Minimally Important Change and Smallest Detectable Change of the Oslo Sports Trauma Research Center Questionnaire in (Half) Marathon Runners

Master thesis

Physiotherapy Science

Program in Clinical Health Sciences

Utrecht University

Name student:	T.P.C. (Thierry) Franke
Student number:	3181057
Date:	30-06-2017
Internship supervisor(s):	B.M.A. (Bionka) Huisstede PhD
Internship institute:	Department of Rehabilitation, Physical Therapy Science & Sport, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands
Lecturer/supervisor University:	J. (Janjaap) van der Net PhD

"ONDERGETEKENDE

Thierry Pierre Chretiên Franke

bevestigt hierbij dat de onderhavige verhandeling mag worden geraadpleegd en vrij mag worden gefotokopieerd. Bij het citeren moet steeds de titel en de auteur van de verhandeling worden vermeld."

Examiner

Dr. M.F. Pisters

Assessors:

Dr. B.M.A. Huisstede

Dr. M.F. Pisters

Masterthesis, Physical Therapy Sciences, Program in Clinical Health Sciences, Utrecht University, Utrecht, 2017

ABSTRACT

Background The Oslo Sports Trauma Research Center Questionnaire (OSTRC) severity score (sum of the four OSTRC question answer scores) measures the impact of a running-related injury (RRI). The smallest detectable change (SDC) and minimally important change (MIC) of the OSTRC severity score are currently unknown.

Aim To evaluate the SDC and MIC of the OSTRC severity score in injured (half) marathon runners.

Methods Data from a prospective cohort study, the SUMMUM-2017 study, was used. A 133 runners who reported the same RRI on two OSTRCs two weeks apart, and filled in the global rating of limitations (GRL) and global rating of change (GRC) anchors were analyzed. Using the anchors, runners were classified as truly improved, unchanged, or truly worsened. SDC values were calculated at a group and individual level. MIC values were evaluated using the visual anchor-based MIC distribution and mean change method.

Results SDC values at a group level and at individual level were ≤0.27 and ≤5.91 respectively. The MIC was calculated using two methods. Visual anchor-based MIC distribution method: MIC for RRIs that truly improved on the GRC and GRL anchors were 13.50 and 18.50 respectively. Mean change method: MIC for RRIs that truly improved on the GRC and GRL anchors were 15.49 (95% CI 7.48-23.49) and 45.38 (95% CI 27.45-63.30) respectively. For RRIs that truly worsened a MIC value was only calculated using the mean change method with the GRL anchor 27.71 (95% CI -64.98-9.55).

Conclusion The OSTRC severity score MIC can distinguish between RRIs that truly improve and RRIs that were unchanged over a two-week period. The SDC was smaller than the smallest difference between any of the answer scores of the OSTRC questions. Furthermore, answers on the OSTRC questions need to improve by at least three options for the OSTRC severity change score to exceed the MIC value of 18.50.

Implications of key findings The MIC of 18.50 can be used in the process of monitoring a runner and help to determine if the RRI status truly improved or not. Furthermore, future studies could use the MIC values for sample size calculations.

Keywords: running; running-related injuries; Oslo Sports Trauma Research Center (OSTRC); Minimally important change; Smallest detectable change

INTRODUCTION

Running as a sport has gained popularity.[1] In the Netherlands, more than two million people participated in running as a sport in 2014.[2] These runners had 710.000 running-related injuries (RRI) which resulted in 2.9 million euros in direct medical costs, and 5.4 million euros in costs due to work absenteeism.[2]

In 2014 a consensus statement was published on injury surveillance in runners.[3] In this statement it was advised to repeatedly measure a runners injury status over time in order to detect RRIs that do not cause time-loss from running, but do lead to a reduction in training intensity, duration, or cause pain during running.[3] Registering these RRIs is possible using the Oslo Sports Trauma Research Centre Questionnaire (OSTRC).[4]

The OSRTC is an easy to perform questionnaire, consisting of four multiple choice questions, regarding the amount of difficulties the runner experienced while running due to their RRI (Table 1).[4] The OSTRC intends to measure and monitor the severity of a RRI.[4] The severity of a RRI is monitored using the OSTRC severity score (sum of the four answer scores, range 0 to 100, higher scores indicates a higher severity).[4] The OSTRC severity score reflects the self-reported assessment of the pain and impact of the runners RRI on their participation, training volume, and sports performance.[4,5] Several studies have shown that the OSTRC can be used to register and monitor both overload and acute onset injuries in different types of sports.[5–7]

Thus, changes in the OSTRC severity score should reflect the actual changes in the runner's participation, training volume, and sports performance. However, to interpret the change scores of the OSTRC severity score, and know if it has truly changed or not, it is necessary to know it's smallest detectable change (SDC) and minimally important change (MIC).[8] The SDC is the smallest change in the OSTRC severity score which can be considered as true change.[8] Whereas, the MIC is the smallest change in the OSTRC severity score which the runner perceives as important.[9]

For example, if a RRI has been registered twice and the OSTRC severity score improved or worsened the runner may try to adapt the training load appropriately. In this manner changes in the OSTRC severity score can be used to influence the decision-making process of runners, trainers, and clinicians. Hence, it is important to know if the OSTRC severity change score has truly changed or not – i.e. if the OSTRC severity change score exceeds the MIC. Therefore, the purpose of this study was to evaluate the SDC and MIC of the OSTRC severity score in injured (half) marathon runners.

