This article has been accepted for publication in BJSM following peer review.
The definitive copyedited, typeset version of record is available online at 10.1136/bjsports-2023-107741

TENDINopathy Severity assessment – Achilles (TENDINS-A): Evaluation of reliability and validity in accordance with COSMIN recommendations.

Myles C Murphy [1,2,] Fergus McCleary [2], Dana Hince [3], Ruth L Chimenti [4], Paola T Chivers [3,5], J. Turner Vosseller [6], Sophia Nimphius [5], Nonhlanhla Mkumbuzi [7-10], Peter Malliaras [11], Nicola Maffuli [12-14], Robert-Jan De Vos [15], Ebonie Rio [16-18]

[1] Nutrition and Health Innovation Research Institute, School of Medical and Health Sciences, Edith Cowan University, Joondalup, Western Australia, Australia

[2] School of Health Sciences and Physiotherapy, The University of Notre Dame Australia, Fremantle, Western Australia, Australia

[3] Institute for Health Research, The University of Notre Dame Australia, Fremantle, Western Australia, Australia

[4] Department of Physical Therapy and Rehabilitation Science, University of Iowa, Iowa City, Iowa, United States of America

[5] School of Medical and Health Sciences, Edith Cowan University, Joondalup, Western Australia, Australia

[6] Jacksonville Orthopaedic Institute, Jacksonville, Florida, United States of America

[7] Department of Sports, Exercise, and Rehabilitation, Northumbria University, Newcastle upon Tyne, United Kingdom

[8] Department of Rehabilitation, Midlands State University, Gweru, Zimbabwe

[9] Department of Human Movement Science, Nelson Mandela University, Summerstrand, Gqeberha, South Africa

[10] NtombiSport Pty Ltd, Cape Town, South Africa

[11] School of Primary and Allied Health Care, Faculty of Medicine Nursing and Health Science, Monash University, Victoria, Australia

[12] Department of Medicine, Surgery and Dentistry, University of Salerno, Baronissi, Salerno, Italy

[13] Faculty of Medicine, School of Pharmacy and Bioengineering, Keele University, Newcastle, United Kingdom

[14] Barts and the London School of Medicine and Dentistry, Centre for Sports and Exercise Medicine, Queen Mary University of London, London, United Kingdom

[15] Department of Orthopaedics and Sports Medicine, Erasmus MC University Medical Centre, Rotterdam, The Netherlands.

[16] La Trobe Sport and Exercise Medicine Research Centre, La Trobe University, Bundoora, Victoria, Australia

[17] Victorian Institute of Sport, Melbourne, Victoria, Australia

[18] The Australian Ballet, Melbourne, Victoria, Australia

***Corresponding author:** Dr Myles C Murphy. Nutrition and Health Innovation Research Institute, School of Medical and Health Sciences, Edith Cowan University, Joondalup, Western Australia, Australia Email: m.murphy@ecu.edu.au

ABSTRACT

Objective: To evaluate the construct validity (structural validity and hypothesis-testing), reliability (test-retest reliability, measurement error and internal consistency) and minimal important change (MIC) of the 13-item TENDINopathy Severity assessment–Achilles (TENDINS-A).

Methods: Participants with Achilles pain completed an online survey including: demographics, TENDINS-A, Foot and Ankle Outcome Score (FAOS), and Victorian Institute of Sport Assessment-Achilles (VISA-A). Exploratory factor analysis (EFA) assessed dimensionality. Confirmatory Factor Analysis (CFA) assessed structural validity [root-mean-squared error of approximation (RMSEA); Comparative Fit Index (CFI); Tucker-Lewis Index (TLI); Standardised Root Measure Square (SRMS)]. Correlations between TENDINS-A and the FAOS or VISA-A assessed hypothesis-testing. Intraclass correlation (ICC) assessed test-retest reliability. Cronbach's α assessed internal consistency. Standard error of the measurement (SEM) assessed measurement error. A distribution-based approach assessed MIC.

Results: Seventy-nine participants (51% female) with a mean (SD) age=42.6 (13.0) years, height=175.0 (11.7) cm and body mass=82.0 (19.1) kg were included. EFA identified three meaningful factors, proposed as pain, symptoms and function. The best model identified using CFA for TENDINS-A had structural validity (RMSEA= 0.101, CFI= 0.959, TLI= 0.947, SRMS=0.068), which included three factors (Pain, Symptoms, and Function), but excluded three items from the original TENDINS-A. TENDINS-A exhibited moderate positive correlation with FAOS (rho=0.598,$p<0.001$) and a moderate, negative correlation with VISA-A (r=-0.639,$p<0.001$). Reliability of the TENDINS-A was excellent (ICC=0.930; Cronbach's α=0.808; SEM=6.54 units), with an MIC of 12 units.

Conclusions: Our evaluation of the revised 10-item TENDINS-A determined it has construct validity and excellent reliability, compared to the VISA-A and FAOS which lack content and construct validity. The TENDINS-A is recommended as the preferred patient-reported outcome measure to assess disability in people with Achilles tendinopathy.

What is already known: The TENDINopathy Severity assessment – Achilles (TENDINS-A) has been co-designed with patients, clinicians and researchers as a new patient-reported outcome measure (PROM) to assess the severity of disability in Achilles tendinopathy, but its psychometric properties have not been previously reported.

What this study adds: The TENDINS-A has content validity, construct validity and excellent reliability in the assessment of the severity of disability in Achilles tendinopathy.

How this study might affect practice: The TENDINS-A is recommended for use in both clinical practice and research to assess the severity of disability in Achilles tendinopathy.

