
Feasibility 4 2 1 57.1%
Roles & Maudsley Score
In COS-LET? 1 3 3 14.3%
Truth (a) 2 2 3 28.6%
Feasibility 4 1 2 57.1%
0% 10% 20% 30% 40% 50% 60% 70% 80% 90% 100%
Yes Unsure No

Figure 8: Participant response for **Quality of Life** Domain



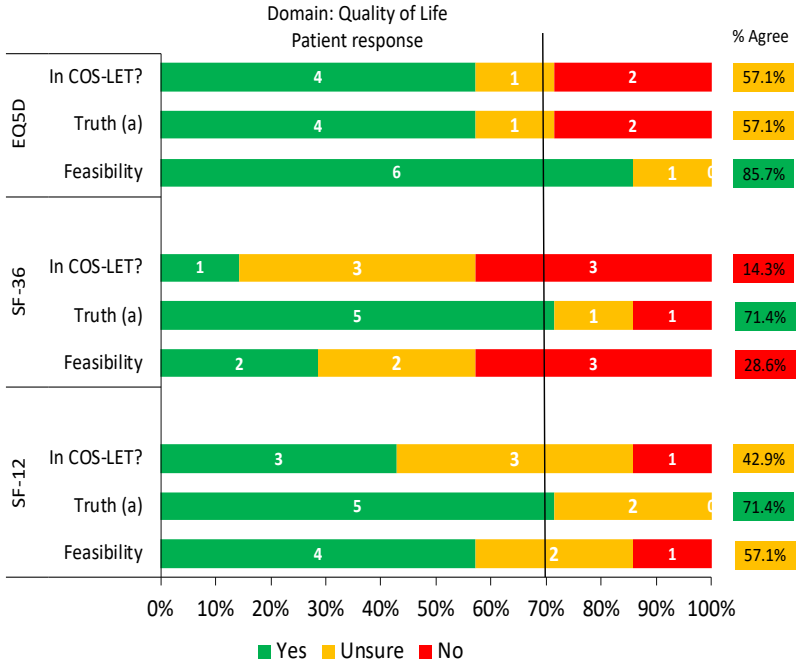
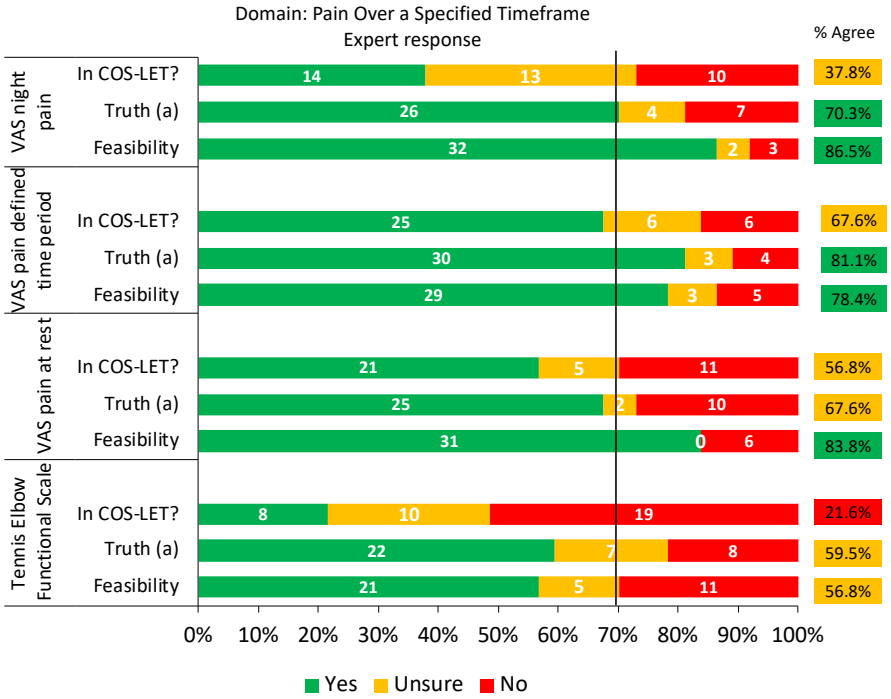


Figure 9: Participant response for **Pain Over a Specified Time** Domain



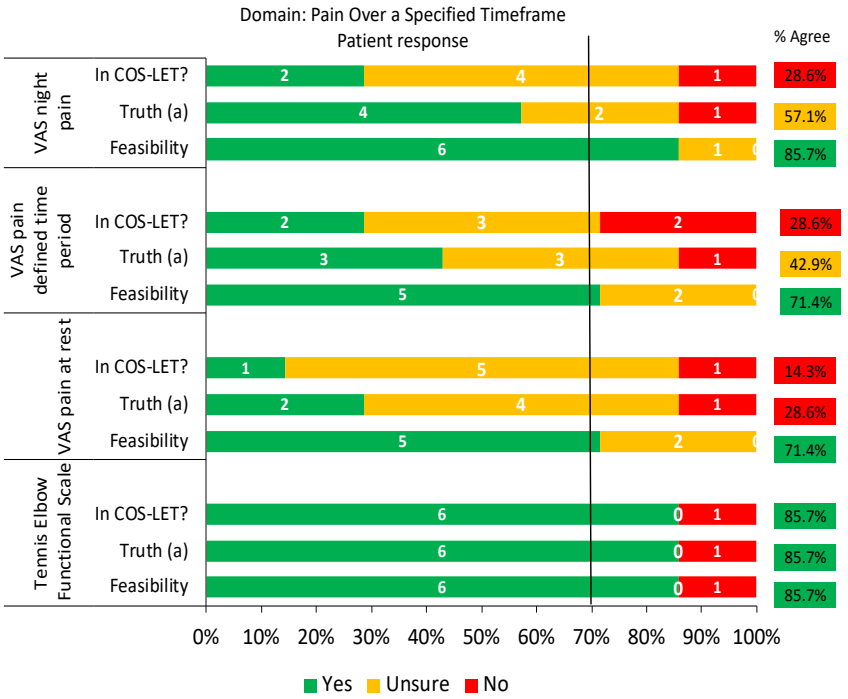
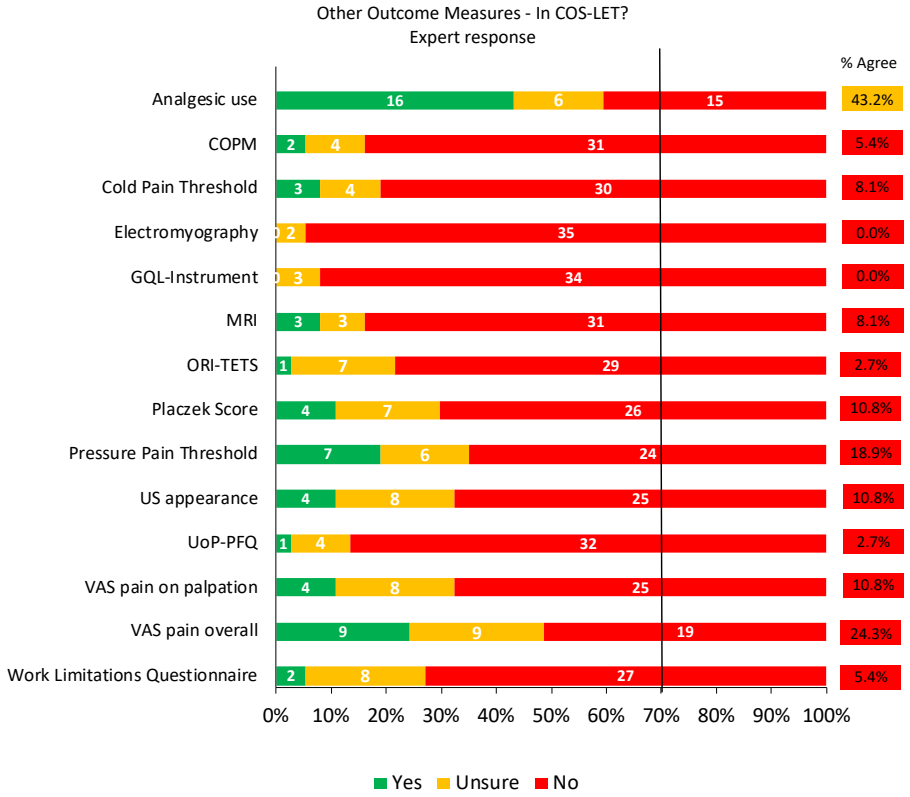
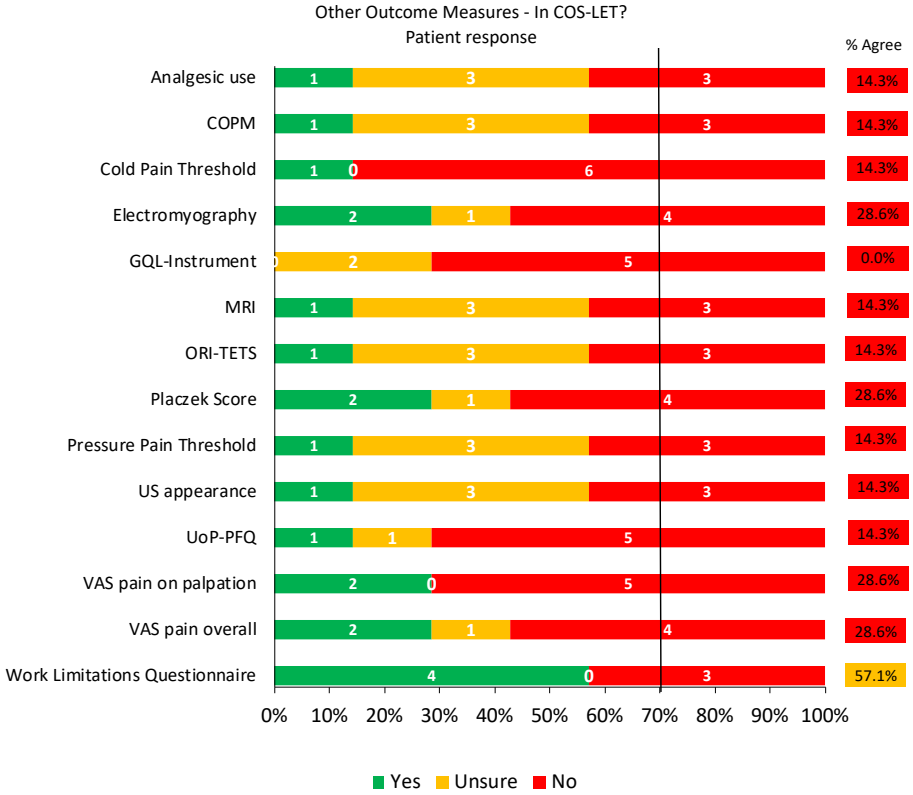


Figure 10: Other outcome measures





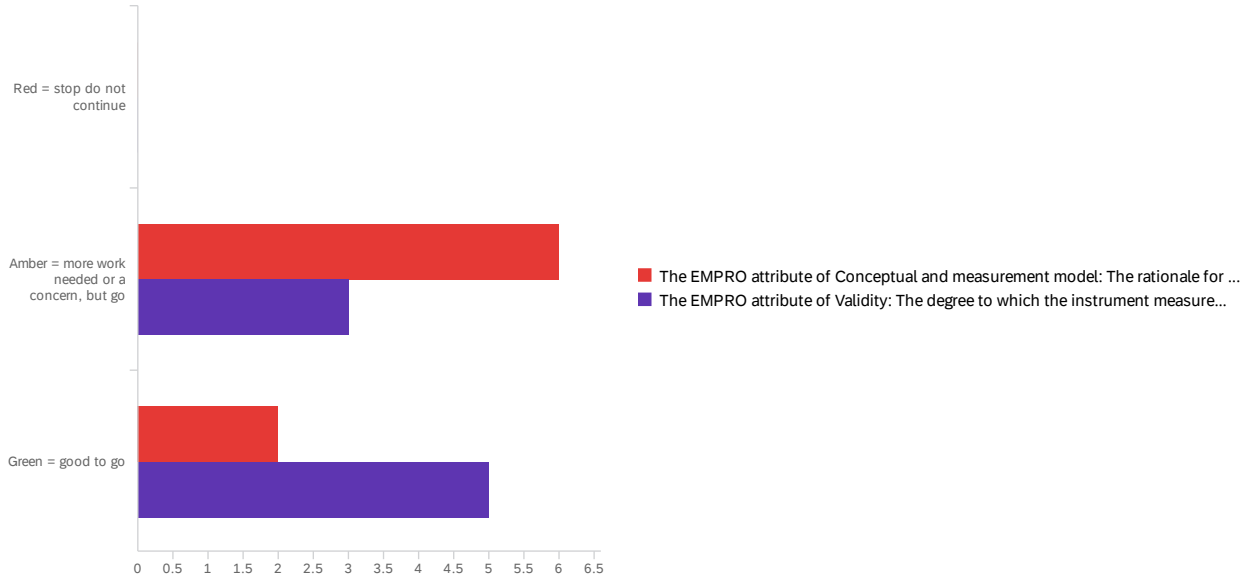
Default Report

COS-LET_Steering_Committee_OMERACT-filter2
January 10, 2021 5:43 PM MST

Q49 - Please leave your details here in case we need to follow up.

