

The medial tibial stress syndrome score: a new patient-reported outcome measure

Marinus Winters,¹ Maarten H Moen,^{2,3} Wessel O Zimmermann,^{4,5}
Robert Lindeboom,⁶ Adam Weir,⁷ Frank JG Backx,¹ Eric WP Bakker⁶

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bjsports-2015-095060>).

¹Department of Rehabilitation, Nursing Science & Sports, University Medical Centre Utrecht, Utrecht, The Netherlands

²Bergman Clinics, Naarden, The Netherlands

³The Sports Physician Group, St Lucas Andreas Hospital, Amsterdam, The Netherlands

⁴Department of Training Medicine and Training Physiology, Royal Netherlands Army, Utrecht, The Netherlands

⁵Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA

⁶Division of Clinical Methods and Public Health, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands

⁷Aspetar Orthopedic and Sports Medicine Hospital, Doha, Qatar

Correspondence to

Marinus Winters, Department of Rehabilitation, Nursing Science & Sports, University Medical Centre Utrecht, P.O. Box 85500, Utrecht 3508 GA, The Netherlands; marinuswinters@hotmail.com

Accepted 1 October 2015

ABSTRACT

Background At present, there is no validated patient-reported outcome measure (PROM) for patients with medial tibial stress syndrome (MTSS).

Aim Our aim was to select and validate previously generated items and create a valid, reliable and responsive PROM for patients with MTSS: the MTSS score.

Methods A prospective cohort study was performed in multiple sports medicine, physiotherapy and military facilities in the Netherlands. Participants with MTSS filled out the previously generated items for the MTSS score on 3 occasions. From previously generated items, we selected the best items. We assessed the MTSS score for its validity, reliability and responsiveness.

Results The MTSS score was filled out by 133 participants with MTSS. Factor analysis showed the MTSS score to exhibit a single-factor structure with acceptable internal consistency ($\alpha=0.58$) and good test-retest reliability (intraclass correlation coefficient=0.81). The MTSS score ranges from 0 to 10 points. The smallest detectable change in our sample was 0.69 at the group level and 4.80 at the individual level. Construct validity analysis showed significant moderate-to-large correlations ($r=0.34-0.52$, $p<0.01$). Responsiveness of the MTSS score was confirmed by a significant relation with the global perceived effect scale ($\beta=-0.288$, $R^2=0.21$, $p<0.001$).

Conclusions The MTSS score is a valid, reliable and responsive PROM to measure the severity of MTSS. It is designed to evaluate treatment outcomes in clinical studies.

INTRODUCTION

The medial tibial stress syndrome (MTSS) is one of the most common exercise-induced leg injuries among running and jumping athletes and military personnel.¹ It is defined as exercise-induced pain along the posteromedial border of the tibia, and when pain is additionally provoked by palpation over five or more consecutive centimetres.²

A recent systematic review showed that there is no conclusive evidence for any effective intervention in the management of MTSS.³ The absence of a specific outcome measure for patients with MTSS disables a valid measurement of injury severity and intervention effects. Studies investigating the effects of interventions in participants with MTSS have used a wide range of outcome measures to quantify their results, for example, time to recovery, visual analogue scales, Likert scale and numeric rating scale.⁴⁻⁶ Differing definitions for the same outcome

measure such as 'time to recovery' are often used.^{6,7}

A standardised assessment instrument that enables a valid and reliable assessment of treatment effects in patients with MTSS is needed.³ The patient's perspective has become increasingly important in the context of determining treatment effects.⁸ Patient-reported outcome measures (PROMs) are recommended to evaluate effectiveness in clinical settings and randomised controlled trials.⁹ Recently, items for a new PROM for patients with MTSS were generated using a Delphi procedure.¹⁰ The objective of this study was to test the methodological properties of these items, select the best ones to form the MTSS score, and assess the MTSS score's validity, reliability and responsiveness.

METHODS

Design and objective

A prospective cohort design was used to select the best items for the MTSS score and to assess its validity, reliability and responsiveness. We followed the consensus-based standards for selection of health measurement instruments (COSMIN) guidelines while validating the MTSS score.¹¹

Participants

Between 1 January 2013 and 1 January 2015, 13 healthcare centres (including 5 sports medicine facilities, 1 military medical centre, 5 sports physiotherapy practices and 2 military physiotherapy centres) in The Netherlands assessed possible eligible participants for study participation. Sports physicians and sports physiotherapists working in the participating facilities assessed potential candidates by applying our inclusion and exclusion criteria. Participants (≥ 16 year) with MTSS for at least 3 weeks were considered eligible for inclusion. MTSS was defined as activity-related pain along the posteromedial tibial border and tenderness on the same site over a length of at least five or more consecutive centimetres.² Participants were excluded when a history of tibial fracture, clinical suspicion of chronic compartment syndrome or stress fracture was present, or when coexisting injuries were present.¹² Participants with concurrent lower extremity symptoms and participants with spoken or written Dutch language comprehension difficulty were excluded. Participants who met the inclusion criteria were informed about the study purpose and participated after signing informed consent. The medical ethics committees of Zuid-West Holland (12-092) and Utrecht (12-542/C), The Netherlands, provided approval before the study's initiation.