INTRODUCTION

Achilles tendinopathy is characterised by focal Achilles tendon pain accompanied by impaired function with mechanical loading,[1] which is typically managed using predominantly exercise rehabilitation.[2 3] Achilles tendinopathy is common in the running population [prevalence of 4.2% (diagnosed using a pain mapping approach) in runners training for an event][4 5] and also prevalent within the general population.[6] People with tendinopathy, expert clinicians and researchers have identified several core health domains as important in tendinopathy, including tendon-related disability.[7] However, multiple different measures to quantify tendon-related disability are used across studies, with no existing consensus.[8 9] In clinical practice measures of tendon-related disability are not commonly used,[10] in part due to inadequate content validity.[11 12] Thus, despite a consensus that tendon-related disability is an important aspect of health,[7] there remains no measure of tendon-related disability that is validated for research or clinical use.

The consensus-based standards for the selection of measurement instruments (COSMIN) guidelines represent international best-practice for the design and appraisal of Patient Reported Outcome Measures (PROMs).[13 14] However, the value of a PROM is not in its availability but in its quality, reflected by its clinometric properties.[15] Valid and reliable PROMs should be used and recommended, but this is not currently possible in Achilles tendinopathy research of the disability domain.[16] Existing PROMs that assess tendon-related disability, such as the Victorian Institute of Sport Assessment – Achilles (VISA-A),[16 17] have historically been recommended in the absence of other tools,[18] but do not satisfy the guidelines set out by COSMIN and suffer from substantial flaws in their design.

Using the VISA-A as an example, it has inadequate content and structural validity. VISA-A development did not involve patients[11] and methodological testing of the VISA-A scale using Rasch analysis has demonstrated it has insufficient structural validity[19] and hypothesis testing of the VISA-A shows it is significantly correlated to participant age and body mass index.[20] The recommendations from several reviews and original research studies highlight the VISA-A is not valid and should not be used in research or clinical practice.[11 12 19 20]

The newly developed TENDINopathy Severity assessment – Achilles (TENDINS-A) assesses the severity of the Tendinopathy Core Health Domain of Disability,[7] and consists of questions covering sub-domains of pain, symptoms, and physical function related to Achilles tendinopathy.[21] The TENDINS-A has adequate content validity, being co-designed by people with Achilles tendinopathy, expert clinicians and researchers.[21] Whilst the TENDINS-A has sufficient content validity, the structural validity and other measurement properties of this new PROM (e.g., reliability) remain unknown.

Objective

The objective of this study was to evaluate the construct validity (including structural validity and hypothesis testing), reliability (including test-retest reliability, measurement error and internal consistency) and minimal important change of the TENDINS-A.

Hypotheses

The TENDINS-A will have adequate criteria-based structural validity, demonstrate moderate significant correlations to commonly utilised PROMS for Achilles tendinopathy[16] [VISA-A and Foot and Ankle Outcome Score (FAOS)], no correlation to unrelated constructs such as baseline characteristics (e.g., age or body mass index) and excellent reliability.

METHODS

Study Design

Cross-sectional cohort study evaluating the psychometric properties of the TENDINS-A.

Participants and setting

We used a network of greater than 20 clinicians known to the research team (exercise scientists, general practitioners, orthopaedic surgeons, physiotherapists, podiatrists, rheumatologists and sport and exercise physicians) and Achilles tendon researchers to identify participants with either mid-portion or insertional Achilles tendinopathy and provide them with our online survey using Qualtrics (convenience sampling). This was accessed via a quick response (QR) code or anonymous web link. Any person over the age of 18 years, who could read the English language, with self-reported Achilles tendinopathy was eligible.

Inclusion criteria

Participants self-identified as having Achilles tendon pain and were provided with a pain map of established locations of Achilles tendon symptoms and asked to select all markings that corresponded to their region of pain (which is superior to both palpation and ultrasound tissue characterisation for pain localisation).[22] Participants

who selected the insertion of mid-portion Achilles tendon were included within this validation study and this methodology provides 'near-perfect' agreement to Sports Physician diagnosis of Achilles tendinopathy.[23]

Outcome measures

All outcome measures were completed by participants within a single survey and the survey forced responses to avoid missing data.

Participant characteristics

Age (years), sex (male/female/intersex), height (cm), body mass (kg), ethnicity, country of residence, languages other than English spoken by the participant (self-reported), whether the participant performed moderate to vigorous physical activity (MVPA) most days (yes/no), highest level of education, work status and total household income were reported.

TENDINopathy Severity assessment – Achilles (TENDINS-A)

The original TENDINS-A (which consists of 13 scoreable items numbered 3, 4, 6, 7, 8, 10, 11, 12, 13A, 13B, 13C, 13D, 13E, and 4 non-scoreable items numbered 1, 2, 5, 9) was provided to participants as the first PROM within the survey and was scored between 0 and 100, with '0' representing a perfect score (no disability) and '100' representing complete disability.[21] If participants were unable to perform one of the pain-with-loading tests (e.g., single leg hop), they were instructed to leave it blank and a score of '10' was provided. Participants completed the TENDINS-A prior to the FAOS and the VISA-A. Directly after completing the VISA-A questionnaire, the TENDINS-A was immediately repeated. This ensured the clinical status of participants was unchanged,[24] whilst still allowing for any potential interference effects. Specifically, the participant would be unable to remember responses to the initial TENDINS-A as they have no option to go back in the survey and 50 questions (i.e., the VISA-A and FAOS) separated the initial and then repeated TENDINS-A.

Foot and Ankle Outcome Score

The 42-item FAOS[25] was performed following the TENDINS-A, with a mean total score calculated for the purposes of this study and used for hypothesis testing. A score of '0' represented a perfect score (no disability), with a score of '100' representing complete disability. The FAOS is a reliable tool and has adequate internal consistency,[25] however it has inadequate content validity (as it was not developed for Achilles tendinopathy) and its construct validity in Achilles tendinopathy is unknown.