Last name	First name
val	jones
Phadnis	Joideep
Watts	Adam
Evans	Jon
Bisset	Leanne
Vicenzino	Bill
Vuvan	Viana
Bateman	Marcus

truth - The PRTEE meets the truth OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

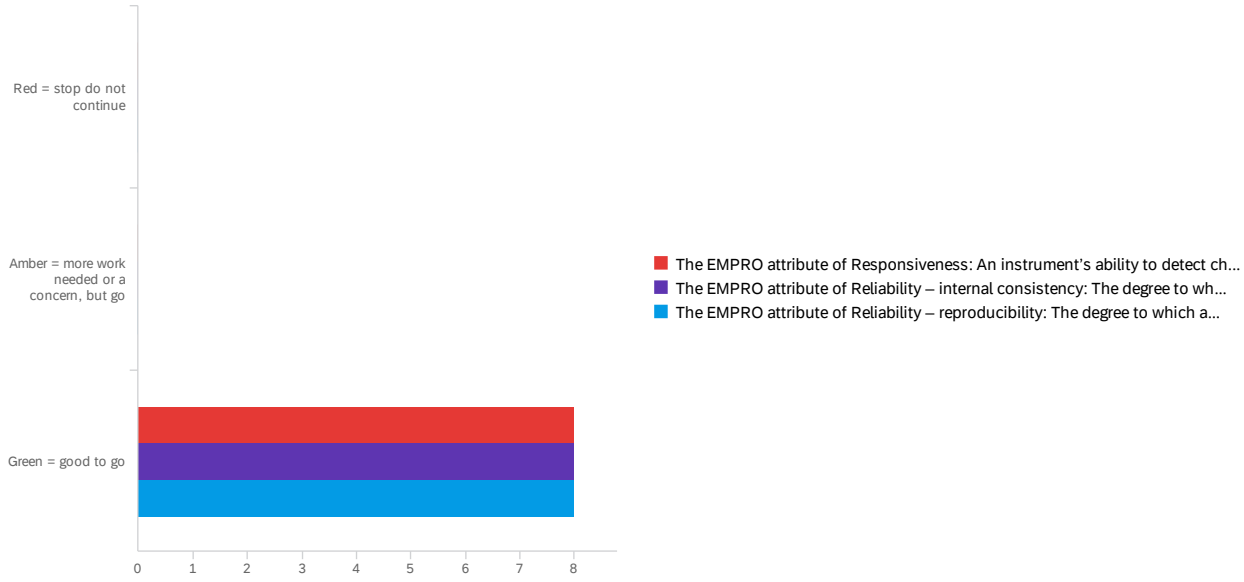


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	2.00	3.00	2.25	0.43	0.19	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%	2.00	3.00	2.63	0.48	0.23	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	0.00% 0	75.00% 6	25.00% 2	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%	0.00% 0	37.50% 3	62.50% 5	8

Showing rows 1 - 2 of 2

discrim - PRTEE meets the discrimination OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.



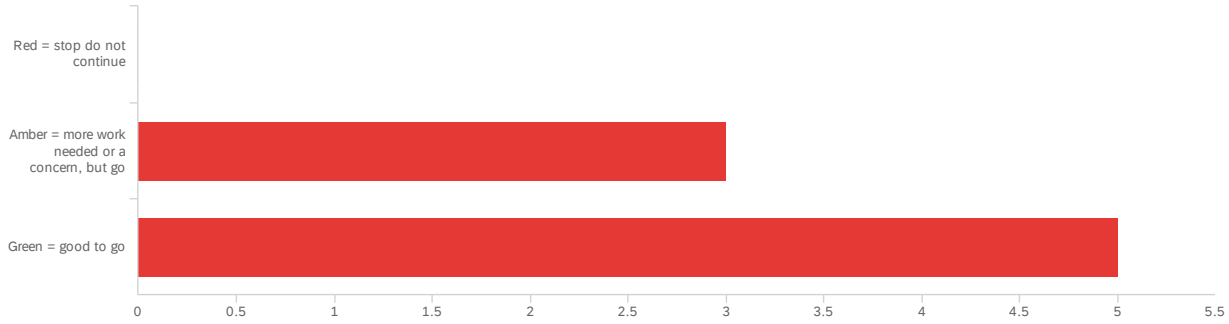
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	3.00	3.00	3.00	0.00	0.00	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	3.00	3.00	3.00	0.00	0.00	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	3.00	3.00	3.00	0.00	0.00	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	0.00% 0	0.00% 0	100.00% 8	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	0.00% 0	0.00% 0	100.00% 8	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	0.00% 0	0.00% 0	100.00% 8	8

Showing rows 1 - 3 of 3

overall - Considering your responses to the truth and discrimination OMERACT filter above, does PRTEE meets the requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

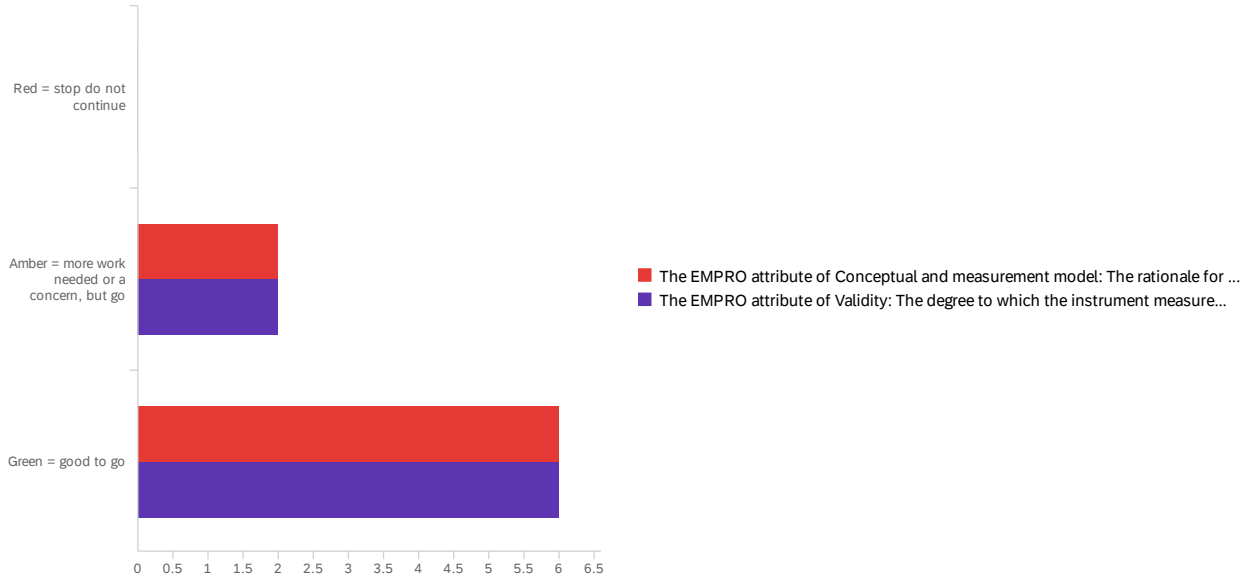


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 57.02%	2.00	3.00	2.63	0.48	0.23	8

#	Field		Choice Count
1	Red = stop do not continue	0.00%	0
2	Amber = more work needed or a concern, but go	37.50%	3
3	Green = good to go	62.50%	5
			8

Showing rows 1 - 4 of 4

Q12 - The qDASH meets the truth OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

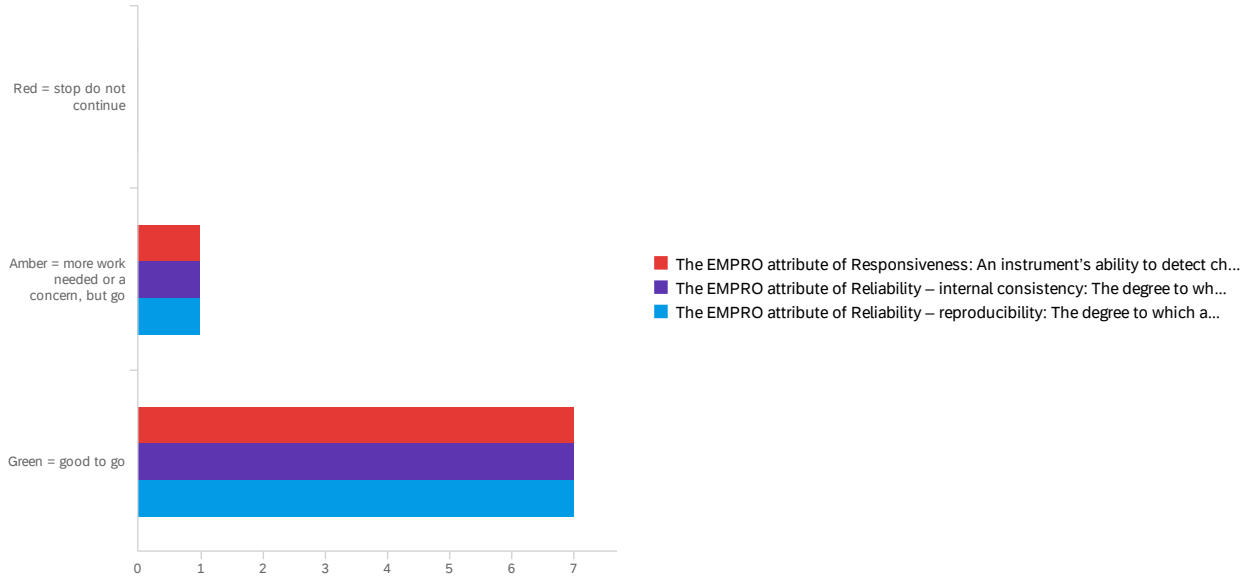


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 76.19%	2.00	3.00	2.75	0.43	0.19	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 72.22%	2.00	3.00	2.75	0.43	0.19	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 76.19%	0.00% 0	25.00% 2	75.00% 6	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 72.22%	0.00% 0	25.00% 2	75.00% 6	8

Showing rows 1 - 2 of 2

Q13 - qDASH meets the discrimination OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.