To cite: Winters M, Moen MH, Zimmermann WO, et al. *Br J Sports Med* Published Online First: [please include Day Month Year] doi:10.1136/bjsports-2015-095060

Procedure

Participants were asked to fill out questionnaires on three occasions. At baseline (T1), participants were asked to fill out a form relating demographic information, preliminary items of the MTSS score, the RAND 36-item Health Survey and to answer questions relating to their sports activities. After 1 week (T2), the primary investigator (MW) contacted participants by telephone and requested them to fill out the preliminary items of the MTSS score again in an online environment. The final measurement was administered at 3 months (T3). Participants were approached by telephone to fill out the MTSS score's preliminary items, a global perceived effect (GPE) scale and to answer questions relating to their weekly sports activities in an online environment. During the study, participants continued standard medical care at their facility. [Figure 1](#) shows the study flow and the administered measures for each occasion.

Measures

Items for the MTSS score

Experts developed items for the MTSS score by means of a Delphi study. These items were then appraised by a total of 20 patients with MTSS who did not participate in the validation study. We reported on the item generation process elsewhere.¹⁰ All items were generated in Dutch. In total, 15 items were generated, assessing limitations in sporting activities, pain while performing sporting activities, pain while performing activities of daily living (ADL) and pain at rest. Items have four response

options with descriptors for each response category. Higher item scores indicate a more severe pain or limitation and hence more severe MTSS symptoms. Participants were asked to fill out the MTSS score with their most painful shin in mind, in case of bilateral symptoms.

Items of the RAND 36-item Health Survey

We used items of the Dutch version of the RAND 36-item Health Survey for assessment of construct validity.¹³ The RAND-36 is widely used to measure a variety of domains, including pain and limitations while performing ADL, and also in musculoskeletal and sports medicine-related research.^{14–16} Of specific interest to this study were items 3G, 3H and 7. Item 3G measures the limitation while walking >1 km. Item 3H measures the limitation while walking 0.5 km. Low non-standardised scores indicate that the activity is more limited for both items. Item 7 of the RAND-36 evaluates the degree of pain in the past week, with higher non-standardised scores indicating less pain.

Transition scale

At T2, the transition scale assesses the perceived change since T1. Participants could indicate if their condition had improved, worsened or remained unchanged.¹¹ Those participants whose condition had remained unchanged were considered 'stable' participants.

GPE scale

The GPE scale assesses the participant's perceived condition at follow-up (T3) compared with T1; 'completely recovered', 'much improved', 'slightly improved', 'not changed', 'slightly worsened', 'much worsened' or 'worse than ever'.¹⁷

Change in intensity and volume of sporting activities

At baseline, participants indicated the number of hours they were able to perform sporting activities, and how much they had reduced their training volume since the onset of their MTSS symptoms. We labelled the difference as 'volume change in sporting activities in hours'. In addition, we asked to what degree the intensity of their exercise had changed since the onset of their symptoms ('severely diminished', 'diminished', 'my exercise intensity has remained unchanged', 'my exercise intensity increased', 'I am unable to perform any type of exercise due to my shin pain'). We labelled this as 'intensity change in sporting activities'.

Data analysis and statistics

All data were analysed with SPSS (V20.0, IBM SPSS Inc, Chicago, USA) by one author (MW). Missing data were handled by imputing item medians of the sample investigated for all analyses. Demographic data were presented with appropriate measures of central tendency and dispersion.

Preliminary data analysis and item reduction

We planned to reduce the item set to have one item for all relevant domains (limitations in sporting activities, pain while performing sporting activities, pain while performing ADL and pain at rest). We used the reliability and responsiveness analysis to identify the best items for the final version of the MTSS score.

We selected the best item for each domain:

- For limitation in sporting activities: item 'current sporting activities', 'current amount of sporting activities' or 'current content of sporting activities';



Figure 1 Flow diagram (GPE, global perceived effect; MTSS, medial tibial stress syndrome).

- ▶ For pain while performing sporting activities: item 'pain while performing sporting activities', 'time to onset of pain during sporting activities', 'pain throughout sporting activities 1', 'pain throughout sporting activities 2' or 'pain after sporting activities';
- ▶ For pain while performing ADL: item 'pain while standing', 'pain while walking', 'pain while walking up or downstairs' or 'pain while performing common daily activities';
- ▶ For pain at rest: item 'pain at rest', 'pain at night' or 'pain to touch'.

We used the following analyses to select the best items:

- ▶ Test-retest reliability as calculated with intraclass correlation coefficients (ICCs);
- ▶ Association between item change scores and the GPE scale.

Test-retest reliability

We used the data of stable participants, collected at T1 and T2, for evaluation of the MTSS score's items and subscale reliability. Test-retest reliability was assessed with a two-way random effects, consistency, single measures ICC for all items. ICCs were presented with their 95% CIs.¹⁸ ICC values of <0.50 were regarded as insufficient, ICCs between 0.50 and 0.75 were considered acceptable, and ICCs >0.75 were labelled as good.¹⁹

Item responsiveness

We used the data collected at T1 (MTSS score) and T3 (MTSS score and GPE scale) for this analysis. We assessed the relation between each item change score (independent variable) and the GPE scale (dependent variable) in a linear regression analysis. We calculated change scores for each item subtracting T3 from T1 for each item of the MTSS score. The β -coefficient and the R^2 expressed the direction and magnitude of the relation between each item and the GPE scale. These measures were used to select the best items for the MTSS score. We considered a p value <0.1 as a significant relation. We hypothesised a greater change to be negatively correlated with GPE (the lower the GPE score, the greater the improvement).