Victorian Institute of Sport Assessment – Achilles

The 8-item VISA-A[17] was performed following the FAOS and used for hypothesis testing. The VISA-A is inversely scored when compared to the TENDINS-A and FAOS with a score of '100' represented a perfect score (no disability), and a score of '0' representing complete disability. Where participants selected multiple responses in item 8 (related to how much physical activity can be performed before cessation due to symptoms is required), the lowest score was retained for analysis. The VISA-A is a reliable tool and has adequate internal consistency,[17] however it has inadequate content validity[11] and construct validity.[19]

Power calculation

Study size was informed by the recommendations of the COSMIN risk of bias checklist[13 14] rather than formal power calculations. Questions 1, 2, 5 and 9 are not scored in the TENDINS-A; therefore the PROM was initially considered to have 13 scorable items as one item has five secondary scales.[21] COSMIN guidelines suggest that six-times the number of persons to items is adequate for assessing structural validity using classical test theory,[13 14] thus the minimum sample size for this study was determined to be 78 persons. We also required all communalities in exploratory factor analysis (EFA) to be >0.3, thus samples of <100 participants are justified.[26] Furthermore, a sample of >50 participants is considered adequate for reliability and internal consistency.[13 14]

Statistical analysis

The different statistical analyses, per each measurement property, are described below. All statistical analyses were performed within IBM SPSS statistics (version 29.0), or IBM SPSS Amos (version 29.0) and we have provided a glossary of statistical terminology used in Table 1. Where appropriate, confidence intervals are presented, and statistical significance was considered when $p<0.05$.

For the purposes of clarity in this manuscript, when referencing an Item (e.g., Item 7) these will always be in reference to the original 13 items of the TENDINS-A. However, item numbers for the final TENDINS-A (presented as an appendix) are different and have been re-ordered to avoid confusion (e.g., labelled 1-10).

Table 1. Glossary of statistical terminology

Term	Definition
Bartlett's test of Sphericity	Assesses the probability that the correlation matrix has significant correlations amongst items in the dataset, which is a requirement for factor analysis.
Communality	Proportion of the common variance for an individual item, relative to all factors.
Comparative Fit Index	Assessment of model fit by quantifying discrepancies between the sample data and the model while adjusting for sample size.
Eigenvalues	Quantification of the proportion of the variance accounted for by an individual

	factor.
Kaiser-Meyer-Olkin Measure of Sampling Adequacy	An assessment of whether data is appropriate for factor analysis by determining the proportion of variance among all items that might be from a common variance.
Root-mean-squared error of the approximation	Estimate of the discrepancy the sample data and the model-implied data covariance matrices per degree of freedom.
Scree Plot	Graphical representation of eigenvalues to visually assess when new factors are unlikely to explain a sufficient level of the variance.
Standardised Root Measure Square	Square root of the calculated discrepancy between the sample covariance matrix and the model covariance matrix.
Tucker-Lewis Index	Measure of goodness of fit in the analysis of covariance structures.

Dimensionality

We performed EFA for all scoreable items using SPSS Statistics to assess dimensionality. To proceed to confirmatory factor analysis (CFA), we required all communalities to exceed 0.3,[27] required the Kaiser-Meyer-Olkin Measure of Sampling Adequacy to exceed 0.7 [28] and for Bartlett's test of Sphericity to be significant ($p<0.01$),[29] which are all recognised criteria for progression to factor analysis.

Factors were eligible for inclusion within EFA where the Eigenvalues exceeded 1.0 (which are deemed substantial). The Scree plot was also visually inspected and additional models were performed to include more/ less factors, based on author (MCM) judgement and with the TENDINS-A being proposed to measure three factors related to disability: pain, symptoms and function.[21] A principal axis factoring method was used to determine our initial factor matrix. As per standard practice, factors were considered meaningful when an item loaded at ≥0.4 for that factor.[30] As 'disability' is a domain that reflects pain, symptoms and function,[7] we hypothesised the TENDINS-A to have three factors identified by EFA.

Structural Validity

Confirmatory factor analysis was performed using SPSS Amos (22) to assess structural validity. We included all three factors (which were identified in the EFA and TENDINS-A content validity study),[21] within the initial model, and then removed variables as needed to achieve the best fit. Items with categorical (yes/ no only) responses were not included within CFA (not appropriate for CFA but had been tested in EFA). Models were

tested and selected based off the best root-mean-squared error of the approximation (RMSEA) and overall model chi-squared statistics, with a lower chi-squared statistic representing a better model fit. An overall acceptable model would require a Comparative Fit Index (CFI) OR Tucker-Lewis Index (TLI) of >0.95 AND RMSEA of <0.06 or Standardised Root Measure Square (SRMS) of <0.08.[13]

The CFA provided standardised factor loading and error variance per item. Composite reliability of >0.70 is considered adequate.[31] Composite reliability was calculated using the formula below when λ = standardised factor loading and ε = error variance:

$$Compositereliability = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + (\sum \varepsilon_i)}$$

Convergent validity was assessed using the average variance extracted (AVE). An AVE of >0.50 is considered adequate.[32] The AVE was calculated using the formula below when λ = standardised factor loading and ε = error variance:

$$Averagevarianceextracted = \frac{\sum_{i=1}^{n} \lambda_i^2}{\sum_{i=1}^{n} \lambda_i^2 + \sum_{i=1}^{n} Var(\varepsilon| \quad |i \quad)}$$

The criteria for adequate structural validity of the final model were:

1. CFI or TLI >0.95,[13]
2. RMSEA of <0.06 or SRMS of <0.08,[13]
3. Composite Reliability >0.70,[31] and
4. AVE >0.50.[32]

