Psychometric Analysis of the University of Wisconsin Running Injury and Recovery Index

By

Evan O. Nelson

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Clinical Investigation)

at the

UNIVERSITY OF WISCONSIN-MADISON

2019

Date of final oral examination: 7/31/2019

The dissertation is approved by the following members of the Final Oral Committee:

Bryan C. Heiderscheit, PT, PhD, FAPTA, Professor, Department of Orthopedics and Rehabilitation, University of Wisconsin-Madison (advisor)
Daniel M. Bolt, PhD, Professor, Department of Educational Psychology, University of Wisconsin-Madison
M. Alison Brooks, MD, MPH, Associate Professor, Department of Orthopedics and Rehabilitation, University of Wisconsin-Madison
Tim A. McGuine, PhD, ATC, Distinguished Scientist, Department of Orthopedics and Rehabilitation, University of Wisconsin-Madison
Mary Sesto, PT, PhD, Associate Professor, Department of Medicine, University of Wisconsin-Madison
David J. Vanness, PhD, Professor, Department of Health Policy and Administration, The Pennsylvania State University

Dedication

The enduring love and support I receive from Kari, Claire, Paige, Spencer, and Anna allowed me to pursue this goal and complete this thesis. They ensure I receive love, laughter, motivation, compassion, guidance, and distraction in appropriate measure. I am happy we can celebrate this milestone of completion. It is done!

Acknowledgements

The achievement of one represents the work of many.

My wife, Kari, offered support and understanding. She is the coordinator of chaos in our house, and kept the ship steady when this process was stormy. I love you, and will forever appreciate your support. Claire, Paige, Spencer, and Anna made sacrifices and supported me as I pursued this personal goal. I look forward to endless opportunities to love and support them in return.

I offer sincere thanks to my primary advisor, Bryan Heiderscheit, who is an exemplary role model. Throughout this process, he offered patient mentoring, and his feedback consistently enhanced the quality of my work. Under his guidance, I was able to transform from a consumer of scientific research into a producer of knowledge.

In big and small ways the members of my committee — Dan Bolt, Alison Brooks, Tim McGuine, Mary Sesto, and David Vanness — changed the course of this project and my learning in the process. I have great respect for the expertise they each possess and how they share it with those around them.

The dedicated efforts of the physical therapists at Advanced Physical Therapy & Sports Medicine, Mountain Land Physical Therapy & Rehabilitation, Movement Systems Physical Therapy, and TRIA Orthopaedic Center were critical to the success of this project.

My faculty colleagues in the University of Wisconsin Doctor of Physical Therapy Program provided support and encouragement. I love being a member of this fantastic team of physical therapy educators.

The students, faculty, and staff at the University of Wisconsin Institute for Clinical and Translational Research allowed me to pursue my professional goal of acquiring skills to conduct research enhancing evidence-based physical therapy practice.

Finally, the network of friends and family who support me, and our family, throughout these years. Your understanding, encouragement, and assistance continues to be greatly appreciated.

Data collection services were supported by the UW Institute for Clinical and Translational Research grant support (Clinical and Translational Science Award (CTSA) program, through the NIH National Center for Advancing Translational Sciences (NCATS), grant UL1TR000427).

Veterans RAND 36 Item Health Survey (VR-36) include: Drs. Lewis Kazis, Alfredo Selim, William Rogers, James Rothendler, Austin Lee, Xinhua S. Ren, Donald Miller, Katherine Skinner,+ Jack Clark, Avron Spiro III., Graeme Fincke, and Ms. Shirley Qian. Historically the VR-12 was developed from the VR-36 (Veterans RAND 36 Item Health Survey, principal developer Dr. Kazis) and the MOS SF-36 Version 1.0. The users of this assessment were granted permission to use the official algorithms and scoring that is credited to the developers. The VR-36 and VR-12 are in the public domain as the development of the instrument was supported by federal funds from the Veterans Health Administration, Health Services Research and Development Service (VHA-HSR&D) and the Centers for Medicare and Medicaid Services (CMS). Dr. John Brazier at the University of Sheffield developed the VR-6D. The VR-36 and VR-12 are in the public domain and there is no cost for their use for non-profit organizations.

IRB approval 4/7/2017 UW-Madison Minimal Risk IRB (Health Sciences), submission id 2016-1544

Table of Contents

Abstract

Background: Patient-reported outcome measures (PROs) evaluating recovery following running-related injury (RRI), have not been properly tested and the impact of RRIs on overall health is unknown. This thesis describes the development and psychometric assessment of the University of Wisconsin Running Injury and Recovery Index (UWRI); it also investigates the relationship between RRIs and health-related quality of life (HRQoL).

Methods: An iterative, interview-based, item generation process included the target audience. Runners seeking care for a RRI completed PROs at the baseline and 12 weeks later during the psychometric assessment. Validity was determined by comparing changes in UWRI score with the global rating of change (GROC), Veterans Rand 12-item (VR-12), and body-region specific PROs. Responsiveness was determined using anchor-based and distribution-based techniques. UWRI items were mapped to the VR-12 component scores to investigate the relationship with HRQoL.

Results: The 9-item UWRI incorporates key elements runners use to self-assess recovery; the maximum score of 36 indicates restoration of pre-injury running ability. Prospective assessment demonstrated UWRI change (mean±SD 7.7±9.0) was correlated with GROC (r=0.67), as well as the changes in VR-12 physical component score (r=0.54) and mental component score (r=0.31). UWRI change was correlated with changes in lower extremity, but not low back specific PROs. The UWRI was responsive to changes in running ability; the minimal clinical important difference was 8 points.

Conclusion: The UWRI is a valid, reliable, responsive evaluative PRO of running ability following RRI. Following RRI, running ability is associated with physical, and to a lesser degree mental, HRQoL.

Introduction

Running-related injuries (RRIs) are a common problem causing runners to reduce training, miss competitive events, or leave the sport altogether. In addition to running-specific impairments, RRIs also cause limitations in daily life and motivate runners to seek medical care. In the absence of a preferred assessment of running ability, patient-reported outcome measures (PROs) specific to a condition (e.g., patellofemoral pain, Achilles tendinopathy) or region of the body (e.g., Lower Extremity Functional Scale, Foot and Ankle Ability Measure) are frequently used to evaluate the response to care. Assessing clinical effectiveness in this way may be not go far enough to determine the true benefits of standard RRI interventions.[1] Common running-specific interventions, such as gait retraining, are administered to improve running ability across injuries and conditions.[2] Most body region-specific or condition-specific PROs include a small number of items assessing athletic tasks, while the majority of items evaluate low-demand functional activities. Thus, the ability of these PROs to measure changes in running ability may be questioned, as they have never been subjected to a psychometric assessment in a population of injured runners.

The inability to accurately determine the effectiveness or meaningfulness of common interventions represents a significant problem. To solve this problem, it is important to understand the critical elements runners use to determine recovery, describe recovery of running ability following injury, evaluate tools measuring the construct, and determine the meaningfulness of the measurements. The clinical research required to solve this problem has high translational potential because the solution will have clinical and research applications.