Table 1 Oslo Sports Trauma Research Center Questionnaire
Per example specified here for knee problems

Part 1: Knee Problems		
Please answer all questions regardless of whether or not you have problems with your knee. Select the alternative that is most appropriate for you, and in the case that you are unsure, try to give an answer as best you can anyway. The term “knee problems” refers to pain, ache, stiffness, swelling, instability/giving-way, locking or other complaints related to one or both knees.		
Question 1		Score
Have you had any difficulties participating in normal training and competition due to knee problems the past two weeks?		
○	Full participation without knee problems	0
○	Full participation, but with knee problems	8
○	Reduced participation due to knee problems	17
○	Cannot participate due to knee problems	25
Question 2		
To what extent have you reduced your training volume due to knee problems during the past two weeks?		
○	No reduction	0
○	To a minor extent	6
○	To a moderate extent	13
○	To a major extent	19
○	Cannot participate at all	25
Question 3		
To what extent have knee problems affected your performance during the past two weeks?		
○	No effect	0
○	To a minor extent	6
○	To a moderate extent	13
○	To a major extent	19
○	Cannot participate at all	25
Question 4		
To what extent have you experienced knee pain related to your sport during the past two weeks?		
○	No pain	0
○	Mild pain	8
○	Moderate pain	17
○	Severe pain	25

METHODS

Design

Data from a prospective cohort study, the SUcces Measurement and Monitoring Utrecht Marathon (SUMMUM) 2017 study, was used to determine the SDC and MIC of the OSTRC. The study was approved by the University Medical Center Utrecht ethics committee (protocol number 16/503).

Participants

Runners participating in the Utrecht Science Park Marathon (USPM) (half) marathon (March 19th 2017) were recruited from September 1st 2016 forward. Recruitment ended on March 19th 2017. Runners were recruited during subscription for the USPM, via a newsletter, and during a symposium on RRIs. All half and whole marathon runners who subscribed for the USPM were asked if they were interested in participating in the SUMMUM-study. If the runner was interested an information letter was sent. Informed consent was provided by the runners before filling in the baseline questionnaire. Runners were eligible for inclusion when they were 1) 18 years or older; 2) in possession of an e-mail address; 3) possess sufficient Dutch language skills.

Procedures

Data collection started on November 25th 2016 (group 1 16-weeks before the USPM), when the first group of runners was invited to fill-in the first questionnaire. From then on, a new group of runners was invited to fill-in the first questionnaire every four weeks, up to the date of the USPM (December 22nd 2016, group 2 12 weeks before the USPM; January 20th 2017, group 3 8 weeks before the USPM; February 17th 2017, group 4 4 weeks before the USPM; and March 20th 2017, group 5 the day after the USPM).

The first questionnaire runners were invited to fill-in was the baseline questionnaire. Runners were provided seven days to complete the questionnaire. From there on, every two-weeks questionnaires were sent to the runners up to the date of the USPM. Runners were asked to complete the questionnaires within five days. During the week after the USPM runners were requested to fill in the post-marathon questionnaire regarding their participation in the USPM. The fifth group only received the baseline and post-marathon questionnaire. Every two to three days reminders were sent to all runners who had not started or completed the questionnaires. All questionnaires were sent via e-mail using NetQ (NetQuestionnaires, NetQ Healthcare B.V., Amsterdam, The Netherlands).

Questionnaires

Baseline Questionnaire

Using the baseline questionnaire each runner’s descriptive characteristics such as age, gender, weight, length, and the anatomical location, type, and duration of the RRI were assessed.

OSTRC

To monitor the (non-)injury status of all runners the Dutch version of the OSTRC was used (Table 1).[4,6] The OSTRC consists of four questions, for which the answer scores are summed to calculate the OSTRC severity score (range 0-100, a higher score indicates a higher severity).[4] If the OSTRC severity score was greater than zero a runner was considered as having a RRI. If a runner reported a RRI, follow-up questions were posed regarding the RRIs anatomical location, type of RRI, duration of the RRI, and the anchor questionnaires. If the same RRI was registered twice within two sequential questionnaires, i.e. within two weeks, a change score was calculated for the OSTRC severity score by subtracting the severity score of the second OSTRC from the first one.

Anchor Questionnaires

To evaluate the MIC and SDC of the OSTRC severity score external criteria were used to determine whether runners had changed over time. Therefore, two anchor questionnaires were added: the global rating of change (GRC) and global rating of limitations (GRL). If a runner reported no RRI, i.e. the OSTRC severity score was zero, the runner was asked if they had registered a RRI two to four weeks ago. If a RRI was reported two to four weeks ago the anchor questions were also filled-in.

The GRC, a retrospective anchor, regarded the changes in the RRI status during the last two weeks compared to when the runner first perceived the RRI. The GRC had seven possible answers "extremely worse" (1), "much worse" (2), "little worse" (3), "unchanged" (4), "slightly improved" (5), "much improved" (6), and "very much improved" (7). Because the GRC inherently contains the perceived change in the RRI status, it was not necessary to calculate a change score. Runners were classified to be truly improved if they answered "much improved" (6) or "very much improved" (7).[10] Runners answering "extremely worse" (1) or "much worse" (2) were considered as truly worsened.[10] Runners answering "little worse" (3), "unchanged" (4), or "slightly improved" (5) were also considered as unchanged to avoid socially desirable answers and to ensure that the measured change was clinically important.[10]

The GRL was used as a five-point prospective anchor measuring the runner's global rating of limitations with running performance due to their RRI. Possible answers (points) were "poor" (1), "fair" (2), "moderate" (3), "good" (4), or "excellent" (5). A change score was calculated by subtracting the second GRL score from the first one. A runner was considered as truly improved or worsened if the GRL score changed ≥2 points.[10]

STATISTICAL ANALYSIS

All data were analysed using SPSS (SPSS version 21, IBM, Armonk, New York, USA.) and Excel (Version 15.32.2 Microsoft, Redmond, Washington, USA). Baseline characteristics were described using descriptive statistics. To test if there were any significant differences between the half and whole marathon runners for the baseline characteristics a Chi-squared test (categorical variables) or T-test (continuous variables) was used. For continuous variables if the assumptions of normality were not met than the non-parametric Mann-Whitney U-test was used.