Test-retest reliability

Test-retest reliability of the TENDINS-A was calculated as a continuous scale. The absolute and relative test-retest reliability (which is a sub-type of reliability and demonstrates how closely someone will generate the same score on a PROM when repeated in a specified timeframe) were determined.[33] Relative test-retest reliability was reported as the intraclass correlation (ICC), with a two-way mixed approach used to assess absolute agreement of a single measure.[34] Absolute test-retest reliability is a sub-type of reliability and demonstrates the degree of uncertainty in a measurement. For example, measurements with more uncertainty

will have greater measurement error, hence it is systematic and random error that is not reflective of true change.[34] The standard error of the measurement[35] was calculated using the following equation:

$$Standarderrorofthemeasurment = SD_{baseline}\sqrt{1 - ICC}$$

Minimal important change

The Minimal important change (MIC) and 95% confidence intervals was calculated using a distribution-based method,[36] using the following equation:

$$Minimalimportantchange = \frac{SD_{baseline}}{2}$$

Internal consistency

Internal consistency is a sub-type of reliability that investigates whether different items that measure the same construct give comparable outcomes. The internal consistency for the entire 13-item TENDINS-A was calculated using the average inter-item correlation and a Cronbach's Alpha was reported with a positive rating given for values >0.7.[13]

Convergent and divergent validity

In addition to the convergent validity assessed using CFA, we also assessed convergent and divergent validity of the TENDINS-A scale using a correlation between scores of the TENDINS-A and another PROM theorised to measure the same overall core health domain of tendinopathy (disability)[7 16] or unrelated constructs, such as age and BMI (given they are continuous variables that should not be related to tendon-related disability). The TENDINS-A, FAOS and the VISA-A were assessed for normality using visual assessment and the one-sample Kolomogorov-Smirnov Test. For normally distributed data the Pearson's correlation coefficient was used and for non-normally distributed data the non-parametric Spearman rho correlation was used. Correlations were considered very weak (between 0-0.25), weak (0.26-0.49), moderate (0.5-0.69), strong (0.7-0.89) and very strong (0.9-1).

Comparison between sub-groups

To ensure that significant difference between sub-groups did not exist we calculated between-group differences for key baseline characteristics [sex (male/ female), MVPA most days (yes/no), multilingual status (yes/ no), and whether tendon loading activities are typically performed daily (yes/no)] using independent t-tests.

Equity, diversity and inclusion

The authorship group are of diverse gender, geographical location and research experience. Furthermore, the participants included in this study were diverse in relation to many demographic variables (e.g., sex, age, BMI, geographical location or education level).

RESULTS

Participants

Seventy-nine (n=79) participants (51% female, 49% male) provided survey data and were included within our analysis with no missing data. Participants had a mean (SD) age of 42.6 (13.0) years, height of 175.0 (11.7) cm and body mass of 82.0 (19.1) kg. Participants were predominantly of 'Australian' ethnicity (79.7%) and were not multi-lingual (64.6%) with those multilingual participants predominantly having English as their first language (53.6%). A significant number of participants reported that they performed MVPA most days (69.6%). Participant education level and employment varied, with most having completed tertiary studies (74.6%) and working full time (63.3%). Income levels ranged from less than $30,000 Australian Dollars (AUD) (5.1%) to greater than $200,000 AUD (27.8%), per annum. The activities that aggravated the pain of participants with Achilles tendinopathy varied: walking slow (15.2%), walking fast (44.3%), walking up and down stairs (43.0%), running up and down stairs (46.8%), running slow (60.8%), running fast (54.4%), hopping and jumping (62.0%), and rapidly changing direction while running (44.3%). The complete breakdown of participant characteristics is provided in Appendix A.

Uni-dimensionality

Bartlett's test of Sphericity suggested that the results of our correlation matrix were not random [$\chi^2(78) = 708$, $p<0.001$]. Our Kaiser-Meyer-Olkin Measure of Sampling Adequacy (0.811) far exceeded the minimum cut-off of 0.70. Thus, our correlation matrix was appropriate. The extracted communalities (Mean=0.67, SD= 0.16, Range= 0.18 to 0.93) for all but three items (items 6, 8 and 12) exceeded our cut-off score for CFA inclusion (Appendix B).

Principal axis factoring analysis was performed and the initial Eigen values of three factors exceeded 1.0 (see appendix C), explained 67.9% of the total variance and the first factor accounted for >20% of the variability. The three identified factors were included within the Factor Matrix, and associated factor loading per item, are presented within Table 2. However, no items were associated with a loading of >0.4 on Factor Three.

Table 2. Factor Loading with Principal Axis Factoring Analysis for all items with loading >0.4 (n=79).

TENDINS- A	Factor One	Factor Two	Factor

			Three
Item 3	0.412		0.400
Item 4	0.496	0.479	
Item 6			
Item 7		0.644	
Item 8		0.504	
Item 10		0.672	
Item 11	0.614		
Item 12			
Item 13A	0.731		
Item 13B	0.888		
Item 13C	0.915		
Item 13D	0.916		
Item 13E	0.917		

*Blank cells represent loading of <0.4

Structural Validity Several Maximum Likelihood Estimates CFA models were trialled, informed by the factor loading from EFA, and subscales from our content validity study to generate our final CFA model. The best model used three factors [a) Pain, b) Symptoms, and c) Function] and eight items (Figure 1). Question 12 was excluded because of poor EFA initial communality, non-meaningful factor loading within EFA and being unable to load onto either of the three factors within CFA. Question 13A and 13B were removed as their exclusion resulted in substantially improved model fit (Appendix D). The scoring for the remaining 12 questions was amended to maintain a score range from 0 to 100, as presented within Appendix E.

In the final model, the chi-squared statistic was $\chi^2(22)= 39.2$, the CFI was 0.959, the TLI was 0.947, the RMSEA was 0.101 and the SRMS was 0.068, indicating sufficient structural validity. The model was also assumed accurate with CMIN/DF= 1.796, parsimony comparative fit index= 0.753 and parsimony normed fit index= 0.717.