Epidemiology of running-related injuries

Globally, running is an extremely popular form of exercise, with an estimated 7.9 million runners participating in running events in 2018.[3] In the United States, distance running has incredible annual participation with over 541,000 people completing marathons, more than 2 million people finishing half marathons, and 1,582,302 high school athletes participating in outdoor track and field or cross country.[4] Although running has many positive benefits, musculoskeletal injuries are unfortunately a common consequence. The risk of sustaining a running-related injury (RRI) varies with race distance, running volume, experience, and method of injury assessment.[5-7] Evaluations of injury in relationship to running exposure estimate RRI incidence in novice runners is 17.8 injuries per 1,000 h of running (95% CI 16.7, 19.1); this is significantly higher than the 7.7 injuries per 1,000 h of running (95% CI 6.9, 8.7) in recreational runners.[8] Similar exposure-associated risk calculations based on running distance determined 1.07 injuries occur per 1,000 km of running (95% CI 1.01, 1.13).[8] Most RRIs affect the lower extremity, primarily affecting the knee, lower leg, ankle, and foot.[6 9-13] Medial tibial stress syndrome, patellofemoral pain syndrome, iliotibial band syndrome, Achilles tendinopathy, plantar fasciitis, and gastrocnemius or soleus muscle injuries are the most common RRI diagnoses; however, RRIs include diagnoses requiring more significant medical or surgical management as well.[6 14 15] Numerous risk factors have been proposed to explain the relatively high RRI incidence; but a review of prospective cohort studies identified “injury in the past 12 months” as the only consistent risk factor for the heterogeneous collection of RRIs.[16] The relatively high risk of injury combined with the popularity of the sport necessitates a better understanding RRI recovery that may avert the potential for re-injury.

Etiology

The causal pathway connecting risk factors with injury is not well understood, but the current perception of etiologic factors leading to RRI provides insight into the restoration of running ability following injury. During recovery, the training principles applied by runners seeking to improve performance or physical fitness are implemented to restore running ability. Within the running community, consensus identifies the etiology for most RRIs is training error because substantial changes in running volume, distance, or pace induce abrupt increases in running-imposed load (i.e., “too much, too soon”).[17 18] However, the criterion for how much is too much has not been quantified.[15 19]

Bertelsen et al elaborated upon this idea in a conceptual framework for the causal mechanism explaining the development of RRIs.[20] The model demonstrates an RRI occurs when the load capacity of a specific structure is exceeded by the cumulative load of a running session.[20] Entering a running session, previous training, diet, prior injury, disease, sleep, age, recovery duration, psychological factors, or other individual characteristics affect the tissue’s ability to withstand the running-imposed load.[20] During a running session, the cumulative load applied to a specific structure is a function of the load magnitude of each stride, distribution of the load, and the total number of strides.[20] Step rate manipulation is a practical method to modify the distribution of running load across tissues, while modifying other determinants of cumulative load may be unsuitable for training goals (e.g., distance, speed) or not reasonably modifiable (e.g., body weight).[21 22]

Theoretically, RRIs may be classified as acute (traumatic) when the force exceeding a specific structure's capacity is applied in a nearly instantaneous fashion, but more commonly the injury is progressively acquired and considered an "overuse injury."[23] Operationally, symptoms originating during a clearly identifiable event differentiate acute (traumatic) injuries from overuse injuries.[24] However, this operational strategy risks misclassifying injuries with abrupt symptom onset following a period of subclinical, cumulative tissue damage. Misclassification is a risk for RRIs because the onset of clinical symptoms often follows a period of subclinical tissue damage.[25] Assessments of recovery require independence from this common classification scheme because athletes often continue to participate despite ongoing symptoms.[26]

Functional consequence

For runners, it is of significant consequence to reduce training, miss competitive events, or receive medical care due to injury; but RRIs may impose more subtle consequences.[5 13 27-31] A recent survey of 1,145 runners in the United Kingdom found 49.8% had a current RRI: 91% of these runners reported the injury caused pain, 89% felt it directly impacted running performance, 86% had to reduce their running volume due to the injury, but only 6% were unable to run because of the injury.[12] Runners frequently continue running despite experiencing symptoms. While RRI symptoms typically decrease running performance, in some cases, running distance, speed, or volume are not reduced in the presence of injury.[32] Valid assessments of RRI recovery must detect changes in symptoms the patient experiences while running. Many injury surveillance systems use a "time loss" definition of injury, which does not represent the experience of injured runners.[33]

In addition to the physical symptoms, athletic injuries, including RRIs, are known to induce dynamic, individualized psychological and emotional responses that may fluctuate over time, but play an important role in recovery.[34] The psychological effects of RRIs have not been specifically evaluated, but runners are represented in some studies.[35] Injury-related emotions, cognitions, and behaviors compose the core themes of the multiple psychosocial factors associated with sports injury rehabilitation.[36] Fluctuations are especially evident at transitions in the recovery process, including the return to sport when sport-specific self-confidence is necessary to avert a fear response.[36 37] Despite consensus regarding the importance of repeated psychological assessment throughout the entire recovery process, there are no preferred assessments.[34] Multiple assessment instruments are available, but most are too long to be clinically practical.[38] Efficient methods to integrate psychological screening into clinical care would be beneficial because positive psychological responses are associated with better outcomes, and interventions for negative psychological responses decrease time lost due to injury.[34 38]

Running has unique physical and psychological demands. Following injury, the process to return to normal health is accompanied by a variety of physical and psychological symptoms. The consequences of these symptoms demonstrate variability, but the vast majority of injured runners continue to train or compete despite ongoing symptoms. Valid assessments of recovery must incorporate the patient-perspective, and respond to meaningful fluctuations in the factors associated with recovery.

Treatment

Initially following injury, RRI symptoms are managed individually by the runner or in collaboration with a medical professional. Once the initial symptoms are reduced to a manageable level, treatment involves a shared decision-making strategy considering short and long term goals, while returning the athlete to pre-injury performance as quickly as possible. The potential for re-injury or complication is minimized by controlling the overall running load while weighing the runner's desire to return to running.[1] During this process, the expectation is to induce functional overreaching to stimulate adaptation while avoiding nonfunctional overreaching and the consequential risk for re-injury or symptom regression.[39] The transition from functional to nonfunctional overreaching is indicated by the runner's perceived recovery. Recovery duration varies with injury severity, but recovery invariably includes a period of increasing running distance, speed, or frequency to increase running load.

Emerging in the past 10 years, gait retraining is now a routine biomechanical intervention to reduce and re-distribute running imposed loads.[2] An abundance of studies examined how gait retraining alters running kinematic and kinetic factors associated with injury. As a result of these studies, there are variations in the core gait retraining techniques used in clinical practice.[21 40-43] However, the intended purpose of all gait retraining techniques is to induce biomechanical changes in running gait to reduce load, redistribute load, or alter key kinematic and kinetic variables. Clinically, gait retraining is used to prevent RRI, reduce injury symptoms during running, enable running despite injury, or minimize re-injury risk with continued running.

Despite a plethora of biomechanical studies, only three studies examined the clinical efficacy of gait retraining. Conflicting results exist, and the overall usefulness of these studies is limited because overly simplistic measurements are used to determine treatment response.[43-45] Strategies to balance clinical research integrity and clinical practice generalizability of gait retraining studies have been debated in the literature.[46 47] The most recent randomized control trial investigating gait retraining used an activities of daily living PRO to evaluate standard of care interventions for runners with patellofemoral pain.[45] In this study, gait retraining reduced vertical loading rate and patellofemoral joint loading.[45] However, the study concluded clinical care was no better than patient education due to the equitable reduction in symptoms and functional limitations.[45] Using invalid assessments to deem treatments ineffective poses a risk to evidence-based practice. Conducting this type of study using valid assessments of the multifactorial patient experience would substantially advance our understanding of the intervention's clinical benefit. Studies of clinical efficacy and clinical effectiveness would be simplified if a valid measure of treatment response could be implemented in research studies and routine clinical practice.

Assessment of recovery

Running ability is arguably the essential component in RRI recovery. Injury heterogeneity and individualized running behaviors make it difficult to assess running ability during recovery. Running ability is further complicated by the disparate physical and psychological impacts the injury may have on different types of running (e.g. race, long run, interval training) which are not comprehensively assessed using training metrics (e.g., weekly volume, pace, long run distance). The recovery process may be influenced by internal factors (e.g., psychological response, confidence, motivation) or external factors (e.g., pressures from coaches, teammates, parents, the competition schedule).[48] Incorporating the patient-perspective is critical because an injury's physical impact occurs concurrently with the multiple internal, external, or psychological factors that influence the recovery process. PROs are the principle tools which incorporate the patient-perspective into functional ability assessments. Limitations in running ability may not be properly assessed by many PRO measures because they fail to evaluate the various factors involved in the recovery process.