Runners were included in the SDC and MIC calculation if 1) on two separate sequential OSTRC's the same RRI was registered (RRIs were considered to be the same if the anatomical locations match and the RRI duration was ≥2 weeks, or if the OSTRC severity score was zero and a RRI was reported two weeks ago on the previous OSTRC); 2) the GRL and GRC anchor questions were filled in.

A-priori α was set at 0.05. Since no clear guidelines exist for sample size calculations for studies determining MIC values the recommendations of the COSMIN checklist were followed.[11] Therefore, the desired sample size was 100 runners.

Smallest Detectable Change

The SDC is the smallest change in scores on the OSTRC that can be considered as a true change, and not as a measurement error.[8] SDC values need to be calculated in a stable sample, i.e. no change in RRI status. Therefore, the SDC was calculated for runners with a GRL change score of zero, and for runners with a GRL change score of minus one, zero, or one. The SDC was also calculated for runners with a GRL change score of minus one, zero, and one because their RRI status was considered not to have changed according to the aforementioned criterion for the GRL anchor. Furthermore, the SDC was calculated on a group level [SDC_{group}= (standard deviation [SD]/√n)/ √n] and individual level [$SDC_{individual}$= 1.96 x √2*(SD/√n)].[9]

Minimally Important Change

The MIC is the smallest change in the OSTRC severity score which runners perceive as important.[9] Because MIC values can vary according to the manner by which they are calculated, two anchor based methods were used: the visual anchor-based MIC distribution method and the mean change method for both RRIs which improve or worsen separately.[9,12] To monitor changes in self-reported assessment of the pain and impact of the RRI on running participation, training volume, and sports performance the SDC of the OSTRC severity score should be smaller than the MIC.[13]

Visual Anchor-Based MIC Distribution Method

To calculate the optimal MIC value and the area-under the curve (AUC) (95% confidence interval [CI]) for both runners who truly improved or worsened separately according to both anchors ROC curves were plotted. In ROC curves the sensitivity and 1-specificity are plotted on the y-axis and x-axis respectively. The optimal MIC value was the point on the ROC curve where the sum of [1-sensitivity] and [1-specificity] was the smallest, yielding the smallest amount of misclassification.[9] To reflect the uncertainty of the MIC estimation a 95% CI upper limit was calculated [mean change+1.645*SD_{change}], using the runners who were unchanged according to the anchors.[13] The AUC reflects the capability of the OSTRC severity score to correctly identify injured runners who have truly changed.[9] An AUC value

>0.70, with a 95% CI lower limit >0.50, is considered good discriminatory.[14]

Furthermore, to visualise the distribution of the OSTRC severity change scores two graphs were plotted for runners who were classified as truly improved, truly worsened, or unchanged according to both anchors separately. On a two-sided vertical graph, the OSTRC severity change scores and the proportional frequencies (number of runners with a specific OSTRC severity change score divided by the total number of runners) were plotted on the y-axis and x-axis respectively. Proportional frequencies of runners who truly improved or worsened according to the anchor questionnaires and runners who were unchanged were plotted on the left and right side respectively. The MIC values (95% CI upper limits) were plotted in both graphs.

Mean change method

The MIC was calculated as the mean change score ± SD (95% CI) of the runners who classified as truly improved or truly worsened on both anchors. To prevent overestimation of the MIC, runners who reported being ‘completely recovered', or had a GRL change score >2 were excluded.[10]

Suitability of the Anchors

To determine the suitability of the anchors (i.e. if the anchors were appropriate to measure the same change in RRI status as the OSTRC severity change score), Spearman’s correlation was calculated between the OSTRC severity change score, and the GRL anchor change score, or the GRC anchor. The anchor was considered as suitable if $r \geq 0.30$.[9]

Dependency of the MIC on the Initial OSTRC Severity Score

To determine if the MIC values were dependent on the initial OSTRC severity score a sensitivity analysis was performed. Runners who reported being truly improved on the GRC anchor were divided into two subgroups using the median initial OSTRC severity score.[10] The OSTRC severity change scores of these subgroups were compared using a T-test.

RESULTS

Sample Characteristics

A total of 1084 runners were invited to participate in this study, 618 runners provided informed consent (Figure 1). The baseline characteristics were completed by 573 runners, after which 133 runners were included in the statistical analysis (group 1 N=66, group 2 N=21, group 3 N=28, group 4 N=18, group 5 N=0). These 133 runners reported the same RRI on at least two OSTRC questionnaires and filled-in the anchor questionnaires. The baseline characteristics of these 133 runners are shown in Table 2. Of these 133 runners, 105 and 28 runners participated in the half and whole USPM respectively. No significant differences were found between the half and whole marathon runners on gender (p=0.119), age (p>0.070), height (p=0.447), or weight (p>0.769). Anatomical locations of RRIs such as the knee (27%) and lower leg (14%) had the highest prevalence. The type of injuries which occurred most were overload (35%) and "muscle or tendon" (32%). In most cases, the RRI had been present for a duration of two to four weeks (33%), followed by longer than eight weeks (25%).