All items were discussed for relevancy and importance by four authors (MW, AW, MHM and EWPB) until consensus was reached on which items should be selected for the final MTSS score. However, when consensus could not be met, we voted for selection of an item. Items were selected when a majority of the authors (3/4) favoured selection. When no majority was reached, a fifth author (FJGB) made the decision.

Further methodological testing of the final MTSS score and statistics

We further assessed the remaining item set for its:

- ▶ Structural validity and internal consistency;
- ▶ Construct validity;
- ▶ Responsiveness of the total score;
- ▶ Test-retest reliability of the total score.

In addition, we calculated:

- ▶ Measurement error and smallest detectable change (SDC);
- ▶ Minimal important change.

We present a summary of item variation at T1 and T3 to further address the interpretability of the MTSS score.

Structural validity and internal consistency

To investigate the structural validity of the MTSS score, we ran a factor analysis on the MTSS score data collected at T1. We estimated the amount of common variance by estimating communality values for all variables using the maximum-likelihood method (MLM) with direct oblique rotation. MLM enables

generalisation of the results beyond the study's population. Direct oblique rotation assumes that underlying (latent) factors of the MTSS score are related.²⁰ Kaiser's criterion (eigenvalues ≥ 1) and a scree plot (point of inflexion) assisted in identifying relevant factors.^{21 22} Items with factor loadings of >0.4 were thought to be important for the factor being studied.²³ We checked the item-rest correlations for the items that were maintained in the MTSS score at T1. Item-rest correlations >0.3 were considered to measure the same construct. We addressed the internal consistency of the item set by calculating Cronbach's α (CA). We considered CA around 0.6 as acceptable, and above 0.75 as good.^{24 25}

Construct validity

We assessed the relationships between items of the MTSS score with three items of the RAND-36, and volume and intensity change in sporting activities, collected at T1. After the item selection process, we formulated a hypothesis for each item of the MTSS score. Spearman's Rank tests were used to assess correlations between items. We regarded correlation coefficients around 0.1 as small, around 0.3 as moderate and those around or above 0.5 as large.²⁶ We recoded item scores of items 3G and 3H (recoded: higher scores indicate more limitation) for this analysis.

Responsiveness of the MTSS score

To determine item responsiveness, we calculated the change in MTSS scores between T1 and T3 (ie, T1-T3). We performed a linear regression analysis with these change scores as the independent variable and the GPE as the dependent variable. The β -coefficient and the R^2 expressed the direction and magnitude of the relationship between the MTSS score and the GPE scale. We considered a p value <0.05 as a significant relationship. We hypothesised a greater change to be negatively correlated with GPE (the lower the GPE score, the greater the improvement).

Test-retest reliability, measurement error and SDC of the MTSS score.

We used the data of 'stable' participants, collected at T1 and T2, for evaluation of the MTSS score's reliability. Test-retest reliability of the total MTSS score was assessed in the same way as individual items. We expressed measurement error by the standard error of measurement (SEM). The SEM was calculated as $SEM = SD_{\text{measurement } 1+2} \sqrt{1 - ICC}$.¹⁸ The SDC was calculated at both the individual ($SEM \times 1.96 \times \sqrt{2}$) and group level ($SEM \times 1.96 \times \sqrt{2} / \sqrt{n}$).^{18 27}

Minimal important change

We used the data of those participants who indicated that their condition had 'slightly improved' or 'slightly worsened' on the GPE scale at T3. The same change scores were used here as in the responsiveness analysis. We considered the mean change score of those participants who indicated 'slightly improved' or 'slightly worsened' to be the minimal important change.

Interpretability

To enhance the interpretability of the MTSS score, we present the means, SDs and distributions of the MTSS score at T1 and T3. Floor or ceiling effects were considered to be present when 15% or more of the participants scored the lowest or highest possible MTSS score.^{11 28}

Cross-cultural translation

We translated all items of the preliminary MTSS score into English. This translation process contained a forward and

backward translation. As for item generation, we report on the cross-cultural translation process elsewhere.¹⁰ We present here the final (Dutch) MTSS score and its English cross-cultural translation.

Sample size

We calculated the required sample size for test–retest reliability analysis and exploratory factor analysis, before the study's start. For test–retest reliability, a sample size of 51 stable participants was required, as well as constructing a two-sided 95% CI and assuming an ICC of 0.80 with a lower limit of 0.70.²⁹ For exploratory factor analysis, a minimum of 100 participants is advised; however, others suggest including 10 participants for each item tested in the analysis.³⁰

RESULTS

Prospective cohort study

A total of 133 participants met the inclusion criteria and agreed to participate in this prospective cohort study. The study comprised 73 men and 60 women, the mean age was 24.2 (SD=7.9), and the mean body mass index was 23.0 (SD=3.0). Forty-six participants (35%) were military personnel and 87 (65%) were athletes. Eighty-two per cent of the participants had bilateral MTSS, and 18% had unilateral MTSS. Table 1 provides further demographic information on our participants.

All 133 participants completed the MTSS score, the RAND-36 and questions concerning their exercise volume and intensity at T1. Seventy participants completed the MTSS score at T2 (the median number of days post T1 was 9 (range 5–20)), of whom 48 were 'stable'. At T3, the MTSS score was completed by 66 individuals, whereas the GPE was completed by 63 participants (median number of days post T1 was 70 (range 44–120)).