Outcome and return to sport

Despite the number of RRI epidemiological studies published in the past decade, there is a limited quantity of research describing the RRI recovery and return to sport. The DANO-RUN study prospectively enrolled 933 novice runners with 254 runners (27.22%) reporting at least 1 RRI.[14] Defining recovery as "no pain in the affected anatomical location following two consecutive running sessions of at least 500m," 86.6% of injured runners reported complete recovery with a median time of 71 days.[14] However, in those who fully recovered, the recovery experience demonstrated significant variability. Full recovery occurred for 220 injuries, but required 9-582 days, with 18.2% reporting medication use, 10.7% receiving conservative care, and 4.7% undergoing surgery.[14] Four percent of the injured runners remained injured at the end of the 24-month follow-up period; 20 people experienced functionally limiting symptoms for longer than 1 year; and 16.4% sustained an additional RRI after fully recovering from an initial injury.[14]

A prospective study of recreational runners registered for a variety of distance running events used surveys to determine the incidence, prevalence, and course of RRIs. The injury incidence was 1 injury per 1,303 km of running, and the mean 3-month incidence was 16.3%.[11] The mean(SD) pain intensity was 4.3(2.03) and 69.1% of injuries recovered within 10 days.[11] The injury incidence is consistent with other studies, but the authors acknowledged the self-reporting of relatively minor injuries likely shortened the recovery time.[8 11]

The recovery period was closely monitored during an economic evaluation conducted by Hesphanhol Junior et al. Thirty percent of the study sample experienced an RRI with 34.1% of injuries requiring medical care.[49] The subgroup seeking medical care had a mean time loss of 6.6 days, mean cumulative time loss of 28 days, and mean recovery time of 9 weeks.[49] The medical care subgroup demonstrated a different injury profile than the overall sample, mean time loss of 2.4 days, mean cumulative time loss of 3.0 days, and mean recovery time of 4 weeks.[49]

Even though the majority of RRIs improve quickly with relatively little intervention, those seeking medical care compose a unique cohort. It makes logical sense that more significant RRI symptoms will cause people to seek medical care, and ultimately these injuries will have prolonged recovery periods. During this period, it is important to monitor recovery using a standardized assessment tool to evaluate the patient-specific response to treatment and also to enable comparison of treatment effects. Recovery is a continuous process; dichotomizing recovery using arbitrary time or distance criteria lacks face validity and fails to incorporate the patient-perspective.

Health-related quality of life

Some runners who sustain an RRI report a negative quality of life impact and, at times, alteration of their self-identity. Recently, a randomized control trial found mental health-related quality of life (HRQoL) to be associated with RRI incidence.[10] The underlying mechanism for this association is unclear, and the quality of life toll imposed by RRIs has not been examined nor quantified. HRQoL is a multi-dimensional concept describing the interaction between a specific health state and quality of life. This concept is depicted by the definition of health in the Constitution of the World Health Organization: “A state of complete physical, mental, and social well-being not merely the absence of disease and infirmity.”[50] Scientific literature offers debate and uncertainty regarding a single HRQoL definition; however, the application of this concept in health research is quite clear.[51] HRQoL measures are a type of PRO measure evaluating an individual’s perceived health state. [52] Preference-based HRQoL measures assign a value, or utility, representing the preference for a specific health state on a scale where 0 is death and 1.0 is perfect health.[53] Changes in health utility reflect how an injury, health condition, or medical treatment affects overall health.[52] It is unknown if RRIs produce measurable change in health utility.

Generic preference-based HRQoL measures are the preferred method to estimate health utility for health technology assessments such as cost utility, cost-effectiveness, or comparative effectiveness analyses. [54 55] Clinicians, however, find generic preference-based measures difficult to use when making individual patient care decisions.[56] Due to this, clinicians favor clinically-oriented PRO measures to preference-based measures, but health utility estimates are necessary to make equitable comparisons of health technologies. Because preference-based measures are not always available in clinical data, a regression analysis may be used to map, or “cross-walk,” a clinically -oriented PRO measure to a preference-based utility measure.[53] Exploring the impact of RRIs on HRQoL is necessary to understand the value of the medical care for RRIs. Generic HRQoL life measures apply to a broad spectrum health states while runners inherently represent a narrow spectrum of relatively healthy individuals. The measurement properties of generic measures have not been evaluated in a running population and may not be responsive to subtle changes in running ability. Assessing the psychometric properties of HRQoL assessments in a sample of injured runners will also provide insight into how HRQoL is affected by RRIs.

Organization of the dissertation

Work included in this dissertation began prior to enrollment in the graduate program in Clinical Investigation. The need for a running-specific PRO was recognized; which lead to an iterative process developing the University of Wisconsin Running Injury and Recovery Index

(UWRI). Interviews with previously injured runners were conducted to understand RRI recovery and the methods runners use to evaluate recovery. Incorporating feedback from runners in each step, the UWRI was developed, tested, and revised to create a PRO outcome measure with good test-retest reliability.

The objective of this dissertation is to conduct a psychometric assessment of the UWRI as an evaluative PRO assessing running ability following RRI, and to investigate the association of running ability and HRQoL. The first chapter will describe the development of the UWRI and the initial assessment of validity. Chapter 2 will expand upon longitudinal validity and provide a detailed responsiveness assessment of the UWRI. Chapter 3 will map the UWRI to the Veterans RAND 12-item Health Survey (VR-12) component scores, describing the relationship between RRIs and HRQoL. The final chapter is a summary of the contributions of this research and an overview of the future work to be done.

Chapter 1. Development of the University of Wisconsin Running Injury and Recovery Index

Evan O. Nelson, Michael Ryan, Erin AufderHeide, Bryan Heiderscheit

(Note: this chapter is in press in the Journal of Orthopaedic and Sports Physical Therapy)

Abstract

Background: Runners experience a high proportion of overuse injuries with extended recovery periods involving a gradual, progressive, return to pre-injury status. A running-specific patient-reported outcome (PRO) measure does not exist, and a questionnaire assessing critical elements of runners' recovery process may have superior psychometric properties. The purpose of this study was to develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after running-related injury (RRI) for clinical practice and research applications.

Methods: Self-identified runners and selected experts participated in an iterative, 6-step development process creating the University of Wisconsin Running Injury and Recovery Index (UWRI). Content-related validity was assessed using open comments. Reproducibility was assessed using Cronbach's alpha (α), intraclass correlation coefficient ($ICC_{(agreement)}$), and standard error of measurement ($SEM_{agreement}$). An anchor-based, construct validity assessment measured the association between the change in the UWRI score and global rating of change (GROC). Responsiveness assessments included floor and ceiling effects.

Results: The 9-item UWRI assesses running ability following RRI with the maximum score of 36 indicating a return to pre-injury running ability. The UWRI demonstrated acceptable internal consistency (α = 0.82), test-retest reliability ($ICC_{(agreement)}$ = 0.934), and $SEM_{agreement}$ = 1.47. Change in the UWRI score was moderately correlated with the GROC (*r* = 0.58, 95% CI [0.37, 0.74]). Floor and ceiling effects were absent. Completion required 3 minutes 15 seconds.

Conclusion: The UWRI is a reliable PRO measure responsive to changes in running function following RRI with minimal administrative burden.