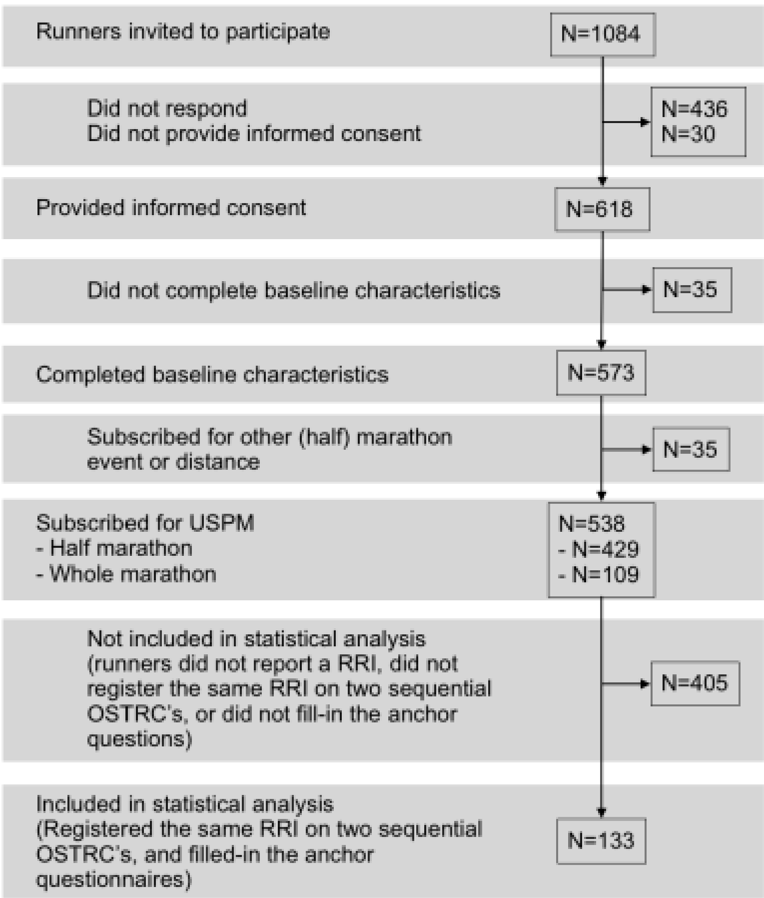

Figure 1 Flow diagram of the study participants
Abbreviations; OSTRC, Oslo Sports Trauma Research Center Questionnaire; RRI, running-related injury; USPM, Utrecht Science Park Marathon.

Table 2 Baseline Characteristics

		All runners (n=133)	Half marathon (n=105)	Marathon (n=28)	P-value*
Gender F/M (%)		60/73 (45/55)	51/54 (49/51)	9/19 (32/68)	0.121
Age (years)	mean ± S.D	38.3 ± 11.0	37.7 ± 11.3	41.3 ± 9.5	0.070
Height (cm)	mean ± S.D	177.2 ± 9.5	176.8 ± 9.5	178.5 ± 9.6	0.399
Weight (kg)	mean ± S.D	70.8 ± 11.2	71.0 ± 11.7	70.3 ± 9.4	0.769
Anatomical location of the RRI	cervical spine	1 (1%)	1 (1%)	0 (0%)	
	chest, ribs	1 (1%)	1 (1%)	0 (0%)	
	thoracic spine	2 (2%)	2 (2%)	0 (0%)	
	lumbar spine	12 (9%)	11 (11%)	1 (4%)	
	pelvic floor	1 (1%)	0 (0%)	1 (4%)	
	upper extremity	1 (1%)	1 (1%)	0 (0%)	
	hip	7 (5%)	6 (6%)	1 (4%)	
	groin	3 (2%)	2 (2%)	1 (4%)	
	gluteal region	3 (2%)	2 (2%)	1 (4%)	
	upper leg	4 (3%)	4 (4%)	0 (0%)	
	dorsal side upper leg or hamstring	9 (7%)	6 (6%)	3 (11%)	
	knee	36 (27%)	31 (30%)	5 (18%)	
	lower leg	18 (14%)	12 (11%)	6 (21%)	
	Achilles tendon	9 (7%)	7 (7%)	2 (7%)	
	ankle	11 (8%)	7 (7%)	4 (14%)	
	foot/toe	15 (11%)	12 (11%)	3 (11%)	
Type of RRI	contusion	3 (2%)	2 (2%)	1 (4%)	
	distortion	9 (7%)	7 (7%)	2 (7%)	
	muscle or tendon	42 (32%)	33 (31%)	9 (32%)	
	fracture	2 (2%)	2 (2%)	0 (0%)	
	chondral injury	2 (2%)	2 (2%)	0 (%)	
	back injury	9 (7%)	9 (9%)	0 (0%)	
	hernia	3 (2%)	2 (2%)	1 (4%)	
	bursitis	2 (2%)	2 (2%)	0 (0%)	
	overload	46 (35%)	35 (33%)	11 (39%)	
	other	10 (8%)	6 (6%)	4 (14%)	
Duration of the RRI	0-2 weeks	30 (22%)	22 (21%)	8 (29%)	
	2-4 weeks	45 (34%)	37 (35%)	8 (29%)	
	4-6 weeks	16 (12%)	12 (11%)	4 (14%)	
	6-8 weeks	9 (7%)	8 (7%)	1 (4%)	
	more than 8 weeks	33 (25%)	26 (25%)	7 (25%)	

Abbreviations; RRI, running-related injury; F, female; M, Male; cm, centimetre; kg, kilogram; Δ, change score
Percentages are rounded to the nearest whole number therefore the sum might not be 100%.
*Tests performed between half and whole marathon runners.

OSTRC and Anchor Questionnaire Responses

Median (inter-quartile range [IQR]) OSTRC severity scores for the first and second time the RRIs were reported are shown in Table 3. The median (IQR) OSTRC severity change score, GRC anchor score, and GRL anchor change score were 8 (0-22), 5 (5-6), and 0 (0-1) respectively (Table 3).