Missing items

For items of the MTSS score, few data were missing: at T1 2%, at T2 1.25%, while at T3 no data were missing. At T1, 7.25% of the data of the three items of the RAND-36 were missing. A minority of the participants did not provide information on sports volume (5.6%) and sports intensity change (6.8%) at T1. No data were missing for the transition scale at T2 or the GPE scale at T3.

Preliminary data analysis and item selection

Test–retest reliability on item level

Forty-eight participants indicated that their symptoms had remained 'unchanged' at T2. We used their data, collected at T1 and T2, to estimate the two-way random effects, consistency, single measures ICCs for all items of the MTSS score. Table 2 provides ICC values for all preliminary items of the MTSS score. All ICCs were acceptable or good, except for items 'pain to touch', 'pain while performing common daily activities', 'pain throughout sporting activities 1' and 'pain throughout sporting activities 2'. These items exhibited low test–retest reliability (ICC<0.50).

Item responsiveness on item level

Change scores between T1 and T3 were calculated for all items of the MTSS score. The change score item 'pain at night' showed an inverse relation with the GPE scale at T3 and was therefore considered invalid. All other change score items showed a relation with the GPE scale at T3; however, this relationship was only significant for items 'pain while standing', 'pain while walking', 'current sporting activities', 'current

Table 1 Demographic information

Demographic variable	Participants (N=133)
Male/female, n	73 (55%)/60 (45%)
Age in years, mean±SD	24.2±7.9
Length in cm, mean±SD	177±10
Weight in kg, mean±SD	74±13
BMI in kg/m ² , mean±SD	23±3
Sports athletes/military personnel, n (%)	87 (65%)/46 (35%)
Sports category, n (%)	
Running	35 (26%)
Fitness	21 (16%)
Hockey	14 (11%)
Soccer	14 (11%)
Athletics (non-distance running)	7 (5%)
Volleyball	6 (4%)
Cycling	5 (4%)
Other	31 (23%)
Hours of exercise a week at T1, median with range (minimum–maximum)	4.0 (0–30)
Duration of symptoms in months, median with range (minimum–maximum)	18 (0.75–144)
Side of symptoms, n (%)	
Both legs	109 (82%)
Only left leg	11 (8%)
Only right leg	13 (10%)

BMI, body mass index; T1, baseline.

content of sporting activities', 'pain while performing sporting activities', 'time to onset of pain during sporting activities' and 'pain after sporting activities'.

Item selection

Limitation in sporting activities

The item 'current sporting activities' was selected for 'limitation in sporting activities'. The item 'current content of sporting activities' showed comparable test–retest reliability (ICC=0.80 vs 0.84) and association with the GPE scale ($\beta=-0.43$ vs -0.38); however, we considered the first to reflect this domain best.

Pain while performing sporting activities

The item 'pain while performing sporting activities' showed the best relation with the GPE scale and exhibited the best test–retest reliability (see table 2) and was therefore selected.

Pain while performing ADL

The item 'pain while walking' was selected for 'pain while performing ADL'. Although the items 'pain while standing' and 'pain while walking up or downstairs' were equally reliable and related to the GPE scale (see table 2), we considered walking more relevant and feasible than standing and walking up or downstairs. More specifically, standing and walking up or downstairs are activities that not all possible participants with MTSS would engage in on a daily basis. 'Pain while performing common daily activities' exhibited a low test–retest reliability (ICC=0.48), but one author considered this item the most relevant to measure this domain. Therefore, the steering committee further discussed item selection for this domain (see Steering committee section).

Table 2 Item selection for the MTSS score

Theoretical domain	Item	Test–retest reliability ICC (95% CI)	Responsiveness analysis*		
			β -coefficient	R ²	p Value
Limitation in sporting activities					
	<i>Current sporting activities</i>	0.80 (0.67 to 0.88)	−0.43	0.065	0.04
	Current amount of sporting activities	0.76 (0.61 to 0.86)	−0.03	0.001	0.78
	Current content of sporting activities	0.84 (0.73 to 0.91)	−0.38	0.114	<0.01
Pain while performing sporting activities					
	<i>Pain while performing sporting activities</i>	0.63 (0.43 to 0.78)	−0.45	0.129	<0.01
	Time to onset of pain during sporting activities	0.72 (0.56 to 0.84)	−0.44	0.201	<0.01
	Pain throughout sporting activities 1	0.44 (0.19 to 0.65)	−0.11	0.018	0.31
	Pain throughout sporting activities 2	0.34 (0.07 to 0.57)	−0.15	0.031	0.18
	Pain after sporting activities	0.74 (0.58 to 0.85)	−0.20	0.048	0.09
Pain while performing activities of daily life					
	Pain while standing	0.72 (0.55 to 0.84)	−0.55	0.098	0.01
	<i>Pain while walking</i>	0.82 (0.70 to 0.90)	−0.50	0.089	0.02
	Pain while walking up or downstairs	0.86 (0.76 to 0.92)	−0.18	0.008	0.48
	Pain while performing common daily activities	0.48 (0.23 to 0.67)	−0.52	0.107	<0.01
Pain at rest					
	<i>Pain at rest</i>	0.60 (0.39 to 0.76)	−0.21	0.019	0.28
	Pain at night	0.91 (0.85 to 0.95)	0.22	0.023	0.22
	Pain to touch	0.50 (0.26 to 0.69)	−0.19	0.059	0.06

Items in italics were selected for use in the MTSS score.