Key Words: running injury, patient-reported outcome measure, psychometric assessment

Introduction

Running-related injuries (RRIs) include a heterogeneous collection of musculoskeletal injuries whose pain and physical limitations cause distance runners to miss competitive events, reduce training time and receive medical care.[5,27-29] To explain the RRI heterogeneity and variable precipitating factors, a conceptual framework has been proposed that incorporates tissue-specific load capacity and running imposed load.[20] Additionally, this conceptual framework acknowledges that stress, fatigue, or psychological factors may contribute to an individual's injury susceptibility. Athletic injuries, including RRIs, produce an emotional response influenced by internal factors (e.g., fear, confidence, motivation) or external factors (e.g., coaches, teammates, parents, the competition schedule) that continues through the recovery process.[34-36,48] Providers balance numerous considerations during rehabilitation, and the recovery from such injuries may not demonstrate equivalent rates of improvement for the different types of running (e.g., daily runs, long runs, speed training, threshold intervals, races).[1,17] Clinical assessments of running ability must account for an injury's disparate impact on different types of running and the dynamic nature of the multiple factors influencing the recovery process.

Patient-reported outcome (PRO) measures are the principle tool used by physical therapists to quantify limitations in functional ability and determine change in patient status due to clinical care.[57 58] It is essential PRO measures demonstrate efficiency and strong measurement properties because these assessments influence subsequent clinical care decisions.[59] A running-specific PRO measure does not currently exist. In the absence of a running-specific measure, PRO measures specific to a body region (e.g., knee, hip) or condition (e.g., Achilles tendinopathy, patellofemoral pain) are used to assess functional limitations imposed by RRIs.[60-65] The psychometric properties of these PRO measures have not been assessed in a running population and these measures may have limited evaluative capability because the majority of items assess low-demand activities or fail to assess the multiple internal and external factors which influence recovery following a RRI.[66-68] A running-specific PRO measure could potentially create a valid, reliable, responsive assessment of running ability to measure RRI severity in prospective clinical studies and standardize clinical effectiveness evaluations in practice and research.[69-71]

The purpose of this study was to develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after RRI for clinical practice and research applications.

Methods

The development of this new instrument consisted of 6 steps: (1) item generation, (2) item reduction, (3) item clarification and content validation, (4) preliminary evaluation and revision, (5) reliability assessment, (6) psychometric assessment. For each step, a RRI was defined as a musculoskeletal problem resulting from running that required the individual to prematurely stop running while training or during a competitive event or to miss a training session or competitive event.[72 73] Rather than employ strict criteria, research participants were allowed to self-identify as a runner in attempt to represent the diversity within the target population. Runners who suffered at least 1 RRI were recruited from running clubs, groups, and retail stores in Wisconsin and the surrounding states (step 1-3 & 5). Individual solicitation recruited physician, physical therapist, and running research experts with at least 5 years of experience conducting patient care in a specialized running clinic or publishing articles about RRIs in peer reviewed journals (step 3). Runners experiencing a current RRI were recruited from the UW Health Runners' Clinic (step 4 & 6). The UW Health Runners' Clinic is comprised of physical therapists who are running experts using a shared decision making process when performing running evaluations including video gait assessments and implementing multimodal intervention plans using a combination of therapeutic exercise, running gait modification, training recommendations, and footwear recommendations.[74] Patients frequently receive concurrent physical therapy care from a separate provider in the UW Health Sports Rehabilitation Clinic.

Questionnaire Development

Semi-structured telephone interviews conducted by a single interviewer (E.A.) explored the spectrum of the RRI experience and investigated how etiological factors influence the recovery process. Questions surveyed the injury description and associated pain, activities while injured, training and racing following injury, emotional impact, recovery and expectations, and an open-ended request for additional information. All responses were recorded, transcribed, and coded before generating partially overlapping potential items. Using an ordinal scale, participants assessed the relevance (0=no; 1=yes) and importance (1=not important; 5=very important) of each potential item. Items were ranked by the importance product, calculated as the sum of the importance divided by the sum of the relevance across participants.[75] Importance product ranking and expert review reduced the item pool to develop a draft questionnaire. Written comments related to the comprehension, clarity, and ease of responding to the draft questionnaire were sought during item clarification. To assess content validity, participants were asked if the draft questionnaire comprehensively assessed the RRI recovery process.[76]

Psychometric Assessment

Item analysis examined the response distribution for each item. Pearson correlation coefficients were calculated to determine how each item correlated with the total score as well as the impact of removing any single item. An exploratory factor analysis using maximum likelihood with promax rotation was performed.[76] To include systematic sources of error, the intraclass correlation coefficient for agreement ($ICC_{agreement}$) evaluated test-retest reliability between paired assessments.[77] Response memory is possible between assessments, but physiologic or neuromuscular change is unlikely and longer response periods would increase the likelihood a change in running ability may occur. The standard error of measurement ($SEM_{agreement}$) was calculated to evaluate score stability between administrations, including systematic error, as $SD * \sqrt{(1 - ICC)}$.[76] Coefficient (Cronbach's) alpha correlation was calculated to evaluate internal consistency for the collective tool and each identified factor.[76 78] Anchor-based validation assessments are a common approach to evaluate patient-reported change in measures which cannot be compared to a gold standard.[76] In the absence of a universal assessment of running ability multiple comparison measures were used through this project. The Lower Extremity Functional Scale (LEFS) is a 20-item patient reported outcome measure with the 80 point maximum composite score indicating higher functional ability.[60] The numeric pain rating scale (NPRS) is an 11 point (0-10) ordinal response scale with endpoint descriptors "No Pain at All" and "Worst Possible Pain."[79] The global rating of change (GROC) is an 11 point (-5 to +5) ordinal response scale with endpoint descriptors "Very Much Worse" and "Completely Recovered."[80] The strength of the association between outcome measures was used to evaluate construct validity using the Pearson correlation coefficient. Correlation coefficients were considered good $|r| > 0.75$, moderate $0.50 < |r| \leq 0.75$, fair $0.25 < |r| \leq 0.50$ and no association $|r| < 0.25$.[81] Responsiveness was assessed as the average change in PRO score reported by individuals reporting improvement (GROC +4 or +5) and slight improvement (GROC +2 or +3) will be compared to evaluate responsiveness using a "between-subjects", and "within-subjects" approach, respectively.[79 82] Statistical significance was set at $P < .05$. All analyses were performed using the psych, irr, psychometric, and xlsx packages in R software (Vienna, Austria).[83-87]

The testing protocols were approved by the Health Sciences and Education and Social/Behavioral Sciences Institutional Review Boards at the University of Wisconsin-Madison. Informed consent was obtained and the rights of participants were protected in accordance with institutional policies.

Results

Step 1: Item Generation Sixteen runners who suffered at least 1 RRI agreed to be interviewed. Five participants were elite or professional runners and most had experience running on a team with an official coach at the high school (14/16), college (13/16), or post-collegiate (5/16) level. The authors conducted axial coding using the interview question categories to identify how runners assess recovery. Data saturation occurred and interviews of recreational runners did not produce new themes. Forty-two potential items were generated.

Step 2: Item Reduction A unique sample of 79 runners who suffered at least 1 RRI completed paper surveys containing 42 potential items and importance product questions (i.e., relevance and importance). Items were ranked by importance product and items above the median were consistent with the primary factors identified in step 1. Items requiring skilled assessment were removed from the item pool. The RRI's impact in daily life was a critical recovery component identified in step 1, however, all items assessing specific non-running activities (e.g., squatting, walking, etc.) were below the median and consolidated into a single item. Items assessing cross-training or physical fitness were removed because they do not directly assess running ability. The conceptual framework incorporating physical symptoms, running performance, and psychological responses was maintained when content consolidation reduced the pool to 12 items. Individual interviews with 6 experts affirmed the relevance of 9 items, but suggested removing 3 items assessing interval training, event participation, and individual running goals as these were not generally applicable or unlikely to assess the spectrum of recovery. Expert panelists provided consensus agreement the remaining items comprehensively assess clinically relevant components of post-injury running ability. The beta version of the 9-item University of Wisconsin Running Injury and Recovery Index ($UWRI_{beta}$) included a 7-point numerical response with unique written endpoint descriptions for each item's response scale.[88] The instructions specified a 7-day recall period and a hard stop question restricted completion of items 4-9 unless the individual ran during the preceding 7 days. A question to enforce the recall period was included because an individual may experience improvement when the medical management advises abstinence from running.