Table 3 OSTRC and Anchor Questionnaires (change) scores

		All runners (n=133)	Half marathon (n=105)	Marathon (n=28)	P-value*
OSTRC severity score T1	median (IQR)	25 (20-55)	28 (20-54)	35.50 (22-66)	0.415
OSTRC severity score T2	median (IQR)	22 (6-44)	22 (8-41)	16 (0-49.25)	0.654
OSTRC severity score Δ	median (IQR)	8 (0-22)	8 (0-22)	13.50 (1.75-25.75)	0.221
GRC anchor score	very much worse	0 (0%)	0 (0%)	0 (0%)	
	much worse	0 (0%)	0 (0%)	0 (0%)	
	slightly worse	13 (10%)	11 (11%)	2 (7%)	
	unchanged	14 (11%)	13 (12%)	1 (4%)	
	slightly improved	44 (33%)	34 (32%)	10 (36%)	
	much improved	37 (28%)	27 (26%)	10 (36%)	
	very much improved	25 (19%)	20 (19%)	5 (18%)	
GRL Δ	-3	2 (2%)	2 (2%)	0 (0%)	
	-2	5 (4%)	3 (3%)	2 (7%)	
	-1	11 (8%)	10 (10%)	1 (4%)	
	0	63 (47%)	52 (50%)	11 (39%)	
	1	39 (29%)	28 (27%)	11 (39%)	
	2	7 (5%)	4 (4%)	3 (11%)	
	3	6 (5%)	6 (6%)	0 (0%)	

Abbreviations; OSTRC, Oslo Sports Trauma Research Center Questionnaire; Δ, change score; GRC, Global Rating of Change Anchor; GRL, Global rating of Limitations Anchor; *Tests performed between half and whole marathon runners.

Smallest Detectable Change

The SDC of the OSTRC severity score for runners with a GRL change score of zero at a group and individual level was 0.27 and 5.91 respectively (Table 4). The SDC of the OSTRC severity score for runners who score minus one, zero, or one on the GLR change score at a group and individual level was 0.18 and 5.45 respectively.

Table 4 Smallest Detectable Change

	Runners with Δ GRL =0 (n=63)	Runners with Δ GRL = -1, 0, or 1 (n=113)
Δ OSTRC severity score (95% CI)	7.68 (3.42 – 11.95)	9.17 (5.27 – 13.06)
SEM*	2,13	1.966
SDC_{group} †	0.27	0.18
$SDC_{individual}$ ‡	5.91	5.45

Abbreviations; OSTRC, Oslo Sports Trauma Research Center; SDC, Smallest Detectable Change; SEM, Standard Error of the Mean, Δ, change score; GRC, Global Rating of Change Anchor; GRL, Global rating of Limitations Anchor; *SEM=SD/√n; †SDC_{group}= (SD/√n)/ √n; ‡$SDC_{individual}$= 1.96 x √2*(SD/√n).

Minimally Important Change

As mentioned above, two anchor-based methods were supposed to be used to calculate the MIC of the OSTRC severity score for RRIs that truly improve or worsened separately. However, zero runners reported a RRI which truly worsened on the GRC anchor. Consequently, it was not possible to calculate MIC values with this anchor for runners who truly worsened.

Visual Anchor-based MIC Distribution Method

For runners who truly improved on the GRC anchor (n=62) the optimal MIC value for the OSTRC severity score and the AUC (95%CI) of the ROC curve were 13.50 (95% CI upper limit 38.74) (Table 5) and 0.772 (95% CI 0.686–0.857) (Figure 2) respectively. For runners who truly improved on the GRL anchor (n=13) the optimal MIC value for the OSTRC severity score and the AUC (95%CI) of the ROC curve were 18.50 (95% CI upper limit 43.56) and 0.831 (95% CI 0.772–0.940) respectively. For runners who truly worsened according to the GRL anchor (n=7) the AUC of the ROC curve was 0.250 (95% CI 0.0–0.52). Because the AUC was lower than 0.70 no MIC value was calculated. Figure 3 shows the visual anchor-based MIC distributions with MIC values and 95% CI upper limits.

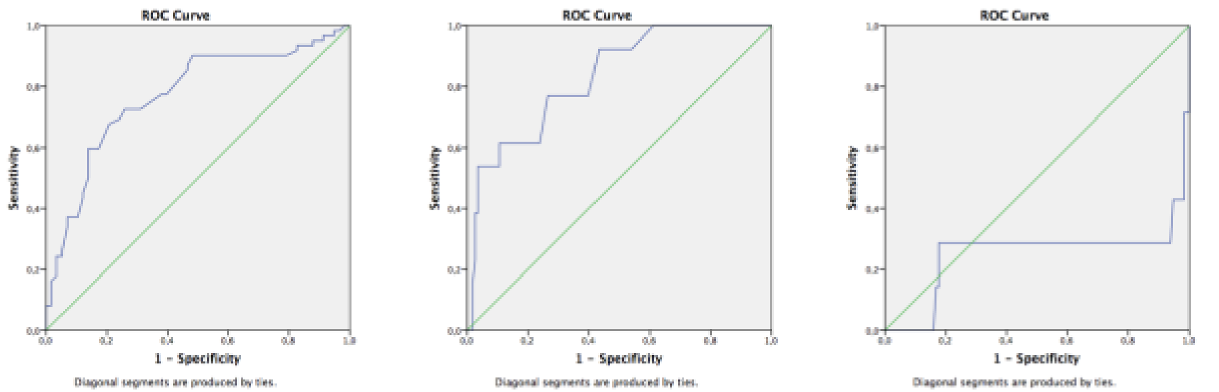

Figure 2 Receiver Operator Characteristic (ROC) Curves
Note; **Left** ROC curve for runners with running-related injuries who truly improve according to the global rating of change anchor; **Middle** ROC curve for runners with running-related injuries who truly improve according to the global rating of limitations anchor; **Right** ROC curve for runners with running-related injuries who truly worsened according to the global rating of limitations anchor.