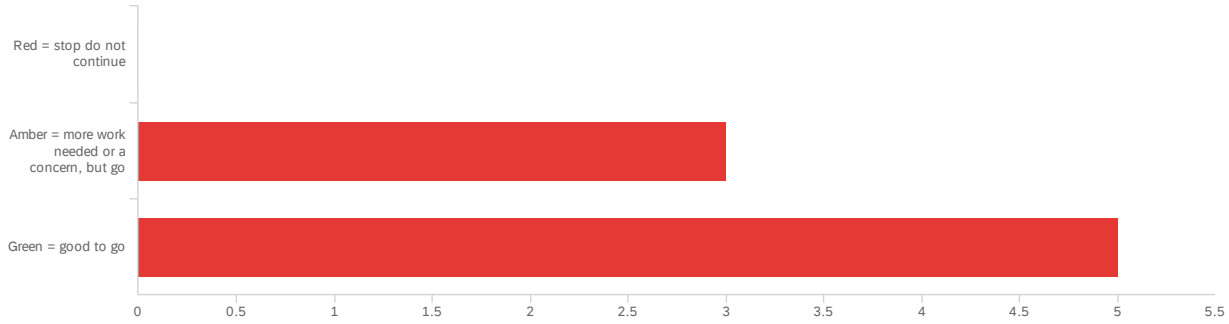
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	2.00	3.00	2.88	0.33	0.11	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	2.00	3.00	2.88	0.33	0.11	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	2.00	3.00	2.88	0.33	0.11	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	0.00% 0	12.50% 1	87.50% 7	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	0.00% 0	12.50% 1	87.50% 7	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	0.00% 0	12.50% 1	87.50% 7	8

Showing rows 1 - 3 of 3

Q14 - Considering your responses to the truth and discrimination OMERACT filter above, does qDASH meets the requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

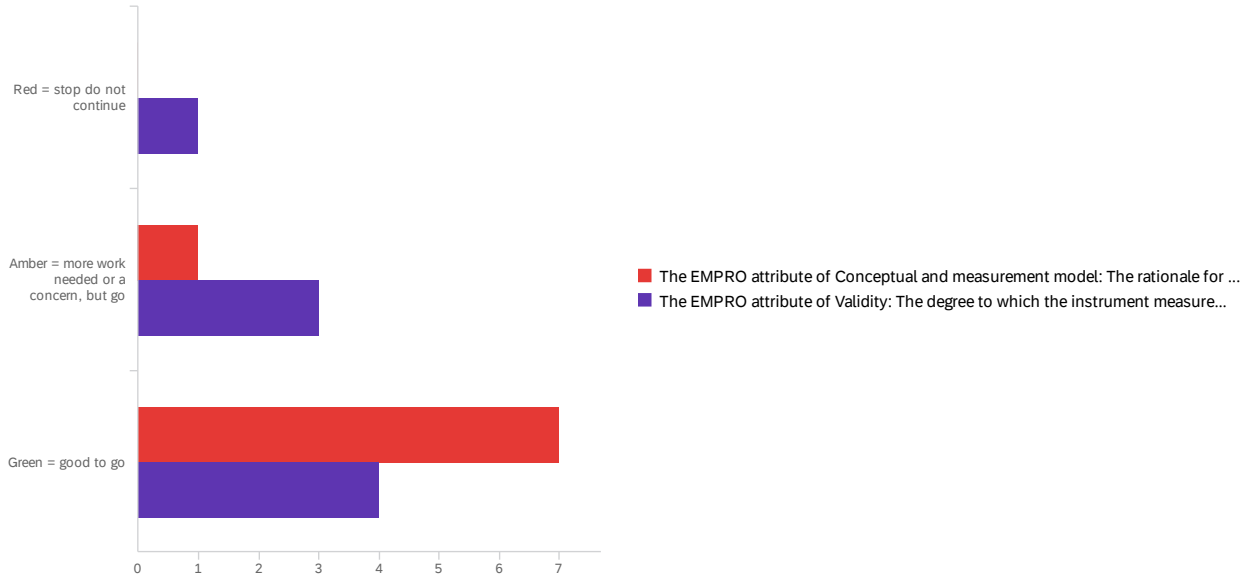


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 57.02%	2.00	3.00	2.63	0.48	0.23	8

#	Field		Choice Count
1	Red = stop do not continue	0.00%	0
2	Amber = more work needed or a concern, but go	37.50%	3
3	Green = good to go	62.50%	5
			8

Showing rows 1 - 4 of 4

Q16 - The DASH meets the truth OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

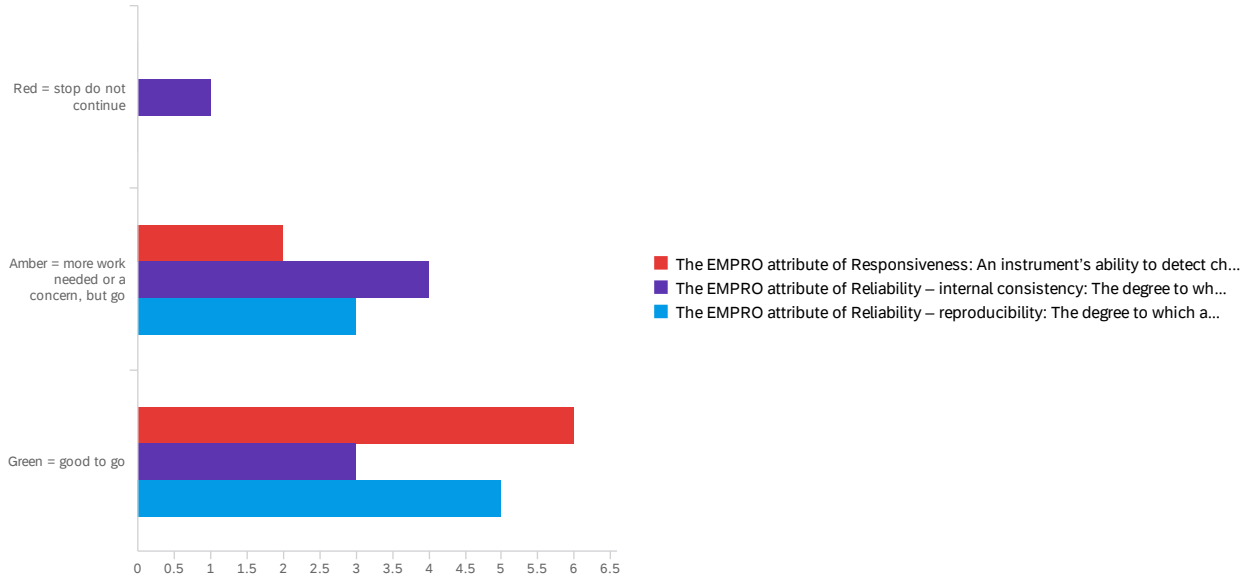


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 76.19%	2.00	3.00	2.88	0.33	0.11	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 50%	1.00	3.00	2.38	0.70	0.48	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 76.19%	0.00% 0	12.50% 1	87.50% 7	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 50%	12.50% 1	37.50% 3	50.00% 4	8

Showing rows 1 - 2 of 2

Q17 - DASH meets the discrimination OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.



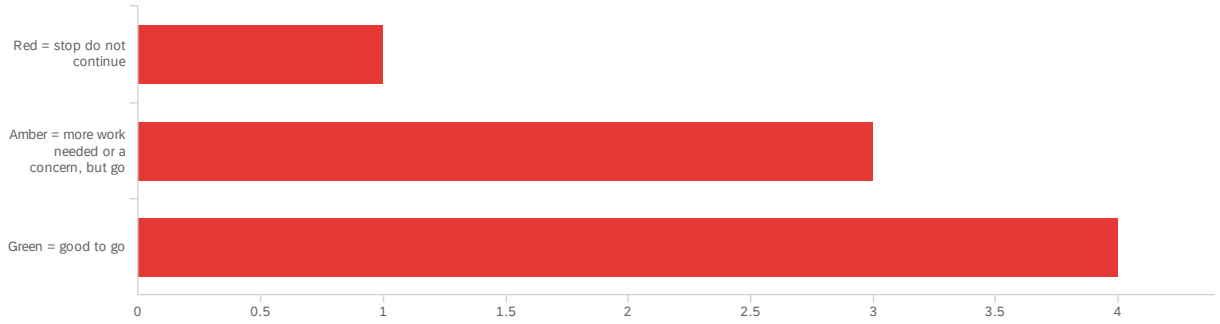
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 66.67%	2.00	3.00	2.75	0.43	0.19	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 50%	1.00	3.00	2.25	0.66	0.44	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	2.00	3.00	2.63	0.48	0.23	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 66.67%	0.00% 0	25.00% 2	75.00% 6	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 50%	12.50% 1	50.00% 4	37.50% 3	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	0.00% 0	37.50% 3	62.50% 5	8

Showing rows 1 - 3 of 3

Q18 - Considering your responses to the truth and discrimination OMERACT filter above, does DASH meets the requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

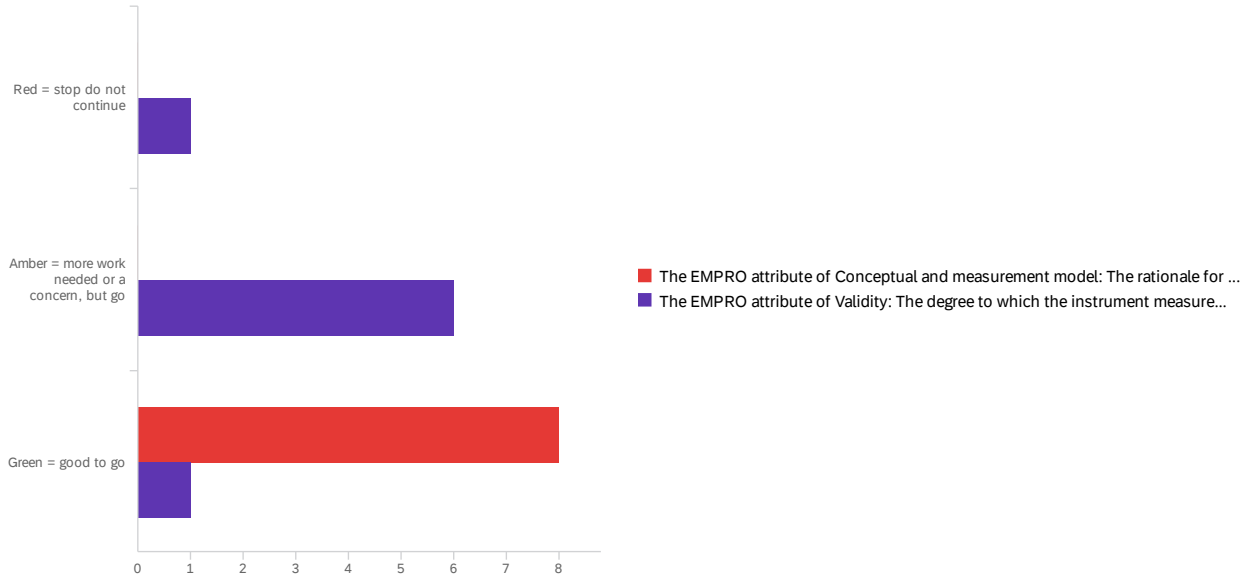


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 66.9%	1.00	3.00	2.38	0.70	0.48	8

#	Field		Choice Count
1	Red = stop do not continue	12.50%	1
2	Amber = more work needed or a concern, but go	37.50%	3
3	Green = good to go	50.00%	4
			8

Showing rows 1 - 4 of 4

Q20 - The OES meets the truth OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

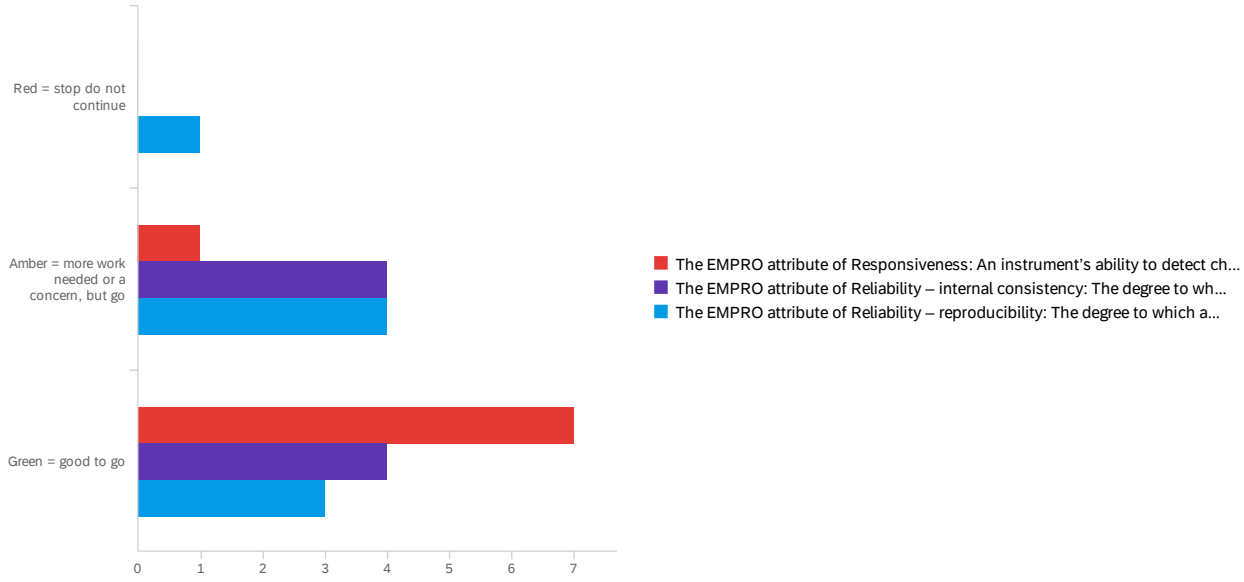


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 85.7%	3.00	3.00	3.00	0.00	0.00	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 44.4%	1.00	3.00	2.00	0.50	0.25	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 85.7%	0.00% 0	0.00% 0	100.00% 8	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 44.4%	12.50% 1	75.00% 6	12.50% 1	8

Showing rows 1 - 2 of 2

Q21 - OES meets the discrimination OMERACT filter requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.