Step 3: Item Clarification and Content Validation A unique sample of 31 runners who suffered at least 1 RRI provided comments regarding $UWRI_{beta}$ item clarity and content validity using an electronic form. Three research team members (E.N., M.R., G.T.), who were runners, performed triangulation by analyst using open coding of all comments until data saturation occurred. The results show the $UWRI_{beta}$ was clear, easy to understand, and comprehensively assessed the recovery of running ability post-injury. However, participants suggested a 5-point response scale because it would be difficult to differentiate perceived changes in running status on the 7-point scale. A 5-point ordinal response system using checkboxes with written anchors at the midpoint and endpoints was created for each question.[89] Editorial changes, that did not alter the item concept, were made to 5 items because greater than 10% of the sample commented on item syntax. Three participants recommended assessing concepts identified in step 1 that were below the median in step 2. Each item was scored 0 to 4 with the maximum score of 36 indicating no deficit in running ability. If an individual had not run during the 7 days preceding completion of the $UWRI_{beta}$, items 4-9 were scored as 0.

Step 4: Preliminary Evaluation and Revision Individuals seeking physical therapy care at the UW Health Runners' Clinic for a RRI were invited to participate in a prospective, preliminary assessment of the $UWRI_{beta}$. Inclusion criteria in this phase included being at least 14 years old, able to read and write in English, and of generally good health. Exclusion criteria included surgery for the injury in the past 12 months, rheumatologic disease, systemic connective tissue disorders or clinical diagnosis of depression. *A priori* sample size calculation indicated 23 individuals would be needed to detect a correlation greater than zero assuming $\alpha = 0.05$, $\beta = 0.20$, $r = 0.55$.

Participants completed the $UWRI_{beta}$, LEFS, and NPRS at the initial visit, prior to receiving physical therapy care, as is common practice in this population.[60 79 90] To assess reliability, all participants were asked to complete the $UWRI_{beta}$ electronically 48-72 hours following the initial assessment. The $UWRI_{beta}$, LEFS, NPRS, and GROC were electronically completed 8 weeks following the initial assessment because measurable functional change is likely to occur in 8

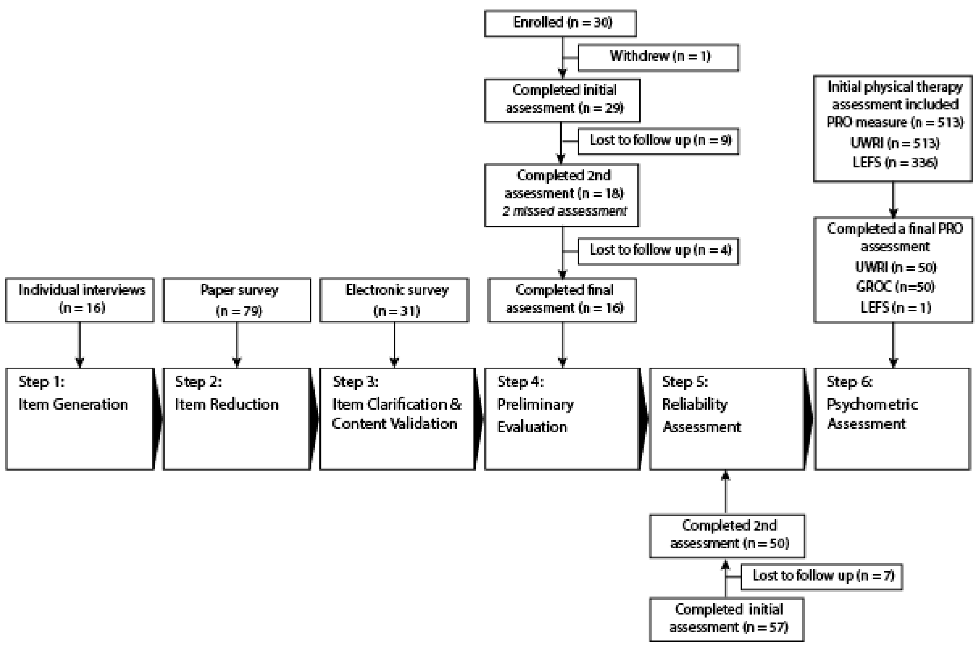

FIGURE 1.1. Participant flow diagram

weeks.[91-93]

Thirty individuals (20 female) agreed to participate and entered the study following the informed consent process (FIGURE 1.1). One individual withdrew for unknown reasons before completing any assessments and 13 individuals were lost to follow-up as they did not respond to requests to complete the 8 week assessment. Baseline characteristics were not different in those lost to follow-up. Participants demonstrated a large age range, a variety of RRI that were predominately chronic in nature, and reduced running volume (TABLE 1.1).

TABLE 1.1. Participant demographics			
	Step 4 Preliminary Evaluation	Step 5 Reliability Assessment	Step 6 Psychometric Assessment
n	30	57	513
Age, y	36.9 ± 11.0 (range 15-57)	33.4 ± 8.4 (range 20-54)	35.4 ± 12.9 (range 12-74)
Gender, n (%)			
Female	20 (66.7%)	32 (64%)	296 (57.7%)
Male	10 (33.3%)	18 (36%)	217 (42.3%)
Running experience, y	2.4 ± 1.2	9.9 ± 7.2	11.5 ± 10.7
Symptom duration, months	8.8 ± 9.9	2.8 ± 1.7	NA
Current running volume, km/wk	21.9 ± 26.1	31.7 ± 26.3	27.6 ± 21.1
Pre-injury running volume, km/wk	37.2 ± 24.6	57.7 ± 35.4	NA
Injury location, n (%)	40 (100%)	50 (100%)	NA
Lumbopelvic	4 (10%)	4 (8%)	
Hip or thigh	6 (15%)	10 (20%)	
Knee	14 (35%)	8 (16%)	
Leg or calf	6 (15%)	4 (8%)	
Ankle	2 (5%)	7 (14%)	
Foot	8 (20%)	13 (26%)	
Other	0 (0%)	4 (8%)	
Abbreviations: NA, not available. Values are mean ± SD unless otherwise indicated.			

The median (range) $UWRI_{beta}$ score was 14 (0, 27) at the initial and 24 (6, 34) at the final assessment creating a statistically significant quantity of change in the Wilcoxon signed rank test of paired data, 10 (-17, 27) points (*P* = .008). The median (range) LEFS score was 69 (42, 79) at the initial assessment and 77 (46, 80) at the final assessment equating to a median change of 5.5 (-20, 24) points that was not statistically significant (*P* = .08). The median (range) NPRS score was 3 (0, 7) at the initial evaluation, 2 (0, 8) at the final assessment, and had a median change of -1 (-4, 3) points that was not significant (*P* = .37). The median (range) GROC was 3 (-5, +5). A good correlation was observed between the change in $UWRI_{beta}$ score over 8 weeks and GROC, $r = 0.75$, 95% CI [0.39, 0.91] ($P = .001$).

The $UWRI_{beta}$ and LEFS demonstrated moderate correlation for the change over the course of the study, $r = 0.55$, 95% CI [0.05, 0.82] ($P = .03$), as well as at the 8 week assessment, $r = 0.73$, 95% CI [0.35, 0.9] ($P = .002$). A correlation was not observed between the $UWRI_{beta}$ and LEFS at the initial assessment, $r = 0.04$, 95% CI [-0.33, 0.4] ($P = .82$). A correlation between the $UWRI_{beta}$ and NPRS was not observed at any assessment. Individuals reporting significant clinical improvement on the GROC (+4 or +5) demonstrated a median (range) UWRI change of 11.5 (9-20) points and a median (range) LEFS change of 5.5 (-2-15) points.[94] Three minutes 15 seconds were required to complete the $UWRI_{beta}$.