The standardised factor loadings and error variance for the final model are included within Table 3. The correlation between ‘Pain’ and ‘Symptoms’ was estimated at 0.642 with significant covariances (Estimate=

4.104, standard error= 1.101 and composite reliability= 3.728, p<0.001). The correlation between 'Pain' and 'Function' was estimated at 0.668 with significant covariances (Estimate= 5.424, standard error= 1.300 and composite reliability= 4.171, p<0.001). The correlation between 'Symptoms' and 'Function' was estimated at 0.275 with significant covariances (Estimate= 2.818, standard error= 1.353 and composite reliability= 2.082, p=0.037).

Table 3. Standardised Factor Loadings of the TENDINS-A (n=79)

TENDINS- A	Factor	Standardised Factor Loading	Error Variance
Item 11	Pain	0.643	0.587
Item 3	Pain	0.605	0.634
Item 4	Pain	0.627	0.607
Item 7	Symptoms	0.741	0.451
Item 10	Symptoms	0.855	0.269
Item 13C	Function	0.921	0.152
Item 13D	Function	0.975	0.049
Item 13E	Function	0.967	0.065

Internal consistency. For the 8-items assessed using CFA included in TENDINS-A theorised to measure 'disability', the internal consistency was 0.842, hence displaying sufficient internal consistency.

Composite reliability. For the 8-items assessed using CFA included in TENDINS-A theorised to measure 'disability', the composite reliability was 0.934, hence displaying sufficient composite reliability.

Convergent validity. For the 8-items assessed using CFA included in TENDINS-A theorised to measure 'disability', the AVE was 0.648, hence displaying sufficient convergent validity.

Thus, our final CFA model could determine the contribution of different items in measuring 'Pain', 'Symptoms' and 'Function' and the TENDINS-A has sufficient structural validity. Consequently, the scores for the remaining items were amended to ensure the TENDINS-A retained a maximal score of 100 (40 units for 'Pain', 30 units for 'symptoms' and 30 units for 'function'). Questions 6 and 8 were also excluded from CFA modelling being dichotomous (Yes/No) items but were retained within the complete TENDINS-A in the 'symptoms' domain. All subsequent analyses therefore relate to the revised TENDINS-A, presented in Appendix E.

Distribution of data

Whilst data for the TENDINS-A (Kolomogorov-Smirnov test statistic=0.09, p=0.097) and the VISA-A (Kolomogorov-Smirnov test statistic= 0.06, p=0.200) were normally distributed, the FAOS were not (Kolomogorov-Smirnov test statistic= 0.12, p=0.019). The TENDINS-A data had a minimum score of 0 and a maximum score of 100. The mean (SD) of the TENDINS-A was 47.89 (24.71) points, the FAOS was 23.60 (18.63) percent, and the VISA was 53.58 (23.96) points. The mean (SD) of the repeated TENDINS-A was 44.34 (24.16) points.

Reliability

Intraclass correlation co-efficient

The 10-item TENDINS-A mixed-effect ICC for absolute agreement of a single measure was ICC= 0.930 (95% CI = 0.881 to 0.959), p= < 0.001.

Internal consistency

The Cronbach's Alpha for the 10-items of the TENDINS-A was reported as α=0.808, representing excellent internal consistency.

Standard error of the measurement

The TENDINS-A standard error of the measurement was calculated as 6.54 units.

Minimal important change

The MIC for the 10-item TENDINS-A was calculated as 12.36 (95% CI= 5.46 to 19.25) units of difference, representing 25.6% points of change from a mean TENDINS-A score of 47.89.

Convergent validity

The TENDINS-A exhibited a moderate positive correlation with the FAOS (rho=0.598, 95%CI= 0.408 to 0.738, p<0.001), and a moderate, negative correlation with the VISA-A (r=-0.639, 95%CI= -0.764 to -0.467, p<0.001).

Divergent validity

The TENDINS-A showed no evidence of a statistically significant association with age (p=0.426) or BMI (p=0.189).

Comparison between sub-groups

Between-group differences (Table 4) were observed for self-reported MVPA most days, with those performing moderate to vigorous physical activity most days having a lower TENDINS-A score (p=0.002) and males reported a lower TENDINS-A score (p=0.044). No differences in TENDINS-A score were observed for those performing specific tendon-loading exercise versus those who did not (p=0.079), or whether participants were multi-lingual (p=0.397). As more males performed MVPA than females in our sample, we performed a *post hoc* linear regression model to assess whether group differences due to sex were a result of higher levels of MVPA in males than females. The model supported this idea, with no association between the TENDINS-A score and sex (p=0.061) after adjusting for MVPA [17.7 (95%CI= 6.46 to 28.97) unit increase in TENDINS-A score, p=0.002].

Table 4. Between group differences in TENDINS-A score.

		Mean (SD), range	Mean Difference (SED)	Test statistic (Independent t-test)	p-value
Moderate to vigorous physical activity most days	Yes (n=55)	37.69 (23.08), 0-91	-18.81 (5.91)	-3.18	0.002*
	No (n=24)	56.50 (26.54), 5-98			
Sex	Male (n=39)	37.56 (23.27), 5-98	-11.54 (5.63)	-2.048	0.044*
	Female (n=40)	49.10 (26.64), 0-91			
Tendon-loading exercise performed most days	Yes (n=43)	38.79 (22.35), 5-86	-10.13 (5.69)	-1.779	0.079
	No (n=36)	48.92 (28.23), 0-98			
Multilingual	Yes (n=28)	48.79 (23.57), 5-91	8.34 (5.97)	1.396	0.167

	No (n=51)	40.45 (26.32), 0-98			

Abbreviations: SD= standard deviation; SED= standard error of difference; n= number; range= min to max

* $p<0.05$

DISCUSSION

The present investigation evaluated the construct validity (consisting of structural validity and hypothesis testing), as well as reliability (consisting of internal consistency, test-retest reliability and measurement error) of the TENDINS-A. Based on this investigation, we present a revised TENDINS-A (Appendix E) that still includes all original questions, but only 10 of the original questions are scoreable.