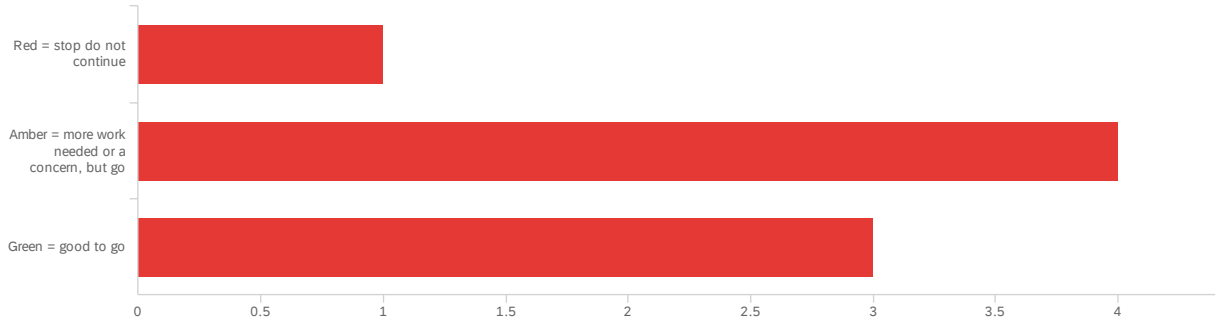
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.8%	2.00	3.00	2.88	0.33	0.11	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 58.3%	2.00	3.00	2.50	0.50	0.25	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 50%	1.00	3.00	2.25	0.66	0.44	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.8%	0.00% 0	12.50% 1	87.50% 7	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 58.3%	0.00% 0	50.00% 4	50.00% 4	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 50%	12.50% 1	50.00% 4	37.50% 3	8

Showing rows 1 - 3 of 3

Q22 - Considering your responses to the truth and discrimination OMERACT filter above, does OES meets the requirements for the outcome measure representing the Disability Domain in the core outcome set for lateral elbow tendinopathy.

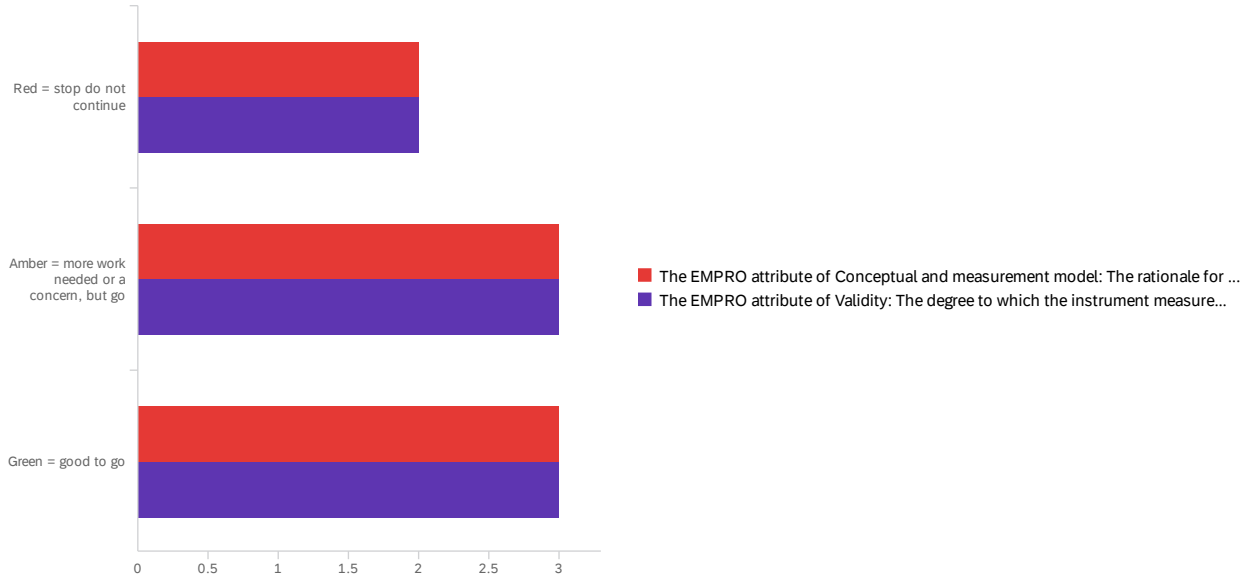


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 66.59%	1.00	3.00	2.25	0.66	0.44	8

#	Field		Choice Count
1	Red = stop do not continue	12.50%	1
2	Amber = more work needed or a concern, but go	50.00%	4
3	Green = good to go	37.50%	3
			8

Showing rows 1 - 4 of 4

Q24 - The PRTEE meets the truth OMERACT filter requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.

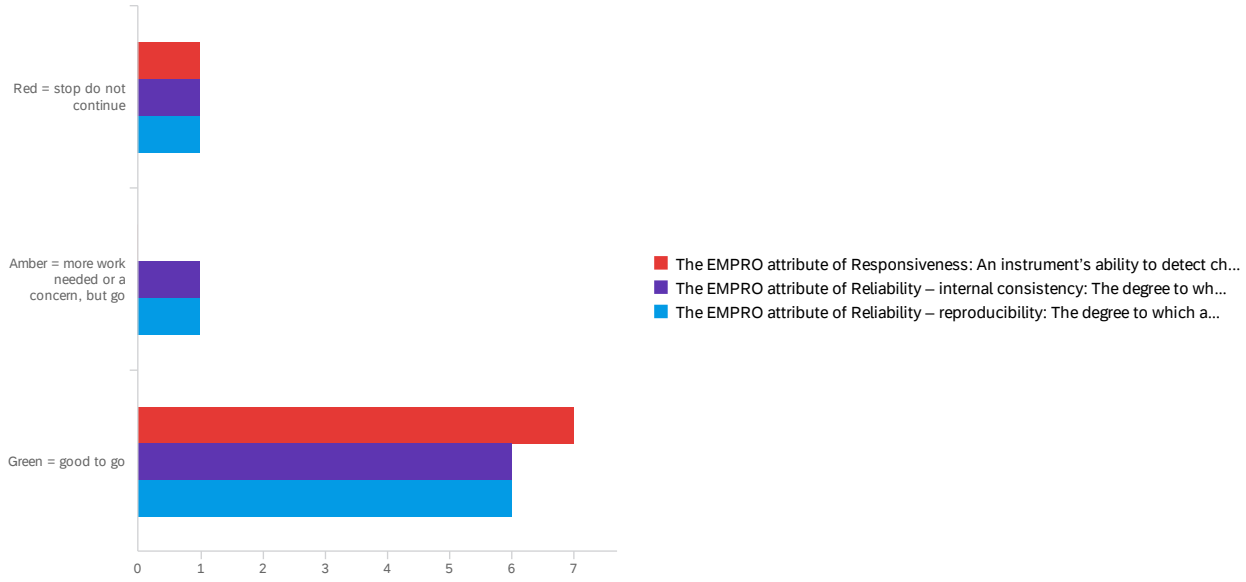


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	1.00	3.00	2.13	0.78	0.61	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%*	1.00	3.00	2.13	0.78	0.61	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	25.00% 2	37.50% 3	37.50% 3	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%*	25.00% 2	37.50% 3	37.50% 3	8

Showing rows 1 - 2 of 2

Q25 - PRTEE meets the discrimination OMERACT filter requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.



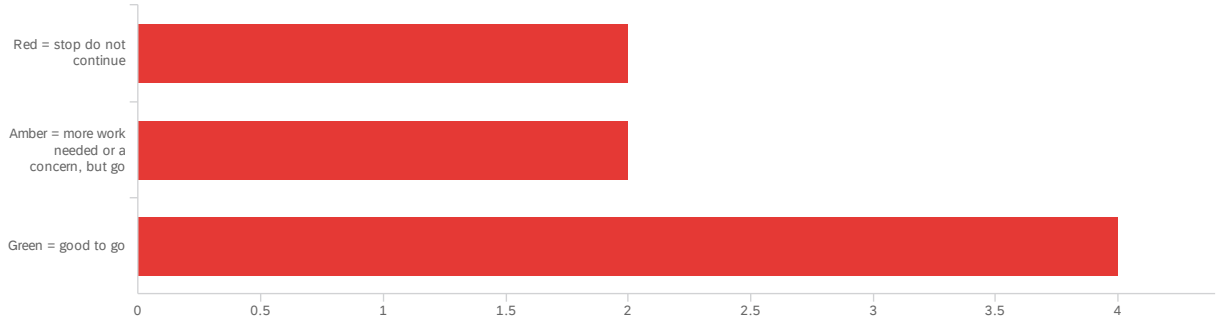
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	1.00	3.00	2.75	0.66	0.44	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	1.00	3.00	2.63	0.70	0.48	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	1.00	3.00	2.63	0.70	0.48	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	12.50% 1	0.00% 0	87.50% 7	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	12.50% 1	12.50% 1	75.00% 6	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	12.50% 1	12.50% 1	75.00% 6	8

Showing rows 1 - 3 of 3

Q26 - Considering your responses to the truth and discrimination OMERACT filter above and the fact that the EMPRO scores will overinflate the validity and overall score due to the lack of validity data for the pain subscale, does PRTEE meets the requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.

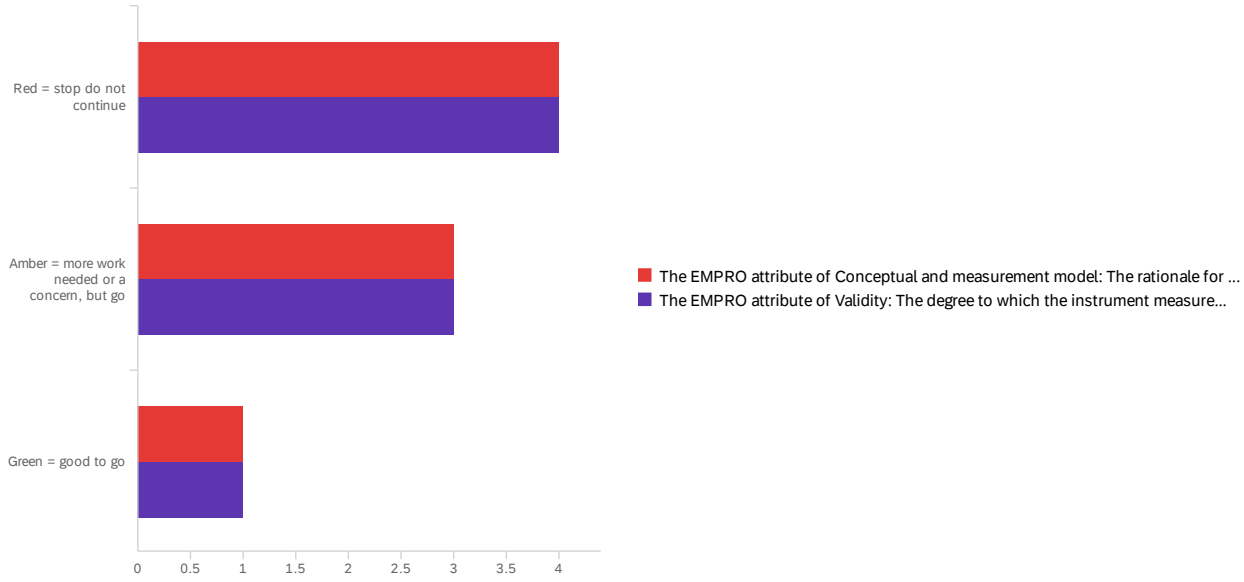


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 57.02%*	1.00	3.00	2.25	0.83	0.69	8

#	Field		Choice Count
1	Red = stop do not continue	25.00%	2
2	Amber = more work needed or a concern, but go	25.00%	2
3	Green = good to go	50.00%	4
			8

Showing rows 1 - 4 of 4

Q28 - The TEFS meets the truth OMERACT filter requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.