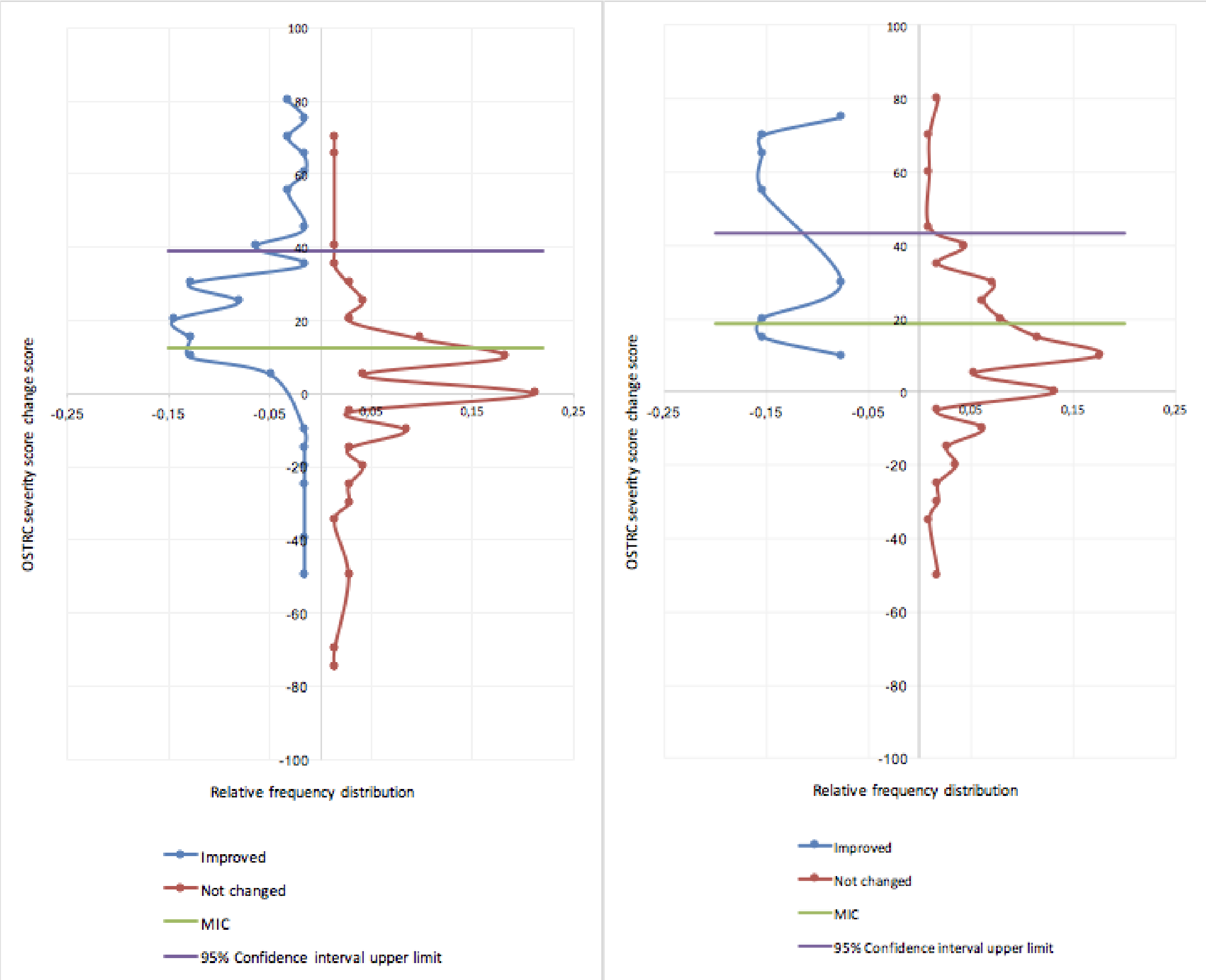

Figure 3 Visual Anchor-Based MIC Distribution according to the global rating of change anchor (left) and global rating of limitations anchor (right)

Note; **Left** graph MIC according to global rating of change anchor (MIC cut-off = 13.50 points, 95% confidence interval upper limit 38.74); **Right** graph MIC according to global rating of limitations anchor (MIC cut-off 18.50 points, 95% confidence interval upper limit 43.56).

Table 5 Minimally Important Change

			AUC (95% CI)	Sensitivity*	Specificity†	MIC ± SD	95% CI
Visual anchor -based MIC distribution	GRC anchor	Truly improved (n=62)	0.772 (0.686 – 0.857)	67.7%	79.3%	13.50	38,74‡
		Truly worsened (n=0)	-	-	-	-	-
	GRL Anchor	Truly improved (n=13)	0.831 (0.772 – 0.940)	76.9%	73.5%	18.50	43,56‡
		Truly worsened (n=7)	0.250 (0.0 – 0.520)	-	-	-	-
Mean change method	GRC Anchor	Truly improved (n=37)	-	-	-	15.49 ± 24.02	(7.48 – 23.49)
		Truly worsened (n=0)	-	-	-	-	-
	GRL Anchor	Truly improved (n=8)	-	-	-	45.38 ± 21.44	(27.45 – 63.30)
		Truly worsened (n=7)	-	-	-	27.71 ± 40.30	(-64.98 – 9.55)

Abbreviations; GRL, Global Rating of Limitations; GRC, Global Rating of Change; OSTRC, Oslo Sports Trauma Research Center; MIC, Minimally Important Change; 95% CI, 95% Confidence Interval; -, item is not applicable; SD, standard deviation;
Note; *sensitivity is the percentage of runners correctly classified as truly improved using the MIC value chosen; †specificity is the percentage of runners correctly classified as unchanged using the MIC value chosen; ‡only upper limit of the 95% CI was selected as proposed by de Vet et al. (2007)[13].

Mean change method

Runners who truly improved on the GRC (n=37) or the GRL anchor (n=8) had a mean OSTRC severity change score of 15.49 (95% CI 7.48–23.49) or 45.38 (95% CI 27.45–63.30) respectively (Table 5). Runners who truly worsened on the GRL anchor (n=7) had a mean OSTRC severity change score of 27.71 (95% CI -64.98-9.55).

Suitability of the Anchors

Spearman's correlation of the both the GRL anchor change score or the GRC anchor score with the OSTRC severity change score were r=0.530 and r=0.487 respectively. Both correlations exceed the predetermined criterion of r≥0.30. Thus, the anchors were suitable to establish the actual change in RRI status.