*As assessed with linear regression analysis. β indicates how the GPE scale changes for each extra unit of the item; R² represents the magnitude of the relation between the GPE scale and each item.

ADL, activities of daily living; GPE, global perceived effect; ICC, intraclass correlation coefficient; MTSS, medial tibial stress syndrome.

Pain at rest

The item ‘pain at rest’ was considered the best item for ‘pain at rest’. ‘Pain at night’ exhibited an inverse relation with the GPE scale ($\beta=0.22$) and was therefore considered invalid. The item ‘pain to touch’ exhibited a low test–retest reliability (ICC=0.50).

Steering committee

Selection was made on the basis of consensus for all items, except for ‘pain while performing activities of daily life’. On this domain, no consensus was reached; we voted for the item ‘pain while performing common daily activities’ or ‘pain while walking’. A majority (3/4 authors) voted for pain while walking.

Methodological testing of the final MTSS score

Structural validity and internal consistency analysis

Data collected at T1 from all 133 participants were used to assess the structural validity of the item set. One factor yielded an eigenvalue of ≥ 1 , explaining 44.4% of the variance in the item set. The scree plot confirmed the unidimensionality of the

item set. All items loaded on this factor satisfactorily (>0.4). We checked the item-rest correlation for each subscale. Item-rest correlations were adequate, $r \geq 0.3$. CA showed acceptable internal consistency, $\alpha=0.58$. Table 3 depicts all results of the factor and the internal consistency analyses.

Construct validity

We checked whether the remaining items of the MTSS score at T1 were associated with items of the RAND-36 and sports volume and intensity change.

We hypothesised that:

1. Item ‘current sporting activities’ would show a moderate-to-large positive correlation ($r=0.3-0.5$) with volume change in sporting activities.
A positive correlation of $r=0.34$ (95% CI 0.17 to 0.50, $p<0.01$) was found.
2. Item ‘pain while performing sporting activities’ would exhibit a moderate to large positive correlation with intensity change in sporting activities ($r=0.3-0.5$).
We found a positive correlation of $r=0.34$ (95% CI 0.17 to 0.50, $p<0.01$).

Table 3 Factor analysis and internal consistency analysis

Items	Factor analysis Factor loadings	Internal consistency analysis	
		Item-rest correlations	Cronbach's α
MTSS score			0.58
Item 1	Current sporting activities	0.40	0.3
Item 2	Pain while performing sporting activities	0.52	0.4
Item 3	Pain while walking	0.64	0.4
Item 4	Pain at rest	0.48	0.3

Factor analysis and internal consistency analysis of definitive items in the MTSS score.
MTSS, medial tibial stress syndrome.

Table 4 Interpretability; item variation of the MTSS score at T1 (N=133)

Item number	Item name	Answer option 1, n (%)	Answer option 2, n (%)	Answer option 3, n (%)	Answer option 4, n (%)	Mean	Median	Missing values n (%)
1	Current sporting activities	16 (12.0)	48 (36.1)	57 (42.9)	8 (6.0)	1.44	2	4 (3.0%)
2	Pain while performing sporting activities	4 (3.0)	65 (48.9)	52 (39.1)	8 (6.0)	1.50	1	4 (3.0%)
3	Pain while walking	41 (30.8)	71 (53.4)	20 (15.0)	1 (0.8)	0.86	1	0 (0.0%)
4	Pain at rest	48 (36.1)	69 (51.9)	14 (10.5)	2 (1.5)	0.77	1	0 (0.0%)

MTSS, medial tibial stress syndrome.

3. Item 'pain while walking' would show a moderate-to-large positive correlation ($r=0.3-0.5$) with items 3G and 3H (degree of limitation while walking >1 km and walking around 0.5 km, respectively).

A large positive correlation was found with items 3G ($r=0.58$, 95% CI 0.43 to 0.70, $p<0.01$) and 3H ($r=0.48$, 95% CI 0.32 to 0.63, $p<0.01$).

4. Item 'pain at rest' would show a moderate-to-large correlation ($r=0.3-0.5$) with item 7 (degree of pain in the past week) of the RAND.

Item 1 showed a large positive correlation ($r=0.53$, 95% CI 0.39 to 0.64, $p<0.01$).

Responsiveness of the MTSS score

A significant negative relation confirmed the responsiveness of the total MTSS score: $\beta=-0.288$, $R^2=0.21$, $t=-3.962$, $p<0.001$.

Test-retest reliability of the total MTSS score

The total MTSS score showed good test-retest reliability: $ICC=0.82$ (95% CI 0.70 to 0.89, $F=9.95$, $p<0.001$).

Measurement error, SDC and minimal important change

We assessed the measurement error by calculation of the SEM and the SDC at the group and individual patient level. The SEM was 1.73. The SDC on the individual level was 4.80. The SDC and the minimal important change at the group level were both 0.69. This means that the MTSS score can measure the minimal important change.

Interpretability

The MTSS score is provided in Dutch and English (cross-culturally translated version) and available online as supplementary material. In addition, [tables 4-6](#) provide information on scoring distributions, means and medians of the MTSS score at T1 and T3. We conclude that floor or ceiling effects are not present for the MTSS score at T1 and T3.

The lowest possible MTSS score is 0, indicating that no MTSS symptoms are present, whereas 10 is the maximum score. This indicates the highest severity of MTSS symptoms. In our study, the mean MTSS scores were 4.58 (± 1.88) and 3.72 (± 2.08) at T1 and T3, respectively.