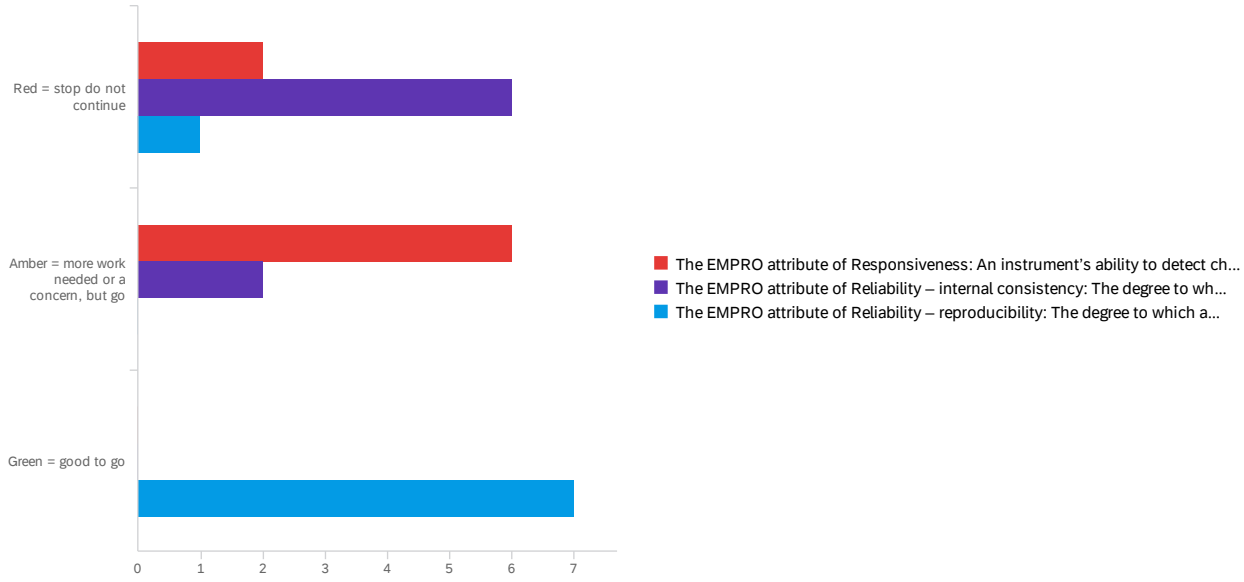


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 33.3%	1.00	3.00	1.63	0.70	0.48	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 33.3%	1.00	3.00	1.63	0.70	0.48	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 33.3%	50.00% 4	37.50% 3	12.50% 1	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 33.3%	50.00% 4	37.50% 3	12.50% 1	8

Showing rows 1 - 2 of 2

Q29 - TEFS meets the discrimination OMERACT filter requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.



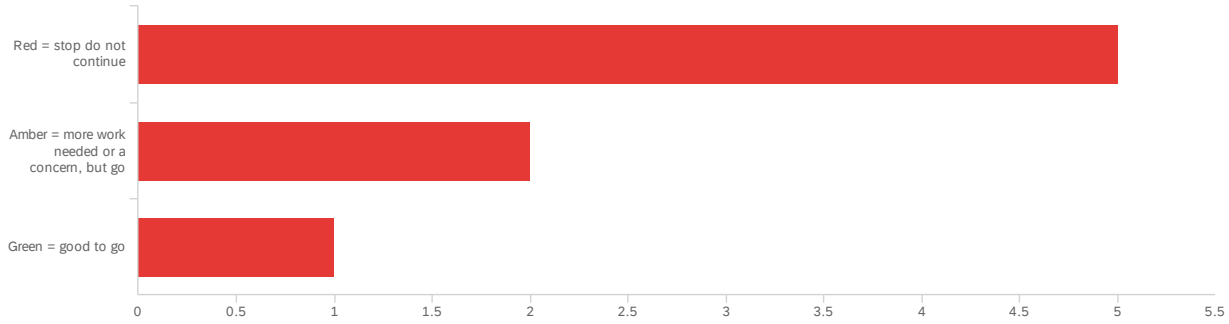
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 44.4%	1.00	2.00	1.75	0.43	0.19	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 8.3%	1.00	2.00	1.25	0.43	0.19	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	1.00	3.00	2.75	0.66	0.44	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 44.4%	25.00% 2	75.00% 6	0.00% 0	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 8.3%	75.00% 6	25.00% 2	0.00% 0	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	12.50% 1	0.00% 0	87.50% 7	8

Showing rows 1 - 3 of 3

Q30 - Considering your responses to the truth and discrimination OMERACT filter above, does TEFS meet the requirements for the outcome measure representing the Pain on Activity/Loading Domain in the core outcome set for lateral elbow tendinopathy.

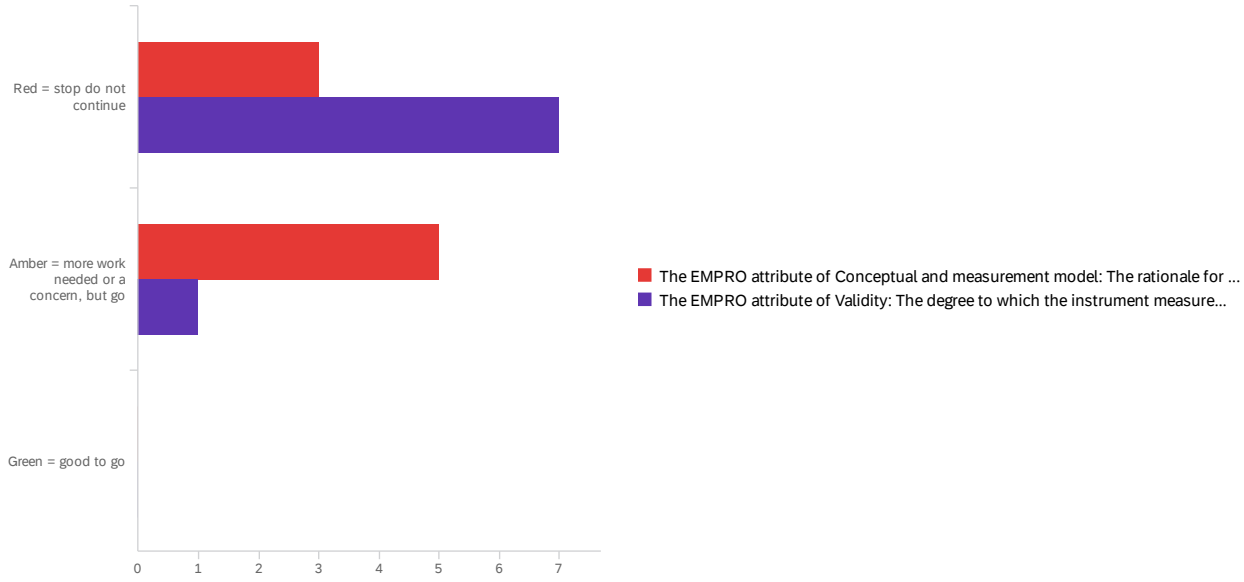


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 41.67%	1.00	3.00	1.50	0.71	0.50	8

#	Field		Choice Count
1	Red = stop do not continue	62.50%	5
2	Amber = more work needed or a concern, but go	25.00%	2
3	Green = good to go	12.50%	1
			8

Showing rows 1 - 4 of 4

Q36 - The TEFS meets the truth OMERACT filter requirements for the outcome measure representing the Pain over a Specified Time Domain in the core outcome set for lateral elbow tendinopathy.

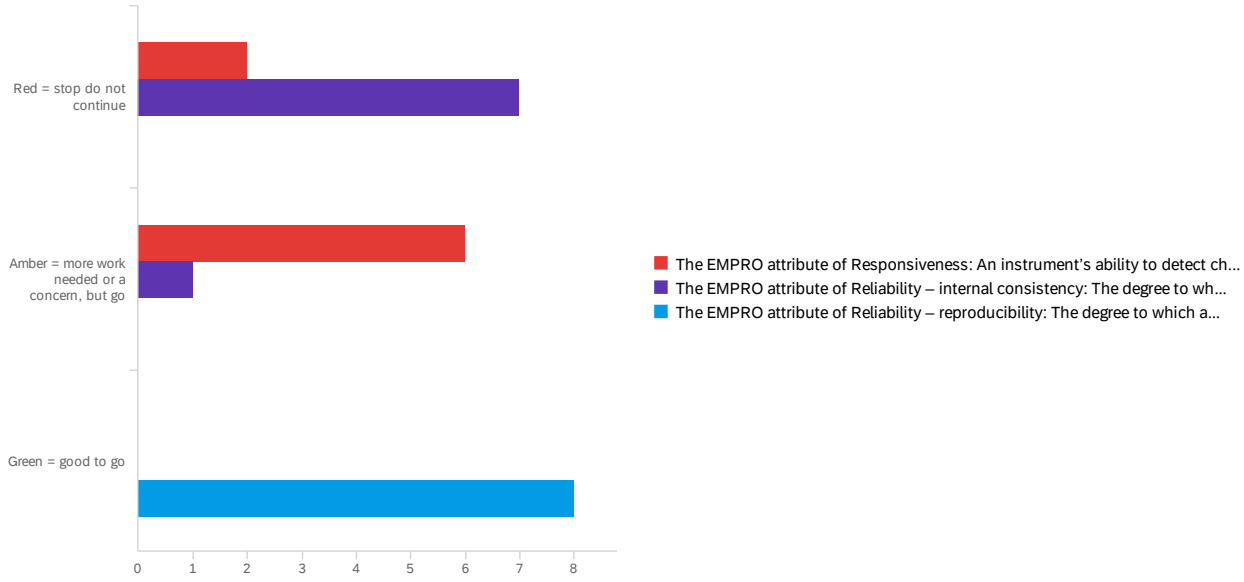


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 33.3%	1.00	2.00	1.63	0.48	0.23	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 33.3%	1.00	2.00	1.13	0.33	0.11	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 33.3%	37.50% 3	62.50% 5	0.00% 0	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 33.3%	87.50% 7	12.50% 1	0.00% 0	8

Showing rows 1 - 2 of 2

Q37 - TEFS meets the discrimination OMERACT filter requirements for the outcome measure representing the Pain over a Specified Time Domain in the core outcome set for lateral elbow tendinopathy.



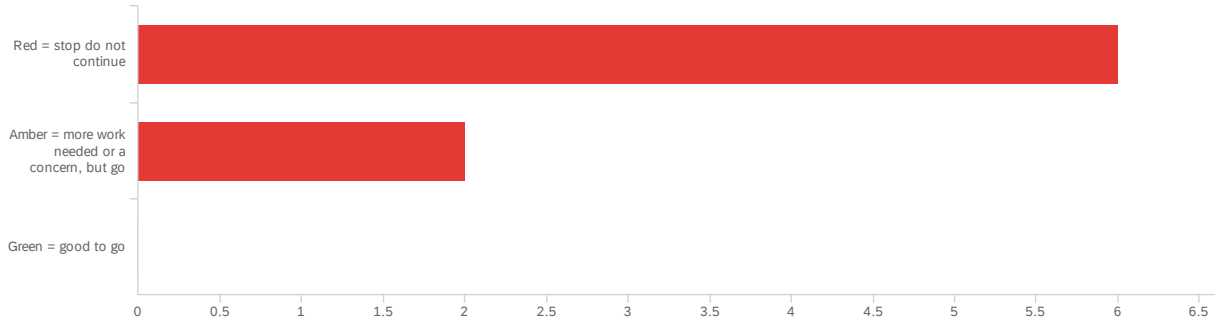
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 44.4%	1.00	2.00	1.75	0.43	0.19	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 8.3%	1.00	2.00	1.13	0.33	0.11	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	3.00	3.00	3.00	0.00	0.00	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 44.4%	25.00% 2	75.00% 6	0.00% 0	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 8.3%	87.50% 7	12.50% 1	0.00% 0	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	0.00% 0	0.00% 0	100.00% 8	8

Showing rows 1 - 3 of 3

Q38 - Considering your responses to the truth and discrimination OMERACT filter above, does TEFS meet the requirements for the outcome measure representing the Pain over a Specified Time Domain in the core outcome set for lateral elbow tendinopathy.