Eighteen individuals completed the $UWRI_{beta}$ at the initial evaluation and day 2 assessment. The difference in mean scores (SD) of these administrations was 2.6 (5.7) and not significant ($P = .07$). The $UWRI_{beta}$ scores were moderately correlated between the 2 assessments, $ICC_{agreement} = 0.56$, 95% CI [0.16, 0.81] ($P = .004$). Cronbach's alpha for reproducibility shows acceptable internal consistency, $\alpha = 0.75$, 95% CI [0.28, 0.93]. The preliminary assessment revealed the test-retest reliability of the $UWRI_{beta}$ was adversely affected because items 4-9 were only presented if the person had run in the past 7 days. The $UWRI_{beta}$ was revised to create the UWRI, presenting all items on each administration (APPENDIX 1). Written anchors accompany each response option, and participants can select "unable to run" if the condition severity or medical recommendation prevent running.

Step 5: Reliability Assessment Fifty-seven English literate, self-identified adult runners in generally good health, with a single, current RRI participated in an evaluation of the UWRI test-retest reliability using a web-based survey. The sole exclusion criteria was having surgery at the injury site in the past 12 months. The web-based survey disassociated this assessment from clinical care because common RRI interventions including gait retraining or educational interventions may have immediate effects that influence perceived running ability.[2 95] Fifty participants reported demographic information, injury duration and location, current and pre-injury running volume and completed the UWRI electronically on 2 occasions separated by 24-72 hours (TABLE 1.1). UWRI test-retest reliability was excellent with an $ICC_{agreement}$ = 0.93, 95% CI [0.89, 0.96] ($P < .001$). UWRI scores were not different between the two administrations ($P = .12$). The $SEM_{agreement}$ was 1.47 UWRI points.

Step 6: Psychometric Assessment

Participants

Retrospective chart review of 513 patients from November 2012-October 2017 was used to conduct a psychometric assessment. Standard clinic procedure administered the UWRI and LEFS during the initial encounter and collected age, gender, running experience and current running volume (TABLE 1.1). The UWRI and GROC were completed electronically at least 8 weeks after the initial encounter. Follow-up data is limited to 50 patients because a systematic process was not implemented until May 2017.

Item analysis

Responses to each UWRI item are spread across the continuum of potential responses and the distribution of the composite UWRI scores is centered at the middle value (TABLE 1.2). All items are correlated with the composite UWRI score at the initial (*r* = 0.51-0.72) and final (*r* = 0.55-0.86) assessments. Each item is responsive to change, and the item-specific change is correlated with the change in composite UWRI score (*r* = 0.63-0.85).

TABLE 1.2. Item analysis of University of Wisconsin Running Injury and Recovery Index from Step 6: Psychometric Assessment

Item	Initial	Final	Change
1	3.12 ± 0.89	3.80 ± 0.63	0.16 ± 1.78
2	1.51 ± 1.10	2.50 ± 1.18	0.76 ± 1.23
3	1.85 ± 1.04	2.70 ± 0.82	1.04 ± 1.14
4	2.39 ± 1.00	3.33 ± 0.50	0.60 ± 0.91
5	2.53 ± 0.95	3.00 ± 0.67	0.52 ± 1.05
6	1.44 ± 1.13	2.30 ± 1.16	0.92 ± 1.12
7	1.50 ± 1.21	2.20 ± 1.23	1.16 ± 1.21
8	1.97 ± 1.26	3.10 ± 0.74	0.60 ± 1.22
9	1.23 ± 1.07	2.60 ± 1.26	1.44 ± 1.63
Score	17.71 ± 6.14	27.07 ± 6.63	7.37 ± 7.84

Values are mean ± SD

Internal consistency

The exploratory factor analysis revealed 51% of the variance was explained by two, moderately correlated factors (*r* = 0.46). Factor 1 (items 6-8) explained 26% of the variance and factor 2 explained 25% of the variance (TABLE 1.3). Cronbach's alpha for the UWRI composite score was 0.82 (TABLE 1.4).

TABLE 1.3. University of Wisconsin Running Injury and Recovery Index factor analysis		
UWRI Item	Factor 1	Factor 2
1: Injury impact on daily tasks	-0.04	0.53
2: Frustration due to injury	0.25	0.51
3: Perceived recovery	-0.21	0.71
4: Pain during running	0.03	0.66
5: Pain after running	-0.08	0.69
6: Weekly running volume	0.95	-0.05
7: Typical run distance	0.95	-0.06
8: Running pace	0.68	0.10
9: Confidence to increase training	0.09	0.55
Items from Step 6: Psychometric Assessment initial evaluations Values are eigenvalues		

Construct validity

In addition to the initial UWRI, 50 people completed the GROC and second UWRI. A statistically significant, moderate correlation was observed between the change in UWRI score and the patient-reported change on the GROC, r = 0.61, 95% CI [0.4, 0.76] ($P < 0.001$). A paired samples t-test showed a statistically significant difference between the initial and final UWRI scores ($P < 0.001$). The correlation of the change in UWRI and LEFS scores was not calculated due to the limited quantity of paired data.

Responsiveness

Individuals reporting significant clinical improvement demonstrated greater change UWRI score than those reporting slight clinical improvement ($P = 0.01$) (TABLE 1.4).[94] At the initial assessment, the maximum score was achieved in 3/451 (0.21%) and 44/354 (12.43%) participants completing the UWRI and LEFS, respectively, whereas the minimum UWRI or LEFS score was not achieved. At the initial evaluation, the average LEFS score was 70.75 out of 80 possible points.

TABLE 1.4. Psychometric properties of the University of Wisconsin Running Injury and Recovery Index

	Step 4 Preliminary Evaluation	Step 5 Reliability Assessment	Step 6 Psychometric Assessment
<u>Item Analysis</u> Range of item response-total score Pearson correlation coefficient	*r* = 0.36-0.81	*r* = 0.47-0.82	*r* = 0.51-0.72
<u>Internal Consistency</u> Cronbach's alpha (95% CI)	0.75 (0.28, 0.93)	0.83 (0.76, 0.89)	0.82 (0.80, 0.84)
Factor 1	0.81 (0.71, 0.93)	0.88 (0.81, 0.94)	0.89 (0.87, 0.91)
Factor 2	0.72 (0.56, 0.88)	0.74 (0.62, 0.85)	0.78 (0.74, 0.80)
<u>Reproducibility</u>			
$ICC_{(agreement)}$ (95% CI)	0.56 (0.16, 0.81)	0.93 (0.89, 0.96)	NA
$SEM_{(agreement)}$	4.46	1.45	NA
<u>Construct Validity</u> UWRI change – GROC Pearson correlation coefficient (95% CI)	*r* = 0.75 (0.39, 0.91)	NA	*r* = 0.61 (0.4, 0.76)
<u>Responsiveness</u>			
UWRI mean ± SD if significant improvement†	13.33 ± 4.93	NA	9.38 ± 5.78
UWRI mean ± SD if slight improvement†	10.42 ± 10.31	NA	4.5 ± 7.55
Ceiling Effect‡	UWRI 0% LEFS 25%	UWRI 0%	UWRI 0.21% LEFS 12.43%
Floor Effect‡	UWRI 3.45% LEFS 0%	UWRI 0%	UWRI 0% LEFS 0%

Abbreviations: LEFS, Lower Extremity Functional Scale; GROC, Global Rating of Change; ICC, intraclass correlation coefficient; SEM, standard error of measurement; CI, confidence interval; NA, not available.