DISCUSSION

This is the first study to assess a PROM for patients with MTSS for reliability, validity and responsiveness. We selected the best items from an item pool generated by a group of experts to be used in the final MTSS score. This new MTSS score is a simple four-item scale that addresses pain at rest, pain while performing ADL, limitations in sporting activities and pain while performing sporting activities. The MTSS score specifically measures pain experienced along the shin and limitations due to shin pain. Its items exhibit four response options with descriptors for the degree of shin pain and limitations. The variation in items, from low-demand activities (resting/walking) to high-demand activities (sports activities), also contributes to the specificity of this new instrument.

Rigorous clinimetric evaluation

A previously performed Delphi study supports the content validity of the MTSS score, as shown by consensus among a group of experts in the field of MTSS. In addition, those items were appraised by a patient panel and were found to be valid, readable and comprehensive.¹⁰ Structural analysis confirmed the unidimensionality of the MTSS score. In addition, the MTSS score showed good construct validity when compared with items of the RAND-36 and the participants' volume and intensity change in sporting activities. The MTSS score's overall scale reliability and responsiveness confirmed the suitability for its use in scientific research. Taken together, this study shows that the MTSS score is a valid, reliable and responsive PROM for the evaluation of the injury severity in patients with MTSS.

In addition to reliability, validity and responsiveness, low measurement error is important for the MTSS score's utility. We found quite a large SDC (4.8, almost 50% of the possible score

Table 5 Interpretability; item variation of the MTSS score at T3 (N=66)

Item number	Item name	Answer option 1, n (%)	Answer option 2, n (%)	Answer option 3, n (%)	Answer option 4, n (%)	Mean	Median	Missing values n (%)
1	Current sporting activities	13 (19.7)	25 (37.9)	24 (36.4)	4 (6.0)	1.29	1	0 (0.0%)
2	Pain during sporting activities	9 (13.6)	37 (56.1)	15 (22.7)	5 (7.6)	1.24	1	0 (0.0%)
3	Pain while walking	31 (47.0)	30 (45.4)	5 (7.6)	0 (0.0)	0.61	1	0 (0.0%)
4	Pain at rest	31 (47.0)	31 (47)	4 (6.0)	0 (0.0)	0.59	1	0 (0.0%)

MTSS, medial tibial stress syndrome.

Table 6 Interpretability; MTSS score at T1 (n=133), T3 (n=66) and MTSS change score (T1–T3, n=66)

	Mean	SD	95% CI	Median	Range	Absolute minimum	Absolute maximum	Floor effects, n (%)	Ceiling effects, n (%)
MTSS score at T1	4.58	1.88	4.26 to 4.90	5	1 to 10	0	10	2 (1.5)	1 (0.8)
MTSS score at T3	3.72	2.08	3.22 to 4.24	4	0 to 9	0	10	3 (2.3)	0 (0.0)
MTSS change score at T1–T3	1.00	1.56	0.62 to 1.38	1	–2 to 5	–10	10	0 (0.0)	0 (0.0)

MTSS, medial tibial stress syndrome.

range) at the individual level. However, analysis at the group level showed that the SDC was equal to the minimal important change (both 0.69 points). This suggests that the MTSS score is an appropriate measure to compare tendencies across different groups, such as in RCTs into the effectiveness of different interventions in the treatment of MTSS.

Another outcome measure for exercise-induced lower leg pain has been validated recently. This outcome measure aims to measure ‘functional impairment and limitation in sports ability’ in runners.³¹ In our opinion, the MTSS score is more valid and feasible for patients with MTSS. Most of the activities that can be scored in the outcome measure developed by Nauck *et al*³¹ may not be relevant to all patients (such as taking off and landing while jumping). In addition, our study suggests that pain at rest and ADL are important limitations to patients with MTSS and should therefore be part of an outcome assessment tool.

Clinical utility of the new MTSS score

Many of the patients in our study had a long duration of symptoms prior to enrolling in our study. This suggests that current interventions and routine care for MTSS are not very effective. The MTSS scores at T1 and T3, and GPE scale at T3, showed that little improvement was made after participants sought medical care in centres with a large clinical experience. This highlights the necessity for new approaches to treating MTSS. The MTSS score can be used in several ways to enhance better treatment outcomes. First, the MTSS score allows for determination of treatment effects as reported by the patient in contrast to determination of treatment effects by the assessor or by physical parameters. Second, the MTSS score is able to reliably and validly track changes in groups. This is predominantly important in randomised clinical trials. Finally, a possible future application could be if the MTSS score was able to predict a window for time to recovery (prognosis). We note that in a 2015 systematic review of risk factors for MTSS, there was no mention of certainty of the clinical diagnosis or any variation in severity of the condition.³² If adopted, our instrument will allow the broad condition of ‘MTSS’ to be sub-categorised according to level of severity of the condition. This instrument may be limited for monitoring individual patients with MTSS.

Strengths and limitations

A strength of the present study is the inclusion of a broad variety of participants with MTSS, athletes and military personnel with short-standing and long-standing symptoms. This strengthens the study’s external validity. The MTSS score is a practical outcome measure; the patient can fill out the MTSS score without any help from a physician or physiotherapist, and it takes little time for the patient to do so.