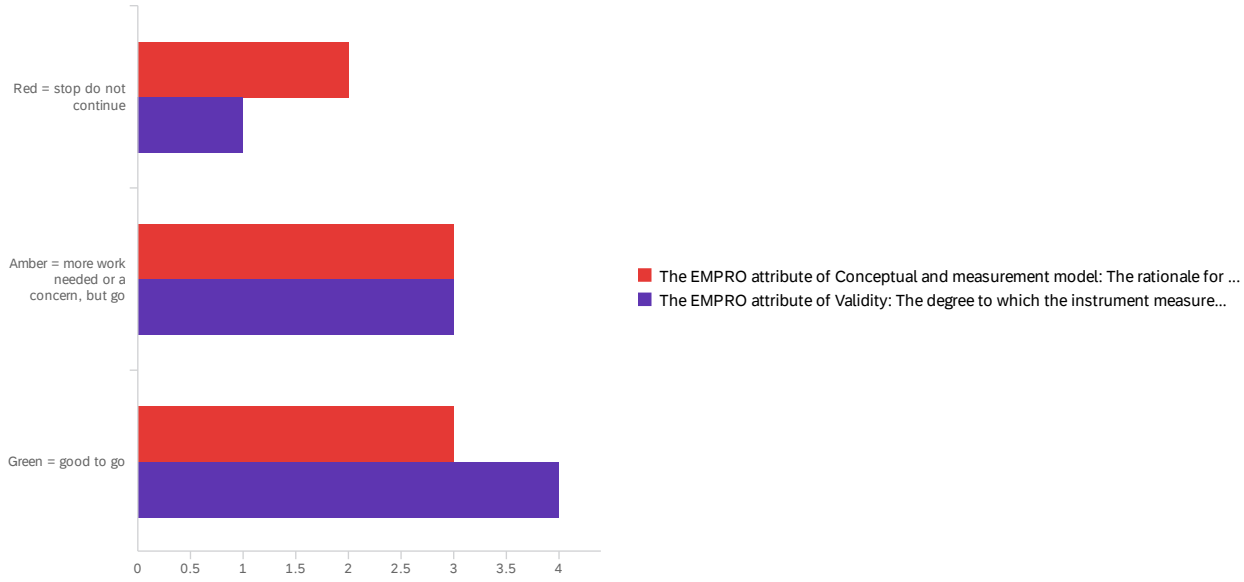


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 41.67%	1.00	2.00	1.25	0.43	0.19	8

#	Field		Choice Count
1	Red = stop do not continue	75.00%	6
2	Amber = more work needed or a concern, but go	25.00%	2
3	Green = good to go	0.00%	0
			8

Showing rows 1 - 4 of 4

Q32 - The PRTEE meets the truth OMERACT filter requirements for the outcome measure representing the Function Domain in the core outcome set for lateral elbow tendinopathy.

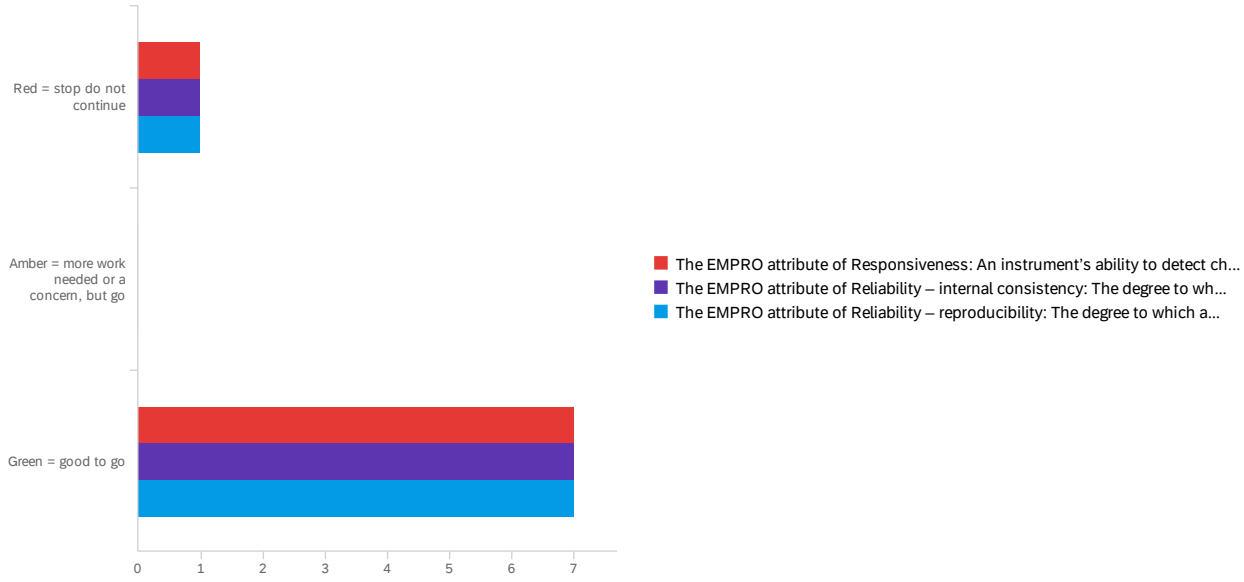


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	1.00	3.00	2.13	0.78	0.61	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%*	1.00	3.00	2.38	0.70	0.48	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 42.86%	25.00% 2	37.50% 3	37.50% 3	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 53.33%*	12.50% 1	37.50% 3	50.00% 4	8

Showing rows 1 - 2 of 2

Q33 - PRTEE meets the discrimination OMERACT filter requirements for the outcome measure representing the Function Domain in the core outcome set for lateral elbow tendinopathy.



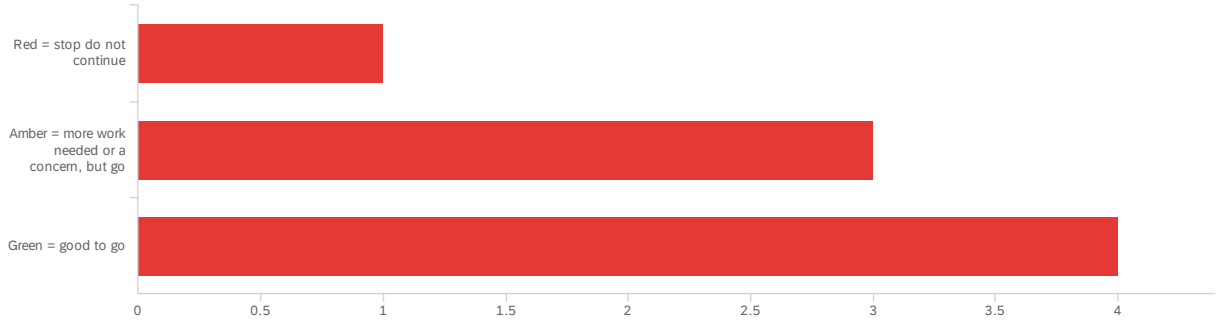
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	1.00	3.00	2.75	0.66	0.44	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	1.00	3.00	2.75	0.66	0.44	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	1.00	3.00	2.75	0.66	0.44	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 77.78%	12.50% 1	0.00% 0	87.50% 7	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 66.67%	12.50% 1	0.00% 0	87.50% 7	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 66.67%	12.50% 1	0.00% 0	87.50% 7	8

Showing rows 1 - 3 of 3

Q34 - Considering your responses to the truth and discrimination OMERACT filter above and the fact that the EMPRO scores will overinflate the validity and overall score due to the lack of validity data for the function subscale, does PRTEE meets the requirements for the outcome measure representing the Function Domain in the core outcome set for lateral elbow tendinopathy.

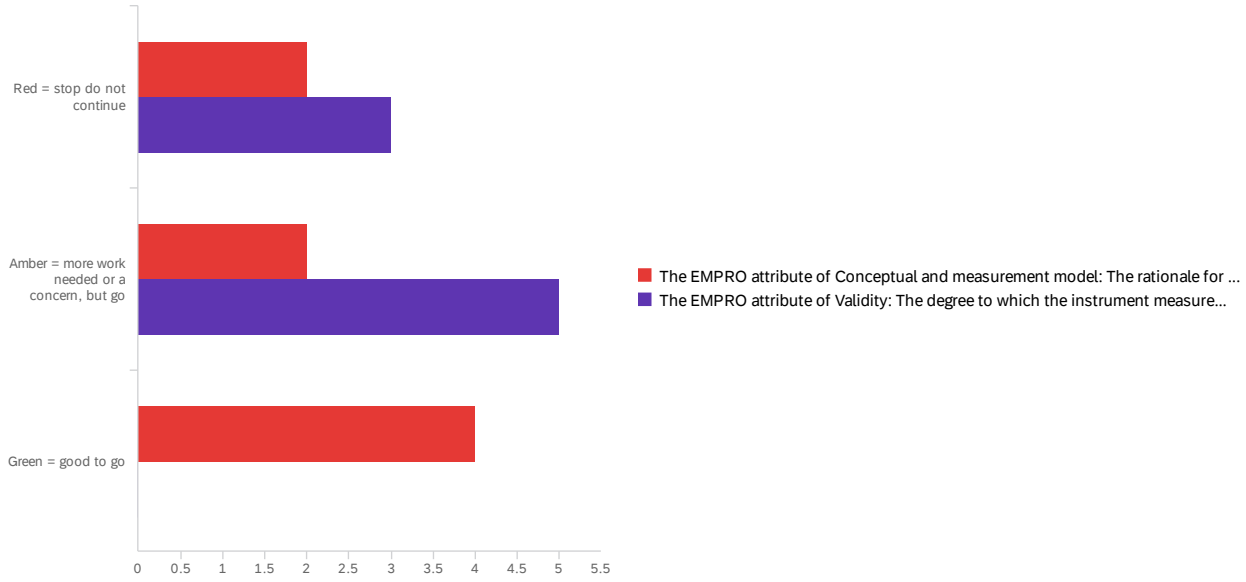


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 57.02%*	1.00	3.00	2.38	0.70	0.48	8

#	Field		Choice Count
1	Red = stop do not continue	12.50%	1
2	Amber = more work needed or a concern, but go	37.50%	3
3	Green = good to go	50.00%	4
			8

Showing rows 1 - 4 of 4

Q40 - The MGS meets the truth OMERACT filter requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.

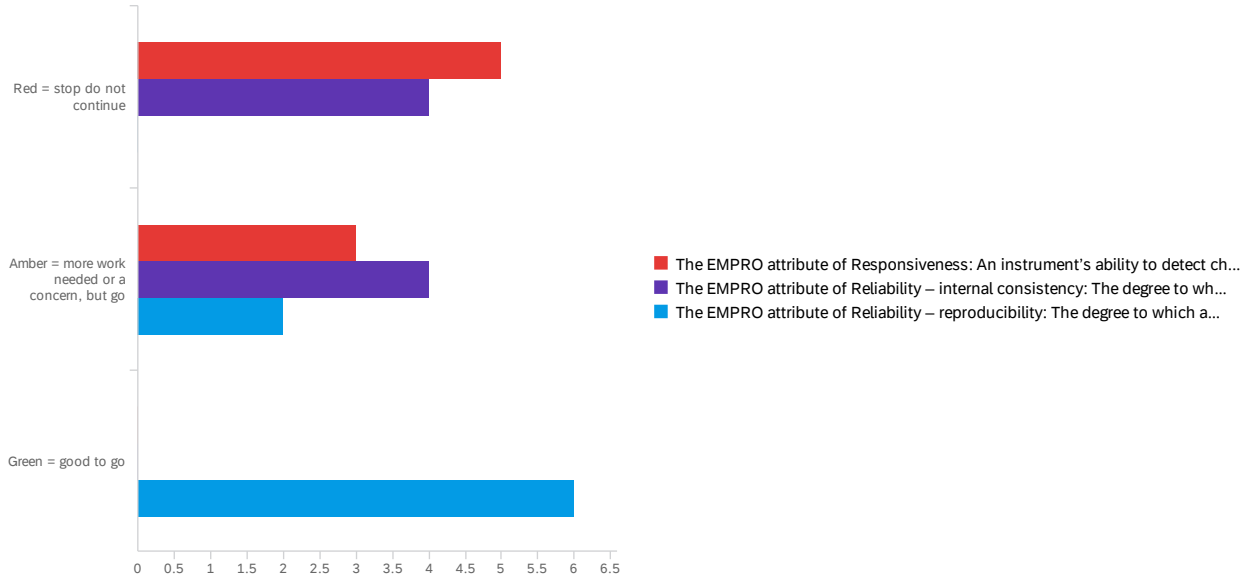


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 47.62%	1.00	3.00	2.25	0.83	0.69	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 0%	1.00	2.00	1.63	0.48	0.23	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 47.62%	25.00% 2	25.00% 2	50.00% 4	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 0%	37.50% 3	62.50% 5	0.00% 0	8

Showing rows 1 - 2 of 2

Q41 - MGS meets the discrimination OMERACT filter requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.