† Significant improvement = GROC +4 or +5; slight improvement = GROC +2 of +3

‡ Ceiling and floor effects are calculated as the percent of respondents achieving the maximum and minimum composite scores, respectively.

Discussion

The purpose of this study was to develop a valid, reliable, and responsive evaluative PRO measure to assess longitudinal change in running ability after RRI for clinical practice and research applications. After sustaining a RRI, runners monitor symptoms and manage the dynamic psychological response during physiologic healing and the progressive physiologic adaptation to restore pre-injury performance. The 9-item UWRI accounts for how runners assess running ability through two components, running progression and symptom surveillance. Running progression (items 6-8) involves assessing different aspects of running through weekly volume, long run distance, and running pace which are unique components in the load application framework proposed by Bertelsen and colleagues.[20] Symptom surveillance incorporates how runners monitor symptoms while running (item 4), assess training response (items 1, 3, 5), and describe the psychological response (items 2, 9). Running performance and experience creates disparity in what is considered normal running, but all levels of runners monitor symptoms to make informed training decisions.

Elite and competitive runners make a significant investment in their running careers and described elaborate, individualized schema for symptom monitoring and profound psychological responses following injury. Non-elite and recreational runners described succinct symptom monitoring processes and greater willingness to abstain from running. Commonalities in symptom surveillance themes were illuminated during item reduction. Predictably, runners monitor symptoms while running, but runners of all performance levels frequently monitor symptoms between running sessions to determine readiness for future training. Psychological responses are inherently individualized; however, frustration is the common emotional theme runners experience during recovery. The progressive increase in an athlete's post-injury confidence is known to play an important role in the successful return to sport and runners consistently reported their running ability was fully restored when they established the confidence to train without fear of re-injury.[96 97] The UWRI is a parsimonious, clinically-relevant measure reflecting runners' self-assessment of running ability and how they subsequently make decisions during the RRI recovery process.[98]

The iterative development process incorporated the target population and produced a sample with injury and runner characteristics that are consistent with other RRI studies. [6 72 99] Including the target population enhanced the UWRI's content validity and enabled data triangulation by source and method to assure the UWRI items accurately and comprehensively represent the recovery of distance running ability following injury.[76] Participants and expert reviewers concluded the UWRI was composed of comprehensible items that are likely to measure the construct due to the detailed, running-specific items, as well as acknowledging the psychological factors associated with recovery. The 5-item response structure was chosen because it was preferred by the target population and is less burdensome to respondents than visual analog scales.[89] Participant responses show the items differentiate between levels of function when used in a clinical setting and are responsive to change over time. Clinical implementation is enhanced by the low administration burden, short completion time, and simple scoring.

During preliminary testing (step 4), the $UWRI_{beta}$ test-retest reliability was adversely affected by a hard stop question blocking access to items 4 through 9 if the participant had not run during the prior 7 days. Excellent test-retest reliability was observed following modifications to present all items during each administration. The error associated with a single score was less than 2 points in the total UWRI score. It was logical to create a single composite score because the two correlated subscales collectively reflect the typical process runners use to gauge their running ability during recovery which is the primary intent of the UWRI. The internal consistency of the UWRI, and each component factor, was within the accepted range during all clinical evaluations indicating item correlation is present without item redundancy.[100]

Construct-related validity was evaluated by comparing the change observed in the UWRI score with measures commonly used to assess RRIs.[60 76] The UWRI demonstrated the ability to assess changes in self-reported running ability and the change in the UWRI total score was correlated with the changes measured by the GROC. The association between the UWRI and LEFS varied throughout the steps of this study, and the LEFS may not have capacity to respond to functional improvements because the initial evaluation scores are skewed toward the maximum potential score. The observed moderate associations were anticipated because the UWRI was designed to be a more critical assessment of running function than other PRO measures with greater temporal stability than the GROC.[101 102]

In this initial study, the UWRI was found to be responsive to change because individuals reporting clinically meaningful improvements on the GROC scale also reported greater change in UWRI scores than those reporting slight, or no change.[76 103] Heightened responsiveness is further supported by the absence of UWRI ceiling or floor effects. The LEFS demonstrated a ceiling effect during step 4 and approached this threshold in the large dataset used in step 6.[76] LEFS responsiveness is further questioned because the majority of respondents, at the initial evaluation, have a score that is too high to achieve the 9-point minimum detectable change.[60] The iterative process used to develop the UWRI enhanced the specificity to running and produced a PRO measure capable of measuring changes in running ability regardless of the body region or injury.

The strengths of this study include the repeated integration of the target population during item development, psychometric assessment of the UWRI using participants from the target population, and the involvement of multiple stakeholders in the development process. The sample population included a diverse representation of age, experience, running volume, speed, and injury. There are several limitations of this study including participants lost to follow up during prospective steps and the use of a retrospective design to conduct a psychometric assessment with clinical data. Further psychometric evaluation is warranted to evaluate the construct validity in combination with measures of running ability and other PRO measures. Clinically useful assessments of responsiveness including minimal detectable change and minimal clinically important difference will be determined in future studies. Future studies will also seek to provide a better understanding of how the UWRI measures change in running ability for different performance abilities or injury types.

Conclusion

The UWRI is a reliable, evaluative measure assessing running ability following RRI with minimal administrative burden. Content and construct-related validity assessments indicate the UWRI is a more exacting evaluation of running ability than other PRO measures currently used in this population. Continued psychometric evaluation in prospective, clinical studies is warranted because this novel measure of running ability is responsive to patient-perceived functional change.

Key Points

Findings: The University of Wisconsin Running Injury and Recovery Index (UWRI) is a novel, running-specific patient-reported outcome measure reflecting how runners assess their own running ability while recovering from a running-related injury. The UWRI is a reliable assessment that can measure the change in running ability during an episode of care.

Implications: The UWRI systematically assesses post-injury distance running ability with very little burden on providers or patients.

Caution: Construct-related validity and responsiveness require additional prospective, psychometric assessment conducted in combination with clinical care.

Chapter 2. Validity and Responsiveness of the University of Wisconsin Running Injury and Recovery Index

Evan O. Nelson, Stephanie Kliethermes, Bryan Heiderscheit

Abstract

Background: The University of Wisconsin Running Injury and Recovery Index (UWRI) is the first running-specific patient-reported outcome measure (PRO). The UWRI evaluates the key elements runners use to self-assess running ability during recovery, a multifactorial experience. This study evaluated the construct-related validity and responsiveness of the UWRI as an evaluative PRO of running ability following running-related injury (RRI).

Methods: Runners seeking care from a physical therapist for a RRI (n=396) completed PROs at baseline and 12 weeks later. UWRI change was validated against the global rating of change (GROC), VR-12 change, and change in body region-specific PROs. Responsiveness was evaluated using anchor-based and distribution-based techniques.

Results: UWRI change (mean±SD 7.7±9.0) was correlated with GROC (r=0.67), as well as the changes in VR-12 physical component score (PCS) (r=0.54) and mental component score (MCS) (r=0.31). UWRI change was correlated with changes in the sport subscale of Foot and Ankle Ability Measure (r=0.75), 12-item international hip outcome tool (r=0.75), and anterior knee pain scale (r=0.48), but not associated with the Oswestry Disability Index 2.0 (r=0.05). UWRI change was significantly different in runners reporting significant improvement (12.2±5.9), slight improvement (7.1±6.6), no change (0.0±9.1), and worsening (-14.6±7.4) on the GROC. The

UWRI minimal important change was 5 points and the minimal clinical important difference was 8 points.

Conclusion: The UWRI demonstrated longitudinal validity (GROC), convergent validity (PCS and body region-specific PROs), and divergent validity (MCS), as well as excellent responsiveness. The UWRI should be used to evaluate patient-perceived running ability during RRI recovery.