The 10-item TENDINS-A, proposed to provide an overall measure of 'disability', demonstrated sufficient construct validity for three factors (pain; symptoms; function) and reliability, supporting its use in clinical and research settings. Furthermore, with normally distributed scores, there was no evidence of floor or ceiling effects within the sample. The TENDINS-A has a median completion time of ~8 minutes and is also likely to be meaningful to people with a lived experience of Achilles tendinopathy, as the research team included such individuals to inform its development.[21]

All continuous variables of the TENDINS-A, after the removal of item 12, 13A and 13B, were able to load onto one of three factors: pain, symptoms or function. The core health domain of 'disability' is defined as the "composite scores of a mix of patient-rated pain and disability due to the pain, usually relating to tendon-specific activities/tasks," and included in tools such as the VISA-A.[7 16] Thus, the TENDINS-A is proposed to have adequate structural validity and measure the overall composite of disability, using scores from items related to 'pain', 'symptoms' and 'function'. Therefore, the TENDINS-A is the only current PROM quantifying Achilles tendinopathy related disability with adequate content[21] and structural validity and should be used in place of all other PROMs. Researchers may still wish to use the VISA-A in addition to the TENDINS-A to allow for comparisons to previous research, but only the TENDINS-A should be used to report primary outcomes.

The reliability of the TENDINS-A is excellent (ICC>0.90). We opted to assess test-retest reliability on a single day as tendinopathy pain and symptoms are known to fluctuate daily.[24] The internal consistency of the TENDINS-A was also excellent ($\alpha>0.80$), ensuring all items were being scored consistently. The terminology surrounding the minimal detectable change, MIC and minimal clinically important difference (MCID) is

confusing and for the purposes of discussion we have included all available comparisons. We calculated the minimal important change (MIC=12.4 units) of the TENDINS-A using a distribution-based method. This is larger than the MIC for the VISA-A, which is reported as between 7.8 to 8.2.[37] A study using an anchor-based method in the VISA-A reported the MCID as 14 units,[38] which is comparable to our estimated measure. Another study using the VISA-A estimated the MCID to be 6.5 units, however this was estimated from a selective sample of 15 participants with insertional Achilles tendinopathy, which may also explain the smaller MCID.[39] To characterize the magnitude of change on the TENDINS-A that is meaningful to individuals with Achilles tendinopathy, future longitudinal studies are needed to estimate the minimal clinically-important difference (MCID).

The TENDINS-A demonstrated moderate correlations to both the VISA-A and FAOS, which is expected given they are proposed to measure the same overall health domain of disability.[16] We did not expect strong correlations to the VISA-A or FAOS given the limitations of the VISA-A,[11 12 15] which we theorised would extend to the FAOS as it is also lacking content validity. This lack of a strong correlation may be due to the TENDINS-A being a superior PROM to the FAOS and VISA-A (given the established content and structural validity of the TENDINS-A). The lack of a strong correlation can also result from the absence of a gold standard outcome measure for disability. Alternatively, we theorised the TENDINS-A should not have been associated with baseline characteristics such as age or BMI, and this was confirmed by our analysis. This differs from the VISA-A, whose total score is significantly associated with age and BMI,[20] reflecting a problem for previous studies using the VISA-A that did not adjust for age or BMI within the statistical analysis.

None of our specified sub-groups (sex, being multilingual, MVPA most days, tendon-loading exercise most days) showed a floor or ceiling effect. We expect this is likely due to items of the TENDINS-A being specifically linked to the function and aggravating tasks typically performed by the participant, making it suitable for both sedentary and physically active people with Achilles tendinopathy. Performing MVPA most days was associated with a lower TENDINS-A score, as expected. This is most likely a consequence of the numerous health benefits of physical activity.[40] Alternatively, the inactive group may undertake less tendon-specific loading exercise and aggravate the Achilles less. However, this seems unlikely, as performing tendon-loading activity was also assessed and no significant effect detected.

Limitations

The sample for this study, whilst diverse in some respects, was heavily reliant on participants living in Australia. While other PROMS for Achilles tendinopathy showed acceptable cross-cultural validity in other languages,[41 42] future research on cross-cultural adaptation of the TENDINS-A is needed. Cross-cultural adaptation of the TENDINS-A is currently being performed by several different research groups internationally and these data can also be used to assess differential item functioning.

Self-report may have resulted in a misdiagnosis of Achilles tendinopathy. However in this study, we enhanced diagnostic accuracy by implementing a standardized pain map.[22] A prior investigation demonstrated a 93% concordance between patient-reported Achilles tendinopathy when using a pain map and diagnoses made by a physician.[23] These findings suggest that, in the majority of cases, self-reported Achilles tendinopathy aligns with the clinical diagnosis.

Finally, further details related to medical co-morbidities and medication usage[43] would have allowed analysis of their influence on total score and would be recommended in future research trials.

CONCLUSION

Our evaluation of the revised TENDINS-A (Appendix E), which includes 10-scoreable items across three sub-domains of pain, symptoms and physical function has adequate construct validity and reliability. These findings, and previously established content validity, ensure that the revised TENDINS-A with 10-scoreable items can be recommended for immediate use in both research and clinical practice, being the preferred tool over the VISA-A and FAOS to assess disability in individuals with Achilles tendinopathy.

Author Contributions: Myles C Murphy, Dana Hince and Ebonie K Rio conceived the study and designed the methods. Fergus McCleary collected all data. Myles C Murphy, Dana Hince, Paola T Chivers and Sophia Nimphius performed all analysis. All authors contributed to interpretation and manuscript preparation.

Competing Interests and Funding: No authors have any competing interests, and this research was not funded.

Acknowledgements: Dr Myles C Murphy is supported by a Western Australian Department of Health Innovation Fellowship (G1006615) and a Near-Miss Award (G1006605).

Data Sharing: The data from this research study will be considered to be shared upon reasonable request of the corresponding author.