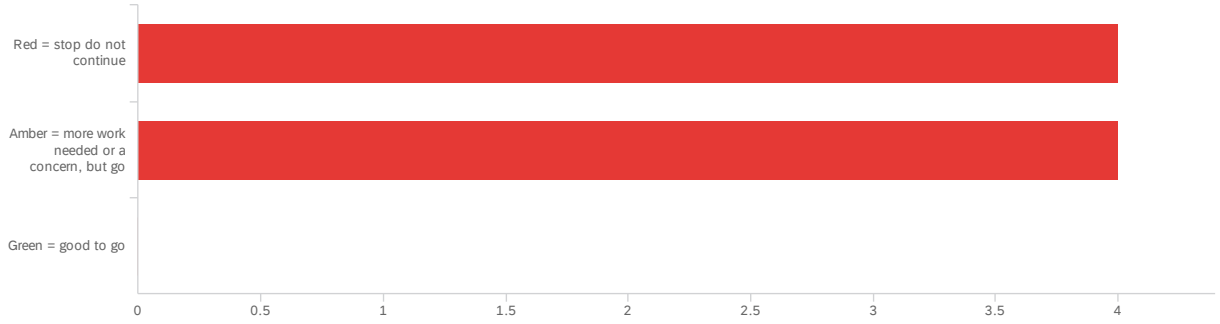
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 22.22%	1.00	2.00	1.38	0.48	0.23	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 0%	1.00	2.00	1.50	0.50	0.25	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	2.00	3.00	2.75	0.43	0.19	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 22.22%	62.50% 5	37.50% 3	0.00% 0	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 0%	50.00% 4	50.00% 4	0.00% 0	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	0.00% 0	25.00% 2	75.00% 6	8

Showing rows 1 - 3 of 3

Q42 - Considering your responses to the truth and discrimination OMERACT filter above and the fact that the EMPRO scores will overinflate the validity and overall score due to the lack of validity data for the function subscale, does MGS meets the requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.

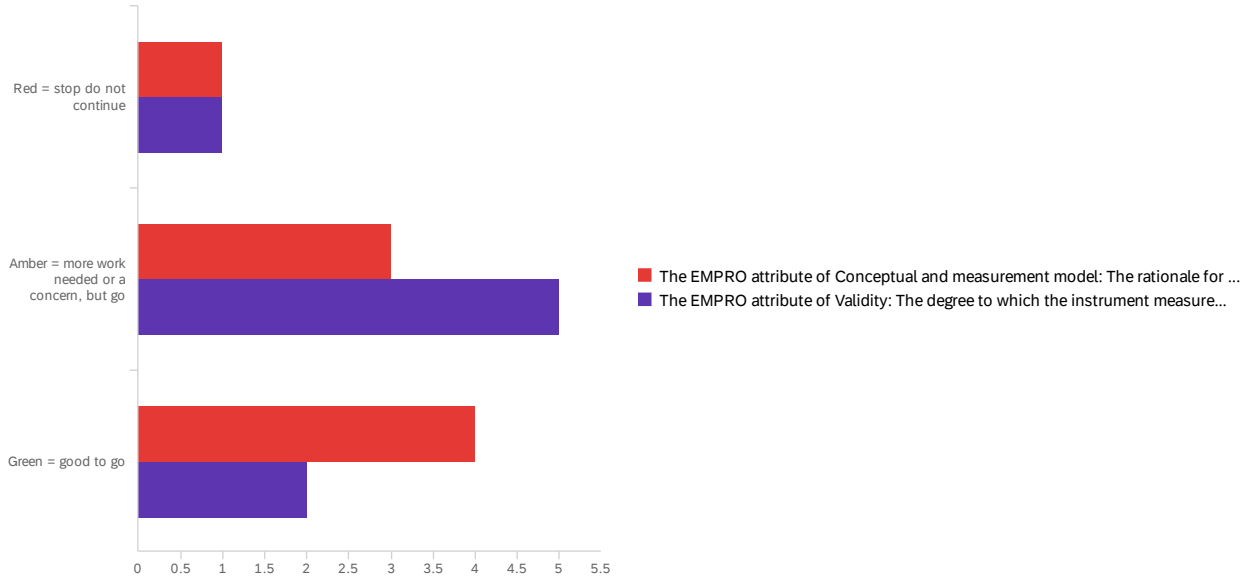


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 25.08%	1.00	2.00	1.50	0.50	0.25	8

#	Field		Choice Count
1	Red = stop do not continue	50.00%	4
2	Amber = more work needed or a concern, but go	50.00%	4
3	Green = good to go	0.00%	0
			8

Showing rows 1 - 4 of 4

Q44 - The PFG meets the truth OMERACT filter requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.

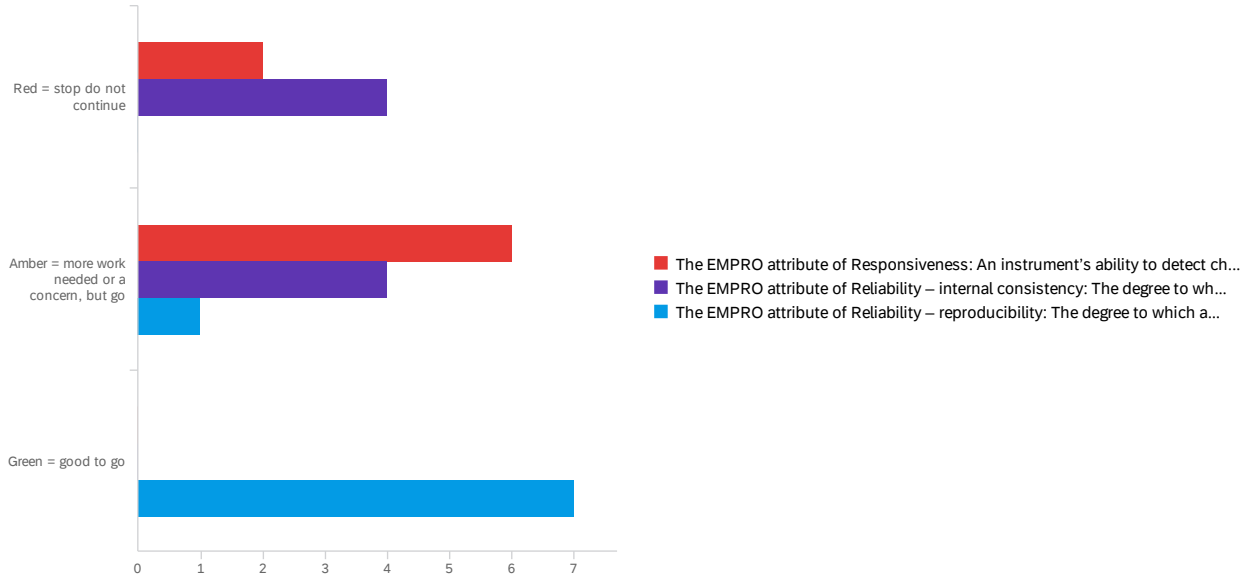


#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 47.62%	1.00	3.00	2.38	0.70	0.48	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 38.89%	1.00	3.00	2.13	0.60	0.36	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Conceptual and measurement model: The rationale for and description of the concept and the populations that a measure is intended to assess and the relationship between these concepts: score of 47.62%	12.50% 1	37.50% 3	50.00% 4	8
2	The EMPRO attribute of Validity: The degree to which the instrument measures what it purports to measure: score of 38.89%	12.50% 1	62.50% 5	25.00% 2	8

Showing rows 1 - 2 of 2

Q45 - PFG meets the discrimination OMERACT filter requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.



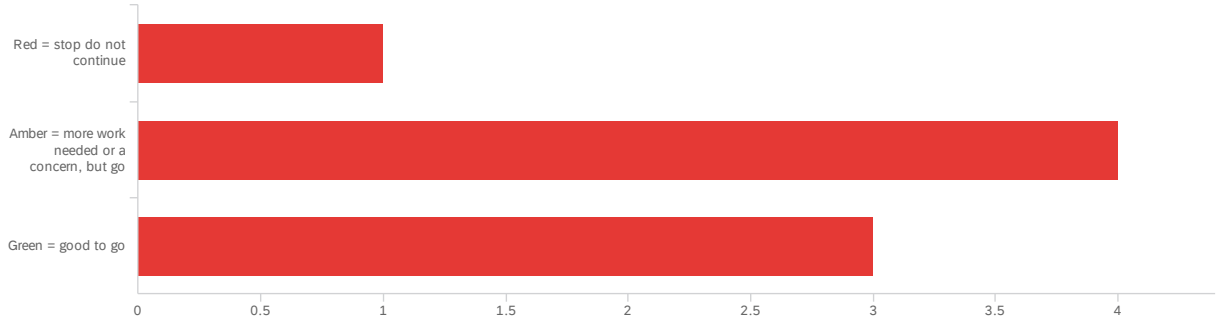
#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 22.22%	1.00	2.00	1.75	0.43	0.19	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 0%	1.00	2.00	1.50	0.50	0.25	8
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	2.00	3.00	2.88	0.33	0.11	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
1	The EMPRO attribute of Responsiveness: An instrument's ability to detect change over time: score of 22.22%	25.00% 2	75.00% 6	0.00% 0	8
2	The EMPRO attribute of Reliability – internal consistency: The degree to which an instrument is free from random error: score of 0%	50.00% 4	50.00% 4	0.00% 0	8

#	Field	Red = stop do not continue	Amber = more work needed or a concern, but go	Green = good to go	Total
3	The EMPRO attribute of Reliability – reproducibility: The degree to which an instrument is free from random error: score of 75%	0.00% 0	12.50% 1	87.50% 7	8

Showing rows 1 - 3 of 3

Q46 - Considering your responses to the truth and discrimination OMERACT filter above and the fact that the EMPRO scores will overinflate the validity and overall score due to the lack of validity data for the function subscale, does PFG meets the requirements for the outcome measure representing the Physical Function Capacity Domain in the core outcome set for lateral elbow tendinopathy.



#	Field	Minimum	Maximum	Mean	Std Deviation	Variance	Count
1	The overal EMPRO Score (a composite of the above attribute scores) is: 32.86%	1.00	3.00	2.25	0.66	0.44	8

#	Field		Choice Count
1	Red = stop do not continue	12.50%	1
2	Amber = more work needed or a concern, but go	50.00%	4
3	Green = good to go	37.50%	3
			8

Showing rows 1 - 4 of 4

End of Report

Delphi Process for Core Outcome Set for Lateral Elbow Tendinopathy (COS-LET): Report of findings from round 1 and 2 Surveys plus input from patient focus groups

In brief, we have now undertaken two surveys and patient focus groups on the matter of what outcome measures will be in the COS-LET. This report contains the results of this process to date – which will be the focus of our forthcoming consensus meeting. As well as the report below, we have attached an agenda, the domains paper, the first survey report (which you received with the second survey) for your reference, and a folder containing the outcome measures that we will be discussing at our consensus meeting.

Context:

1. There is a high level of heterogeneity in outcome measures used in trials of lateral elbow tendinopathy (LET), which makes evidence synthesis across studies difficult.
2. Previous work in the field of tendinopathy has established through a consensus exercise nine core health-related domains that should be measured in tendinopathy research.
3. The aim of this study is to develop a Core Outcome Set for Lateral Elbow Tendinopathy (COS-LET) mapping to these core domains.