Dependency of the MIC on the initial OSTRC severity score

Runners with an initial OSTRC severity score lower than the median initial OSTRC severity score had a significantly lower median OSTRC severity change score than runners with an initial OSTRC severity score higher than the median(p=0.013) (Table 6). Thus, the MIC value calculated, using the mean change method, is dependent upon the initial severity of the RRI within the subgroup of runners who truly improved on the GRC anchor.

Table 6 Baseline Dependency of the MIC

		MIC† (Mean ± SD)	95% CI	P-value*
Runners with initial OSTRC severity score lower than median (<22)	N=15	3.87 ± 21.01	(-7.77 – 15.50)	0.013
Runners with initial OSTRC severity score higher than median (≥22)	N=22	23,41 ± 23.08	(13.18 – 33.64)	

Abbreviations; OSTRC, Oslo Sports Trauma Research Center; MIC, Minimally Important Change; 95% CI, 95% Confidence Interval; SD, standard deviation

Note; *test performed using T-test, †, MIC values were calculated using the mean change method.

DISCUSSION

The aim of this study was to evaluate the MIC and SDC of the OSTRC severity score in injured (half) marathon runners. The OSTRC severity score has an adequate responsiveness and interpretability. This means that when a RRI has been registered at least twice, the OSTRC severity change score can be used to distinguish between RRIs that have truly improved and RRIs that have not changed in a two-week period. The SDC of the OSTRC severity score for individual runners was equal to or smaller than 5.91. Thus, the SDC and the measurement error are smaller than the smallest difference between any of the answer scores on the individual OSTRC questions. Consequently, every change in answer scores on the OSTRC exceeds the measurement error. Furthermore, the SDC values for individual runners are smaller than the MIC values. Therefore, the measurement error is sufficiently small for the MIC to be detected.[9] By calculating the MIC in multiple ways a range of MIC values is proposed for RRIs that truly improve (range 13.50-45.38). Within this study it was only possible to calculate a MIC value for RRIs that truly worsened using the mean change method with the GRL anchor (MIC 27.71).

To the author's best knowledge, no other studies have investigated the MIC and SDC of the OSTRC severity score in runners or other populations. Furthermore, no other studies on instruments which measure the severity of a RRI in the same manner as the OSTRC are available. Standard RRI registration methods quantify the severity of a RRI using the number of days of absence from full running participation.[3] This makes it impossible to compare the results of this study to any other.

MIC values can vary depending on the methodology used, the manner by which they were calculated, and initial severity of the RRI.[12,15] A strength of this study is that two anchor-based methods to determine the MIC of the OSTRC severity score were used. Anchor-based methods have a higher validity than distribution based methods, because the anchors inherently contain the observed change in RRI status.[9,12,15] Further, the visual anchor-based MIC distribution method combines an anchor and distribution based method.[13] The distribution of the OSTRC change scores and ROC curves are plotted considering the OSTRC as a diagnostic test and the anchors as the gold standard. This provides a graphical representation of the amount of (in)correctly classified runners by the instrument. Thus, as shown in this study the MIC of the OSTRC severity score is not a fixed value. Therefore, multiple methods to calculate the MIC were used.[15]

Another strength of this study is that two methods to calculate the MIC of the OSTRC severity score were used: the visual anchor-based MIC distribution method and the mean change method. The visual anchor-based MIC distribution method is preferred over the mean change method. Because the mean change method regards the mean change as MIC no matter how large the SD is.[9] Whereas, the visual anchor-based MIC distribution method searches for the optimal MIC value to maximize the amount of runners correctly classified as truly improved or unchanged.[9] Based upon the latter method a MIC value of 18.50 is preferred for use in individual injured runners. For the OSTRC severity change score to exceed 18.50 points the answers on the four OSTRC questions need to improve by at least three answer options on one or more of the OSTRC questions.

Coaches and healthcare providers could use the OSTRC severity score MIC value of 18.50 in the process of monitoring one or multiple athletes. For example, in a study from Clarsen et al. (2014) a group of Olympic and Paralympic athletes filled in the OSTRC every week during their preparation for the Olympic and Paralympic games.[5] If an athlete reported the same injury two or more times, the MIC values of the OSTRC severity score could have helped the staff to make informed decisions on whether or not the change in the injury status was truly important for the athlete.

Limitations

Since no RRIs were classified as truly worsened according to the GRC anchor. It was not possible to evaluate the MIC of the OSTRC severity score for RRIs that truly worsened. Furthermore, for RRIs that truly worsened according to the GRL anchor the AUC of the ROC curve (AUC 0.250) did not meet the criterion of 0.70 with a lower limit 95% CI of 0.5. Therefore, no MIC value was calculated for the RRIs that truly worsened according to the GRL anchor. Using the mean change method, it was only possible to determine a MIC value for RRIs that truly worsened on the GRL anchor. However, the 95% CI of the MIC value for RRIs that truly worsened according to the GRL anchor reflects the low precision of the MIC estimation. Therefore, it is not recommended to use this MIC value, calculated using the GRL anchor, to determine if a RRI truly worsened. Future research could be performed to determine this MIC value.

Terwee et al. (2010) showed that MIC values can vary strongly between different populations.[12] Therefore, it is not advisable to use these MIC values in other athletes than (half) marathon runners. Nevertheless, this is the first study to investigate the interpretability of the OSTRC severity score in runners.

CONCLUSION

This study shows that the OSTRC severity score has adequate responsiveness and interpretability in (half) marathon runners. The SDC of the OSTRC severity score was found to be smaller than the smallest difference between any of the answer scores on the four OSTRC questions. Thus, every change in the OSTRC severity score exceeds the measurement error. Based upon the results of the visual anchor-based MIC distribution method this study proposes to use a MIC value of 18.50 to distinguish between RRIs that truly improved and unchanged RRIs in a two-week period. For the OSTRC severity change score to exceed the MIC value of 18.50 the answers on the OSTRC questions need to improve by at least three options. Additionally, the MIC values could be used by future studies for sample size calculations. Future research is needed to determine the MIC of the OSTRC severity score for athletes in other sports.