Ethical Approvals: This research was approved by the University of Notre Dame Australia Human Research Ethics Committee (ID: 2022-175F) and all participants gave informed, electronic consent.

Patient involvement: The TENDINS-A was co-designed with people having a lived experience of Achilles tendinopathy and was found to have adequate content validity, ensuring that the TENDINS-A is relevant to patients.

Equity, diversity and inclusion: The authorship group are of diverse gender, geographical location and research experience. Furthermore, the participants included in this study were diverse in relation to many demographic variables (e.g., sex, age, BMI, geographical location or education level).

REFERENCES

1. Scott A, Squier K, Alfredson H, et al. ICON 2019: International Scientific Tendinopathy Symposium Consensus: Clinical Terminology. *British journal of sports medicine* 2020;54(5):260-62. doi: 10.1136/bjsports-2019-100885
2. Murphy M, Travers MJ, Gibson W, et al. The rate of improvement of pain and function in mid-portion Achilles tendinopathy with loading protocols: A Systematic Review and Longitudinal Meta-Analysis. *Sports medicine (Auckland, NZ)* 2018;48(8):1875-91. doi: 10.1007/s40279-018-0932-2.
3. Murphy MC, Travers MJ, Chivers P, et al. Efficacy of heavy eccentric calf training for treating mid-portion Achilles tendinopathy: a systematic review and meta-analysis. *British journal of sports medicine* 2019;53(17):1070-77. doi: 10.1136/bjsports-2018-099934
4. Chen W, Cloosterman KLA, Bierma-Zeinstra SMA, et al. Epidemiology of insertional and midportion Achilles tendinopathy in runners: A prospective cohort study. *Journal of sport and health science* 2023 doi: 10.1016/j.jshs.2023.03.007 [published Online First: 20230323]
5. Lopes AD, Hespanhol LC, Yeung SS, et al. What are the Main Running-Related Musculoskeletal Injuries? *Sports medicine (Auckland, NZ)* 2012;42(10):891-905. doi: 10.1007/bf03262301
6. Albers IS, Zwerver J, Diercks RL, et al. Incidence and prevalence of lower extremity tendinopathy in a Dutch general practice population: a cross sectional study. *BMC Musculoskelet Disord* 2016;17(1):16. doi: 10.1186/s12891-016-0885-2
7. Vicenzino B, de Vos R-J, Alfredson H, et al. 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. *British journal of sports medicine* 2019;54(8):444-51. doi: 10.1136/bjsports-2019-100894
8. Murphy M, Rio E, Debenham J, et al. Evaluating the progress of mid-portion Achilles Tendinopathy during rehabilitation. A Review of Outcome Measures for muscle structure and function, tendon structure and neural and pain associated mechanisms. *Int J Sports Phys Ther* 2018;13(3):537-51.
9. Murphy M, Rio E, Debenham J, et al. Evaluating the progress of mid-portion Achilles Tendinopathy during rehabilitation. A Review of Outcome Measures for Self- Reported Pain and Function. *Int J Sports Phys Ther* 2018;13(2):283-92. doi: 10.26603/ijspt20180283

10. Murphy M, Debenham J, Bulsara C, et al. Assessment and monitoring of Achilles Tendinopathy in clinical practice: A qualitative descriptive exploration of the barrier's clinicians face. *BMJ Open Sport Exerc Med* 2022;ePub ahead of print
11. Korakakis V, Kotsifaki A, Stefanakis M, et al. Evaluating lower limb tendinopathy with Victorian Institute of Sport Assessment (VISA) questionnaires: a systematic review shows very-low-quality evidence for their content and structural validity-part I. *Knee Surg Sports Traumatol Arthrosc* 2021;29(9):2749-64. doi: 10.1007/s00167-021-06598-5 [published Online First: 20210521]
12. Korakakis V, Whiteley R, Kotsifaki A, et al. A systematic review evaluating the clinimetric properties of the Victorian Institute of Sport Assessment (VISA) questionnaires for lower limb tendinopathy shows moderate to high-quality evidence for sufficient reliability, validity and responsiveness—part II. *Knee Surg Sports Traumatol Arthrosc* 2021 doi: 10.1007/s00167-021-06557-0
13. Gagnier JJ, Lai J, Mokkink LB, et al. COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Qual Life Res* 2021;30(8):2197-218. doi: 10.1007/s11136-021-02822-4 [published Online First: 20210405]
14. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018;27(5):1159-70. doi: 10.1007/s11136-018-1829-0 [published Online First: 20180317]
15. Korakakis V, Whiteley R, Kotsifaki A, et al. Tendinopathy VISAs have expired—is it time for outcome renewals? *Knee Surg Sports Traumatol Arthrosc* 2021 doi: 10.1007/s00167-021-06596-7
16. Grävare Silbernagel K, Malliaras P, de Vos R-J, et al. ICON 2020—International Scientific Tendinopathy Symposium Consensus: A Systematic Review of Outcome Measures Reported in Clinical Trials of Achilles Tendinopathy. *Sports medicine (Auckland, NZ)* 2021 doi: 10.1007/s40279-021-01588-6
17. Robinson JM, Cook JL, Purdam C, et al. The VISA-A questionnaire: a valid and reliable index of the clinical severity of Achilles tendinopathy. *British journal of sports medicine* 2001;35(5):335-41. [published Online First: 2001/10/02]
18. Murphy MC. Exercise rehabilitation for mid-portion Achilles tendinopathy: a critique of evidence and assumptions (PhD Academy Award). *British journal of sports medicine* 2022 doi: 10.1136/bjsports-2022-106024 [published Online First: 20221124]