Our study also has limitations. First, we followed the classical test theory for all analyses, whereas the item response theory

would have been more appropriate. Item response theory analyses, however, require large sample sizes, up to 200–500 participants, depending on the type of analysis.²⁸ This was not possible within our network of healthcare providers and budget.

Another limitation is the sample size in relation to the number of statistical tests performed. We acknowledge that 18 tests is a large amount. Statistically, this may have introduced one significant result due to chance. Our methods were, however, in accordance with the COSMIN guidelines, a methods criterion in this field of research.¹¹

The MTSS score exhibits one factor (it is unidimensional) which explained 44% of the variance in the item set. Some would regard this as moderate or low. However, to the best of our knowledge, no hard cut-off values for when this value is sufficient exist in the field of clinimetrics. The MTSS score yielded a value similar to those of other PROMs successfully validated in the field of musculoskeletal pain.^{33–35}

We used the CA statistic to assess for internal consistency. The MTSS score’s CA was 0.58 and we considered this as acceptable. Other classification systems may rate this as moderate or poor.²⁸ Cortina³⁶ showed that a high number of items may inflate CA and a low number of items may deflate CA. Given the relatively low number of items in the MTSS score (N=4), we are confident that the internal consistency is acceptable, also given the sufficient item-rest correlations (all ≥ 0.3).

With regard to test–retest reliability, there are some methodological issues to address: first, 70 of the 133 participants filled out the MTSS score at T1 and T2. Although we attempted to contact all participants for the second measurement, we have not succeeded in reaching them all. It is unclear how this may have affected the test–retest reliability results exactly. However, we were still able to find sufficient test–retest reliability levels for all items of the MTSS score as well as for the overall MTSS score. Second, we used ICCs for categorical data instead of weighted κ . Among the many advantages of ICC over weighted κ , the most important ones are that ICC is able to deal with (the presence or absence of) various sources of error and with missing values.³⁷ Therefore, it is most likely that the MTSS score’s test–retest reliability is estimated more precisely with ICCs, and consequently, conclusions can be drawn more robustly. The direction and magnitude of the β -coefficient and R^2 of the linear regression analysis were used to select the most responsive items. In view of the moderate sample size used in this analysis (N=66), we set the threshold for significance to <0.1 to avoid missing true significant relations between the GPE and ‘MTSS change score’.³⁸ Finally, the cross-cultural English translation should be validated in English-speaking MTSS populations.

We conclude that the MTSS score is a valid, reliable and responsive PROM to evaluate injury severity in patients with MTSS. We recommend its use in studies of MTSS treatment.

What are the findings?

- ▶ The medial tibial stress syndrome (MTSS) score is a new patient-reported outcome measure that measures injury severity in a practical way.
- ▶ The MTSS score has been shown to be valid, reliable and responsive.
- ▶ The MTSS score can detect relevant group tendencies.

Acknowledgements The authors would like to thank all sports medicine physicians and sports physiotherapists who assisted with including patients for this study: Carl Barten, Sandra Chung, Jan-Willem Dijkstra, Frank Franke, Simon Goedegebuurne, Pieter Graber, Floor Groot, Nick van der Horst, Nienke Hulsman, Hilde Joosten, Wout van der Meulen, Robert Oosterom, Victor Steeneken, Karin Thys, Peter van Veldhoven, Joost Vollaard, Niels Wijne and Rahmon Zondervan.

Contributors EWPB and MHM conceived the idea for the study. MW drafted the manuscript. MW, MHM, RL, AW and EWPB were responsible for the study concept and design. MW, MHM, WOZ, FJGB and EWPB collected the data. MW, RL and EWPB were responsible for analysis and interpretation of the data. All the authors critically revised the manuscript.

Competing interests None declared.