Methods:

The development of the COS-LET is being developed as per the following process:

1. Systematic review of studies investigating LET has revealed a comprehensive list of all instruments that have previously been used to quantify treatment effect or outcome.
2. These instruments were matched to the list of nine core tendinopathy outcome domains by a Steering Committee of clinicians and researchers with a specialist interest in LET resulting in a set of candidate instruments.
3. You then responded to the first survey that asked you about the outcome measures for each domain. Seven patients also completed the first survey.
4. The committee then collated your responses, systematically reviewed the clinimetric/psychometric literature and rated each instrument using the EMPRO score – this information was then included in the second survey you completed.
5. You then responded to the second survey to determine what measures will be in the COS-LET, and also what may we consider in the interim (for those measures that did not make it into the COS-LET).
6. The committee then collated your responses and presented these results to two focus group meetings with some patients in the UK and Australia.
7. The results of your survey responses and the patient focus groups have now been collated and are now presented to you herein in the lead up to our consensus meeting. The report from the first survey is also appended for your reference.

Results (of second survey and patient focus group):

The results of the second survey of healthcare professionals and that of the patient focus groups are shown herein in Table 1 and 2. **Table 1** shows the characteristics of the healthcare professionals and patients participating in this consensus process. **Table 2** is a snapshot of the results of the second survey and the patient focus group meetings.

In summary, there was only one outcome that was considered (voted) to be in the core outcome set for lateral elbow tendinopathy (COS-LET) – the Patient Rated Tennis Elbow

Evaluation (PRTEE) for the Disability domain – and patients agreed. This will be ratified at our consensus meeting.

This then leaves us to make some decisions about which, if any, measures we will recommend as interim measures for the remaining domains – and importantly to plan for ongoing work in developing the COS-LET. To this end, there was agreement for PRTEE (some items on the pain subscale) to be used in the interim as a measure of pain on activity/loading domain. This was also the case for the function domain – PRTEE function subscale – though one group of patients (AUS) indicated that some of the items may not cover their specific issues and that other activities/functions may be more relevant to their specific case. We plan to commence the meeting with discussion about the PRTEE as an interim measure for the pain and function domains.

As you can see in Table 2 there are 4 domains that have no clinimetric properties and some discordance with survey results and patient views – this will be a focus of discussions at the consensus meeting. An exception to this is grip strength in that it did have some clinimetric information to consider, but discordant views between survey response and patients.

Table 1: Participant Characteristics (n (%) unless otherwise stated) of those who completed the full survey and provided these details (39 participants commenced, but 2 did not complete). UK focus group patients were a sub-group of those who completed the survey, whereas the AUS patient sub-group had not completed the survey.

Characteristics	Healthcare Professionals (N=37)	Patients (N=7)	AUS focus group patients (N=2)	UK focus group patients (N=3)
Sex: Male	25 (67.6)	2 (28.6)	2 (100)	1 (33)
Age: median (IQR; min-max) years	51 (43-57; 34-68)	48 (47.5-54.5; 26-59)	36.5 (36-37)	51.7 (48-59)
Role:				
Clinician	2 (5.4)			
Researcher	5 (13.5)			
Clinician Researcher	30 (81.1)			
Not a Clinician or Researcher		7 (100)	2 (100)	3 (100)
Highest academic qualification:				
PhD	21 (56.8)			
Master	6 (16.2)	2 (28.6)		1 (33)
Doctor of Medicine	6 (16.2)			
Postgraduate Diploma/Certificate			1 (50)	
Bachelor	3 (8.1)	3 (42.9)	1 (50)	1 (33)
Undergraduate Diploma/Certificate		1 (14.3)		1 (33)
Not specified	1 (2.7)			
No university qualification		1 (14.3)		
Profession:				
Physiotherapist	16 (43.2)			
Orthopaedic surgeon	14 (37.8)			
Sports & Exercise Medicine Physician	3 (8.1)			
Not specified	3 (8.1)			
Rheumatologist	1 (2.7)			
Patient		7 (100%)	2 (100)	3 (100)
Lateral elbow tendinopathy:				
Current history	1 (2.7)	5 (71.4)	2 (100)	1 (33)
Past history	10 (27.0)	4 (57.1)		2 (67)
Country where work:				
Australia	11 (29.7)	2 (28.6)	2 (100)	
United Kingdom	10 (27.0)	5 (71.4)		3 (100)
USA	3 (8.1)			
Canada and Norway each:	2 (5.4)			

Belgium, Finland, Greece, Israel, Italy, Netherlands, Spain, Sweden, and Turkey each:	1 (2.7)

Table 2: Summary of second survey and patient focus groups. [Note: Strong message to avoid over-burdening with too many questionnaires (UK patients)]

Insufficient psychometric evidence In COS-LET Interim suggestion Interim suggestion - patients Unsure Not in COS-LET	In COS-LET? Yes	No	Unsure	Interim Suggestion? Yes	No	UK Focus Group (n=3/5)	UK Comments	AUS Focus Group (n=2/4)	AUS comments
Disability									
DASH	8.11	67.57	24.32						
Oxford Elbow Score	16.22	51.35	32.43						
PRTEE (Patient Rated Tennis Elbow Evaluation)	70.27	13.52	16.22						
Quick DASH	59.46	24.32	16.22						
Pain on activity/loading									
Tennis Elbow Functional Scale	10.81	72.97*	16.22	18.92	81.08				
PRTEE	64.86	18.92	16.22	83.78	16.22				Items in PRTEE-pain too specific, may not cover everyone. Gripping nominated as main provocative movement
Function									
PRTEE	64.86	13.51	21.62	89.19	10.81				Some items in PRTEE-function may not represent their experience

Insufficient psychometric evidence In COS-LET Interim suggestion Interim suggestion - patients Unsure Not in COS-LET	**In COS-LET?**			**Interim Suggestion?**		UK Focus Group (n=3/5)	UK Comments	AUS Focus Group (n=2/4)	AUS comments
Pain over a specified timeframe	**Yes**	**No**	**Unsure**	**Yes**	**No**				
Tennis Elbow Functional Scale	8.11	72.97*	18.92	24.32	75.68*		It was felt that this domain had been covered by PRTEE and EQ5D		Descriptors seem reasonably similar to PRTEE. Not sure how valuable would be to have a single snap shot of a day. Depends on what doing that day. PRTEE better
Proposed by expert: ?NRS or VAS about pain over specified time									
Participation (daily activities, work, sport)							? Covered by section B of PRTEE but only reflects the last week		
Return to sport				56.76	43.24		One not sporty, so not relevant to them		
Time off work				70.27	29.73		2 participants didn't have any time off		May not be sensitive. Could there be a qualifying question regarding occupation/physical activity involved? Changed to 'time off from daily activities' might be more applicable
Total Elbow Scoring System				10.81	89.19*				

Insufficient psychometric evidence In COS-LET Interim suggestion Interim suggestion - patients Unsure Not in COS-LET	In COS-LET?			Interim Suggestion?		UK Focus Group (n=3/5)	UK Comments	AUS Focus Group (n=2/4)	AUS comments
	Yes	No	Unsure	Yes	No				
Patient rating of condition									
Global perceived effect score				35.14	64.86				
Global Rating of Change				56.76	43.24		More inclusive having words and numbers		
Patient Satisfaction Scale				45.95	54.05		Satisfaction is different to the effect of treatment		
Physical function capacity (including strength)									
Grip strength (maximum)	16.22	45.95	37.84	32.43	67.57		How can you accurately measure max strength if inhibited by pain?		
Pain free grip strength	40.54	29.73	29.73	64.86	35.14				
Psychological factors									
Hospital Anxiety and Depression Scale				37.84	62.16		Lacks relevance		A lot of items don't reflect psyc status related to the condition
Tampa Scale of Kinesophobia				43.24	56.76		Concerns regarding too many questions (TSK-17) but more condition-specific		Both felt more relevant to their condition

Insufficient psychometric evidence In COS-LET Interim suggestion Interim suggestion - patients Unsure Not in COS-LET	**In COS-LET?**			**Interim Suggestion?**		UK Focus Group (n=3/5)	UK Comments	AUS Focus Group (n=2/4)	AUS comments
Quality of life	**Yes**	**No**	**Unsure**	**Yes**	**No**				
EQ5D				59.46	40.54				
SF-12				37.84	62.16		I see that and switch off! It just looks bad. It is an assault on the eyes! People would just glaze over		More broader and wholistic? More context when trying to answer questions. 5 vs 12 items not an issue

Core Outcome Set for Lateral Elbow Tendinopathy (COS-LET) Tool 2022

Patient-rated tennis elbow evaluation:

*The questions below will help us understand the amount of difficulty you have had with your arm in the past week. You will be describing your **average** arm symptoms **over the past week** on a scale 0-10. Please provide an answer for all questions. If you did not perform an activity because of pain or because you were unable ,then you should circle a "10". If you are unsure please estimate to the best of your ability. Only leave items blank if you never perform that activity. Please indicate this by drawing a line completely through the question.*

1. PAIN in your affected arm											
*Rate the average amount of pain in your arm **over the past week** by circling the number that best describes your pain on a scale from 0-10. A **zero** (0) means that you **did not have any pain** and a **ten** (10) means that you had **the worst pain imaginable.***											
RATE YOUR PAIN: No Pain											Worst Imaginable
When your are at rest	0	1	2	3	4	5	6	7	8	9	10
When doing a task with repeated arm movement	0	1	2	3	4	5	6	7	8	9	10
When carrying a plastic bag of groceries	0	1	2	3	4	5	6	7	8	9	10
When your pain was at its least	0	1	2	3	4	5	6	7	8	9	10
When your pain was at its worst	0	1	2	3	4	5	6	7	8	9	10

Please turn the page.....

Bateman M, *et al. Br J Sports Med* 2022;0:1–10. doi: 10.1136/bjsports-2021-105044

2. FUNCTIONAL DISABILITY

A. SPECIFIC ACTIVITIES

*Rate the **amount of difficulty** you experienced performing each of the tasks listed below, over the past week, by circling the number that best describes your difficulty on a scale of 0-10. A zero (0) means you did not experience any difficulty and a **ten (10)** means it was **so difficult you were unable to do it at all**.*

	No Difficulty										Unable To Do
Turn a doorknob or key	0	1	2	3	4	5	6	7	8	9	10
Carry a grocery bag or briefcase by the handle	0	1	2	3	4	5	6	7	8	9	10
Lift a full coffee cup or glass of milk to your mouth	0	1	2	3	4	5	6	7	8	9	10
Open a jar	0	1	2	3	4	5	6	7	8	9	10
Pull up pants	0	1	2	3	4	5	6	7	8	9	10
Wring out a washcloth or wet towel	0	1	2	3	4	5	6	7	8	9	10

B. USUAL ACTIVITIES

*Rate the **amount of difficulty** you experienced performing your **usual** activities in each of the areas listed below, over the past week, by circling the number that best describes your difficulty on a scale of 0-10. By "usual activities", we mean the activities that you performed **before** you started having a problem with your arm. A **zero** (0) means you did not experience any difficulty and a **ten** (10) means it was so difficulty you were unable to do any of your usual activities.*

1. Personal activities (dressing, washing)	0	1	2	3	4	5	6	7	8	9	10
2. Household work (cleaning, maintenance)	0	1	2	3	4	5	6	7	8	9	10
3. Work (your job or everyday work)	0	1	2	3	4	5	6	7	8	9	10
4. Recreational or sporting activities	0	1	2	3	4	5	6	7	8	9	10

Comments:

Pain subscale total = …… (max 50)

Function subscale total (specific activities + usual activities /2) = …… (max 50)

Total (pain subscale + function subscale) = …… (max 100)

Pain items 1, 4, 5 total = …… (max 30)

Pain-free grip strength:

Pain-free grip is measured using a handgrip dynamometer that measures grip pressure in pounds or kilograms. This is done with the patient's shoulder in the neutral position, the elbow in 90 degrees of flexion, and the forearm in the neutral position of supination/pronation. The patient slowly squeezes the dynamometer handle until he or she feels pain at the elbow. The force generated is recorded, and the average value of 3 repeated assessments is used for data analysis.

Left Arm		Right Arm	
Measurement 1	……	Measurement 1	……
Measurement 2	……	Measurement 2	……
Measurement 3	……	Measurement 3	……
Mean	……	Mean	……

Please rate from 0-10 how much pain you have when gripping strongly.

No pain Worst imaginable

0 1 2 3 4 5 6 7 8 9 10

In the last month, how many days have you had off work due to your elbow pain?

☐ days Not applicable (I don't work) ☐

Bateman M, *et al. Br J Sports Med* 2022;0:1–10. doi: 10.1136/bjsports-2021-105044