19. Comins J, Siersma V, Couppe C, et al. Assessment of content validity and psychometric properties of VISA-A for Achilles tendinopathy. *PloS one* 2021;16(3):e0247152. doi: 10.1371/journal.pone.0247152 [published Online First: 20210311]
20. Travers N, Murphy MC, Wand B, et al. Does the Victorian Institute of Sport Assessment – Achilles assess disability in people with Achilles tendinopathy: a Rasch Measurement Theory analysis. *Under Review* 2023
21. Murphy MC, Newsham-West RJ, Cook J, et al. TENDINopathy Severity assessment – Achilles (TENDINS-A): Development and Content Validity assessment of a new Patient-Reported Outcome Measure for Achilles Tendinopathy. . *The Journal of orthopaedic and sports physical therapy* 2023;Epub Ahead of Print
22. Rio EK, Rabusin C, Munteanu S, et al. Where is your pain? Pain location on loading is different to palpation, imaging and recall location: a study of Achilles tendinopathy research participants. . *The Journal of orthopaedic and sports physical therapy* 2023;Under Review
23. Sleeswijk Visser TSO, van Es EM, Meuffels DE, et al. Standardized pain mapping for diagnosing Achilles tendinopathy. *Journal of science and medicine in sport* 2022;25(3):204-08. doi: 10.1016/j.jsams.2021.10.010 [published Online First: 20211021]
24. Tavakkoli Oskouei S, Malliaras P, Hill KD, et al. Evaluating daily physical activity and biomechanical measures using wearable technology in people with Achilles tendinopathy: A descriptive exploratory study. *Musculoskeletal science & practice* 2022;58:102534. doi: 10.1016/j.msksp.2022.102534 [published Online First: 20220222]
25. Golightly YM, Devellis RF, Nelson AE, et al. Psychometric properties of the foot and ankle outcome score in a community-based study of adults with and without osteoarthritis. *Arthritis Care Res (Hoboken)* 2014;66(3):395-403. doi: 10.1002/acr.22162
26. MacCallum RC, Widaman KF, Zhang S, et al. Sample size in factor analysis. *Psychological Methods* 1999;4(1):84-99. doi: 10.1037/1082-989X.4.1.84
27. Child D. The essentials of factor analysis: A&C Black 2006.
28. Kaiser HF. An index of factorial simplicity. *Psychometrika* 1974;39(1):31-36. doi: 10.1007/BF02291575
29. Bartlett MS. A Note on the Multiplying Factors for Various χ² Approximations. *Journal of the Royal Statistical Society Series B (Methodological)* 1954;16(2):296-98.
30. Stevens JP. Applied multivariate statistics for the social sciences: Routledge 2012.

31. Psychometric theory : Nunnally and Bernstein, 3rd edition; 2007.
32. Bagozzi RP, Yi Y. On the evaluation of structural equation models. *Journal of the Academy of Marketing Science* 1988;16(1):74-94. doi: 10.1007/BF02723327
33. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737-45. doi: 10.1016/j.jclinepi.2010.02.006
34. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 2016;15(2):155-63. doi: 10.1016/j.jcm.2016.02.012 [published Online First: 20160331]
35. de Vet HCW, Terwee CB, Mokkink LB, et al. Measurement in Medicine: A Practical Guide. Cambridge: Cambridge University Press 2011.
36. Mouelhi Y, Jouve E, Castelli C, et al. How is the minimal clinically important difference established in health-related quality of life instruments? Review of anchors and methods. *Health Qual Life Outcomes* 2020;18(1):136. doi: 10.1186/s12955-020-01344-w [published Online First: 20200512]
37. Challoumas D, Zouvani A, Creavin K, et al. Determining minimal important differences for patient-reported outcome measures in shoulder, lateral elbow, patellar and Achilles tendinopathies using distribution-based methods. *BMC Musculoskelet Disord* 2023;24(1):158. doi: 10.1186/s12891-023-06261-9 [published Online First: 20230302]
38. Lagas IF, van der Vlist AC, van Oosterom RF, et al. Victorian Institute of Sport Assessment-Achilles (VISA-A) Questionnaire-Minimal Clinically Important Difference for Active People With Midportion Achilles Tendinopathy: A Prospective Cohort Study. *The Journal of orthopaedic and sports physical therapy* 2021;51(10):510-16. doi: 10.2519/jospt.2021.10040
39. McCormack J, Underwood F, Slaven E, et al. The minimum clinically important difference on the VISA-Aand LEFS for patients with insertional Achilles tendinopathy. *Int J Sports Phys Ther* 2015;10(5):639-44.
40. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne* 2006;174(6):801-9. doi: 10.1503/cmaj.051351

41. Tu X, Tu Z, Lin W, et al. The Victorian Sports Assessment Institute-Achilles Tendinopathy Questionnaire (VISA-A): Chinese cross-cultural adaptation and psychometric validation. *Health Qual Life Outcomes* 2022;20(1):111. doi: 10.1186/s12955-022-02025-6 [published Online First: 20220723]

42. Bahari M, Hadadi M, Vosoughi AR, et al. Cross-cultural adaptation, reliability and validity of the Persian version of the Victorian Institute of Sport Assessment-Achilles questionnaire (VISA-A). *Disability and rehabilitation* 2022;44(6):983-91. doi: 10.1080/09638288.2020.1781268 [published Online First: 20200714]

43. Rio EK, Mc Auliffe S, Kuipers I, et al. ICON PART-T 2019–International Scientific Tendinopathy Symposium Consensus: recommended standards for reporting participant characteristics in tendinopathy research (PART-T). *British journal of sports medicine* 2019;54(11):627-30. doi: 10.1136/bjsports-2019-100957

Figure 1. Factor Loading within Confirmatory Factor Analysis

Legend: Small circles to the left of the diagram (e.g., E11) represent error related to each item. Medium sized circles (e.g., Item 11) represent scores for each TENDINS-A item (known as indicators). Large circles to the right of the diagram (e.g., PAIN) represent factors causing disability (known as unobserved or latent variables). Arrows linking factors and items represent the standardised factor loading for each item. Arrows linking factors represent covariance between factors.