REFERENCES

1. Scheerder J (Jeroen), Boen F. *Vlaanderen Loopt! : Sociaal - Wetenschappelijk Onderzoek Naar de Loopsportmarkt*. Academia Press; 2009. https://books.google.nl/books/about/Vlaanderen_loopt_Sociaal_wetenschappelij.html?id=5CX8usTKzAMC&source=kp_cover&redir_esc=y. Accessed May 27, 2017.

2. VeiligheidNL. *Hardloopblessures; Blessurecijfers 2014*.; 2015. www.veiligheid.nl.

3. Timpka T, Alonso J-M, Jacobsson J, et al. Injury and illness definitions and data collection procedures for use in epidemiological studies in Athletics (track and field): consensus statement. *Br J Sports Med*. 2014;48(7):483-490. doi:10.1136/bjsports-2013-093241.

4. Clarsen B, Myklebust G, Bahr R. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology. *Br J Sports Med*. 2013;47:495-502. doi:10.1136/bjsports-2012-091524.

5. Clarsen B, Rønsen O, Myklebust G, Flørenes TW, Bahr R. The Oslo Sports Trauma Research Center questionnaire on health problems: a new approach to prospective monitoring of illness and injury in elite athletes. *Br J Sports Med*. 2014;48(9):754-760. doi:10.1136/bjsports-2012-092087.

6. Pluim BM, Loeffen FGJ, Clarsen B, Bahr R, Verhagen EALM. A one-season prospective study of injuries and illness in elite junior tennis. *Scand J Med Sci Sport*. 2016;26(5):564-571. doi:10.1111/sms.12471.

7. Clarsen B, Bahr R, Heymans MW, et al. The prevalence and impact of overuse injuries in five Norwegian sports: Application of a new surveillance method. *Scand J Med Sci Sport*. 2015;25(3):323-330. doi:10.1111/sms.12223.

8. Mokkink et al LB. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63:737-745. doi:10.1016/j.jclinepi.2010.02.006.

9. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*.; 2011. doi:http://dx.doi.org.myaccess.library.utoronto.ca/10.1017/CBO9780511996214.

10. Sierevelt IN, van Eekeren ICM, Haverkamp D, Reilingh ML, Terwee CB, Kerkhoffs GMMJ. Evaluation of the Dutch version of the Foot and Ankle Outcome Score (FAOS): responsiveness and Minimally Important Change. *Knee Surgery, Sport Traumatol Arthrosc*. 2016;24(4):1339-1347. doi:10.1007/s00167-015-3941-9.

11. Mokkink et al LB. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *Biomed Cent Med Res Methodol*. 2010;10(22).

12. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol*. 2010;63(5):524-534.

doi:10.1016/j.jclinepi.2009.08.010.

13. de Vet HCW, Ostelo RWJG, Terwee CB, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res*. 2007;16(1):131-142. doi:10.1007/s11136-006-9109-9.

14. Terwee C, Bot SDM, de Boer M, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012.

15. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5):395-407. doi:10.1016/S0895-4356(03)00044-1.

SAMENVATTING

Doelstelling Het bepalen van de minimaal klinische relevante verandering (engels; minimally important change, MIC) en het minimaal detecteerbare verschil (engels; smallest detectable change, SDC) van de Oslo Sports Trauma Research Center vragenlijst (OSTRC) score bij geblesseerde (halve) marathonlopers.

Methode Data van een prospectief cohort studie genaamd de SUMMUM-2017 studie werd gebruikt. Gegevens van 133 lopers die tweemaal achter elkaar dezelfde blessure rapporteerden en de global rating of limitations (GRL) en global rating of change (GRC) anker vragen invulden zijn gebruikt. De ankers werden gebruikt om lopers te classificeren als werkelijk verbeterd, onveranderd, werkelijk verslechterd. De SDC werd berekend op groeps en individueel niveau met lopers die onveranderd waren volgens het GRL anker. De MIC werd bepaald middels de visual anchor-based MIC distribution en gemiddelde verandering methode voor zowel blessures die werkelijk verbeteren als verslechteren met behulp van beide ankers.

Resultaten SDC waardes op groeps en individueel niveau waren ≤0.27 and ≤5.91 respectievelijk. De MIC-waardes voor lopers die werkelijk verbeterden volgens de GRC en GRL ankers waren 13.50 en 18.50 punten respectievelijk (visual anchor-based MIC distribution method). Met behulp van de gemiddelde veranderings methode waren de MIC-waardes voor lopers die werkelijk verbeterden volgens het GRC en GRL anker 15.49 (95% Bhi 7.48-23.49) and 45.38 (95% Bhi 27.45-63.30) respectievelijk. Voor lopers die werkelijk verslechterden kon alleen een MIC-waarde worden berekend met behulp van de gemiddelde veranderdings methode met het GRL anker 27.71 (95% Bhi -64.98-9.55).

Conclusie De OSTRC-score heeft een adequate responsiviteit en interpreteerbaarheid en kan gebruikt worden om geblesseerde lopers met een werkelijk verbeterde blessure te onderscheiden van lopers met een onveranderde blessures in een periode van twee weken. Voor gebruik in individuele lopers wordt voorgesteld om een MIC-waarde van 18.50 te gebruiken. Om deze MIC-waarde te overschrijven moeten de antwoorden op een of meerdere OSTRC vragen met minimaal drie opties verbeteren.

Klinische relevantie Een MIC-waarde van 18.50 wordt voorgesteld om te bepalen of blessures van individuele lopers werkelijk relevant of niet veranderen. De MIC-waarde kan gebruikt worden bij het monitoren van geblesseerde lopers. Verder kunnen de MIC-waardes gebruikt worden voor steekproefomvang berekeningen.