Ethics approval The medical ethics committees of Zuid-West Holland (12-092) and Utrecht (12-542/C), The Netherlands, provided approval before the study's initiation.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- 1 Taunton JE, Ryan MB, Clement DB, *et al*. A retrospective case-control analysis of 2002 running injuries. *Br J Sports Med* 2002;36:95–101.
- 2 Yates B, White S. The incidence and risk factors in the development of medial tibial stress syndrome among naval recruits. *Am J Sports Med* 2004;32:772–80.
- 3 Winters M, Eskes M, Weir A, *et al*. Treatment of medial tibial stress syndrome: a systematic review. *Sports Med* 2013;43:1315–33.
- 4 Johnston E, Flynn T, Bean M, *et al*. A randomized controlled trial of a leg orthosis versus traditional treatment for soldiers with shin splints: a pilot study. *Mil Med* 2006;171:40–4.
- 5 Rompe JD, Cacchio A, Furia JP, *et al*. Low-energy extracorporeal shock wave therapy as a treatment for medial tibial stress syndrome. *Am J Sports Med* 2010;38:125–32.
- 6 Moen MH, Rayer S, Schipper M, *et al*. Shockwave treatment for medial tibial stress syndrome in athletes; a prospective controlled study. *Br J Sports Med* 2012;46:253–7.
- 7 Nissen LR, Astvad K, Madsen L. Low-energy laser treatment of medial tibial stress syndrome. *Ugeskr Laeger* 1994;156:7329–31.
- 8 World Health Organization. *Regional Office for Europe. Patient outcome measures in mental health: report on a WHO consensus meeting, Stockholm, 23–24 November 1995*. Copenhagen WHO Regional Office for Europe, 1996.
- 9 Davidson M, Keating J. Patient-reported outcome measures (PROMs): how should I interpret reports of measurement properties? A practical guide for clinicians and researchers who are not biostatisticians. *Br J Sports Med* 2014;48:792–6.
- 10 Winters M, Franklyn M, Moen MH, *et al*. The MTSS-score: item generation for a new patient reported outcome measure. [In press]
- 11 Mokkink LB, Terwee CB, Knol DL, *et al*. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010;10:22.
- 12 Edwards PH Jr, Wright ML, Hartman JF. A practical approach for the differential diagnosis of chronic leg pain in the athlete. *Am J Sports Med* 2005;33:1241–9.
- 13 van der Zee K, Sanderman R. *Het meten van de algemene gezondheidstoestand met de RAND-36. Een handleiding*. Groningen: NCG, 1993.
- 14 Thorborg K, Hölmich P, Christensen R, *et al*. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med* 2011;45:478–91.
- 15 Christensen CP, Althausen PL, Mittleman MA, *et al*. The Nonarthritic Hip Score: reliable and validated. *Clin Orthop Relat Res* 2003;406:75–83.
- 16 Martin RL, Kelly BT, Philippon MJ. Evidence of validity for the Hip Outcome Score. *Arthroscopy* 2006;22:1304–11.
- 17 Kamper SJ, Ostelo RW, Knol DL, *et al*. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol* 2010;63:760–6.
- 18 Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res* 2005;19:231–40.
- 19 Portney LG, Watkins MP. *Foundations of clinical research: applications to practice*. 3rd edn. Upper Saddle River, NJ: Pearson/Prentice Hall, 2009.
- 20 Costello AB, Osborne J. Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Pract Assess Res Eval* 2005;10:1–9.
- 21 Kaiser HF. The application of electric computers to factor analysis. *Educ Psychol Meas* 1960;20:141–51.
- 22 Cattell RB. *The scientific analysis of personality*. Chicago: Aldine, 1966.
- 23 Stevens JP. *Applied multivariate statistics for the social sciences*. 4th edn. Hillsdale, New York: Erlbaum, 2002.
- 24 Davis FB. *Educational measurements and their interpretation*. Belmont, California, USA: Wadsworth, 1964.
- 25 Nunnally JC. *Psychometric theory*. 1st edn. New York: McGraw-Hill, 1967.
- 26 Cohen J. *Statistical power analysis for the behavioural sciences*. 2nd edn. New York: Academic Press, 1988.
- 27 de Vet HC, Terwee CB, Knol DL, *et al*. When to use agreement versus reliability measures. *J Clin Epidemiol* 2006;59:1033–9.
- 28 Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use*. 4th edn. New York: Oxford University Press, 2008.
- 29 Bonett DG. Sample size requirement for estimating intraclass correlations with desired precision. *Stat Med* 2002;21:1331–5.
- 30 Field A. *Discovering Statistics. Using SPSS*. 3rd edn. London: Sage Publications Inc, 2009.
- 31 Nauck T, Lohrer H, Padhiar N, *et al*. Development and validation of a questionnaire to measure the severity of functional limitations and reduction of sports ability in German-speaking patients with exercise-induced leg pain. *Br J Sports Med* 2015;49:113–17.
- 32 Hamstra-Wright KL, Bliven KC, Bay C. Risk factors for medial tibial stress syndrome in physically active individuals such as runners and military personnel: a systematic review and meta-analysis. *Br J Sports Med* 2015;49:362–9.
- 33 Waddell G, Newton M, Henderson I, *et al*. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. *Pain* 1993;52:157–68.
- 34 Wheeler AH, Goolkasian P, Baird AC, *et al*. Development of the Neck Pain and Disability Scale. Item analysis, face, and criterion-related validity. *Spine (Phila Pa 1976)* 1999;24:1290–4.
- 35 Roos EM, Roos HP, Ekdahl C, *et al*. Knee injury and Osteoarthritis Outcome Score (KOOS)—validation of a Swedish version. *Scand J Med Sci Sports* 1998;8:439–48.
- 36 Cortina JM. What is coefficient alpha? An examination of theory and applications. *J Appl Psychol* 1993;78:98–104.
- 37 Berk RA. Generalizability of behavioral observations: a clarification of interobserver agreement and interobserver reliability. *Am J Ment Defic* 1979;83:460–72.
- 38 Altman DG. *Practical statistics for medical research*. 1st edn. London: Chapman & Hall, 1999.



The medial tibial stress syndrome score: a new patient-reported outcome measure

Marinus Winters, Maarten H Moen, Wessel O Zimmermann, Robert Lindeboom, Adam Weir, Frank JG Backx and Eric WP Bakker

Br J Sports Med published online October 28, 2015

Updated information and services can be found at:

<http://bjsm.bmj.com/content/early/2015/10/28/bjsports-2015-095060>

These include:

Supplementary Material

Supplementary material can be found at:

<http://bjsm.bmj.com/content/suppl/2015/10/28/bjsports-2015-095060.DC1.html>

References

This article cites 26 articles, 9 of which you can access for free at:

<http://bjsm.bmj.com/content/early/2015/10/28/bjsports-2015-095060#BIBL>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

[Statistics and research methods](#) (50)

[Physiotherapy](#) (157)

[Physiotherapy](#) (212)

Notes

To request permissions go to:

<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:

<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:

<http://group.bmj.com/subscribe/>