Keywords: running-related injury, running injury, patient-reported outcome measure, validation, University of Wisconsin Running Injury and Recovery Index

Background

Running is a popular, accessible form of exercise associated with numerous physical and psychological health benefits that may protect against disabilities later in life.[104-106] However, runners of all experience and performance levels are susceptible to injury with incidence proportions as high as 84.9%.[72 107] Running-related injuries (RRIs) have a have a complex multifactorial etiology, but the causal mechanism remains unclear.[108] A conceptual framework depicts an RRI occurs when the cumulative load per run exceeds the tissue-specific capacity, which is influenced by a number of variables (e.g., age, recovery, diet, previous training).[20] Once injured, runners frequently reduce running, suffer functional limitations, receive medical care, or experience negative psychological consequences.[12 14 29 36 49] Due to the popularity of running and the high RRI incidence, the net injury cost may be significant and injury is the primary reason runners leave the sport.[9]

During RRI recovery, runners manage running-induced load by manipulating running volume and speed until pre-injury performance is restored. During this process, runners monitor physical symptoms to make informed training decisions and may experience fluctuating psychological responses throughout recovery.[35 109] It is very common for runners to continue training despite injury symptoms; in some cases, symptoms do not necessitate a reduction in training volume, and only 6% of injuries prohibit running.[12] Patient-reported outcome measures (PROs) incorporate the patient-perspective into functional assessments, but many PROs include items assessing low demand functional activities which cannot account for symptoms experienced during running.[60 110 111] In the absence of a measure of running ability, studies defined recovery using arbitrary distance criteria and used activities of daily living PROs to assess the efficacy of RRI interventions.[14 45]

The University of Wisconsin Running Injury and Recovery Index (UWRI) is a 9-item running-specific, evaluative PRO assessing running ability following RRI.[109] Runners and providers participated in an iterative process to assure the UWRI includes the key elements injured runners use to assess running ability. The UWRI has excellent test-retest reliability; however, its validity as a measure of longitudinal change and responsiveness to changes in running ability are unknown. As this is a new type of PRO, global assessments of change are recommended to determine the meaningfulness of the change measured by the UWRI.[112] Understanding the validity and responsiveness of the UWRI will establish its utility as an evaluative PRO. Monitoring RRI recovery using valid assessments of the multifactorial patient experience will substantially advance our ability to determine an intervention's clinical benefit and enable comparison of treatment effects.

The purpose of this study was to evaluate the construct-related validity and responsiveness of the UWRI as an evaluative PRO of running ability following RRI.

Methods

Design

A prospective cohort study was conducted in accordance with COnsensus-based Standards for selection of health Measurement INstruments (COSMIN) guidelines in five physical therapy organizations (1 academic medical center, 1 comprehensive orthopedic center, and 3 physical therapist owned private practices) in the United States.[113] Physical therapy patients were recruited through the organizations' running clinics and invited to participate during an office visit when RRI care began. A physical therapist confirmed the presence of an RRI, which was defined as a musculoskeletal problem resulting from running that required the individual to (1) prematurely stop running while training or during a competitive event, or (2) to miss a training session or competitive event.[73] Participants (≥14 years) with at least one musculoskeletal RRI of the lumbopelvic region or lower extremity who were able to read English were eligible for inclusion. Participants who had received surgery at the site of the injury in the past 12 months were excluded. Eligible participants were informed of the study purpose and procedures and included in the study after providing informed consent. The minimal risk institutional review board at the University of Wisconsin-Madison (submission ID #2016-1544) approved the study.

Data Collection

Study data was collected using Research Electronic Data Capture (REDCap™) hosted at the University of Wisconsin-Madison.[114] Participants completed the study questionnaires electronically during the initial office visit (baseline). Participants then received an email containing a link to the second questionnaire 12 weeks after the initial questionnaire (follow-up), and up to three reminder messages were sent to non-responders. The physical therapist provided the injury location, diagnosis, and injury acuity.

Participants self-selected their running status, defined as novice: new to the sport of running, recreational: regular running with or without race participation for enjoyment or health and wellness purposes, competitive: purposeful training with the intent to improve performance without financial gain, and professional: purposeful training to improve performance with financial gain. Participant characteristics included height, weight, body mass index, age, years of running experience. Participation in a running club or team, coaching, type of shoe, and orthotic use were also recorded. Self-reported running metrics (i.e., weekly volume, long run distance, typical run distance, pace, and frequency) were provided for the pre-injury state, current at baseline, and current at follow-up.

Measures

The UWRI was used to measure running ability; scores range from 0-36 with higher scores indicating greater running ability. The test-retest reliability $ICC_{agreement}$ is 0.934.[109] In the absence of a gold standard, multiple comparator PROs used to assess convergent and divergent construct-related validity as has been done in other measures of lower extremity functional impairment.[60]

The Veterans RAND 12-item Health Survey (VR-12) is a generic health-related quality of life (HRQoL) measure assessing 8 domains of health which are used to calculate physical (PCS) and mental (MCS) component scores; higher scores represent better HRQoL.[115] The VR-6D is a preference-based utility measure derived from select VR-12 items; the utility values describe health states on a scale where 0 equates to death and 1.0 is perfect health.[54] Body-region specific PROs were selected based on their previous use in RRI research, acceptance in clinical practice, or application to the construct; the Oswestry Disability Index 2.0 (ODI), International Hip Outcome Tool 12-item (iHOT-12), anterior knee pain scale (AKPS), Foot and Ankle Ability Measure sport-specific subscale ($FAAM_{Sport}$).[61 62 111 116 117] All body-region PROs are scored 0-100 with higher scores indicating better function. Body-region specific PROs were completed based on the participant's reported area(s) of injury and analyzed according to the PT-reported injury location. A detailed description of the comparison measures is available. (SUPPLEMENT 2.1)

The 11-item global rating of change (GROC) is a valid, reliable method to assess patient-perceived change (-5="very much worse", 0="unchanged", +5="very much improved") during anchor-based evaluations of musculoskeletal conditions.[79 80] GROC ratings were consolidated to significant improvement (+4,+5), slight improvement (+2,+3), no change (-1 to +1), or worsened (≤-2).[103]

Data analysis and statistics

Structural validity

A description of the item analysis and exploratory factor analysis are provided in SUPPLEMENT 2.2.

Reliability

Internal consistency was calculated using coefficient alpha (Cronbach's α) applied to baseline data, and considered adequate if the value was between 0.70 and 0.95.[100] Standard error of measurement (SEM = $SD_{observed}$ * 1-*r*) was calculated using the standard deviation of the observed scores and the test-retest correlation acquired from published data.[61 62 109 110 117]

Construct validity

Construct validity was evaluated by establishing *a priori* hypotheses about the association between the UWRI and comparator PRO measures, expecting confirmation of at least 75%.[76] In this study, validity aligns with responsiveness because evaluative outcome measures are intended to measure longitudinal change. Pearson correlation coefficients were considered good $|r| > 0.75$, moderate $0.50 < |r| \leq 0.75$, fair $0.25 < |r| \leq 0.50$ and no association $|r| < 0.25$.[81] The Pearson correlation coefficient was used to assess convergent and divergent validity according to the following *a priori* hypotheses:

(H1) The change in UWRI score will demonstrate a good, positive correlation with the GROC.

(H2) The change in UWRI score will demonstrate a moderate, positive correlation with the change in ODI score, iHOT-12 score, AKPS score, and FAAM score.

(H3) The change in UWRI score will demonstrate a moderate correlation with the change in PCS.

(H4) The change in UWRI score will demonstrate a weak correlation with the change in MCS.

Responsiveness

Applying Beaton's taxonomy, within-person change over time was defined as change; whereas, difference described between-group differences of within-person change.[118] The ability of the measure to characterize change, internal responsiveness, was evaluated using the paired t-test, Cohen's d effect size ($ES = (\overline{X}_2 - \overline{X}_1)/sd(\overline{X}_1)$), standardized response mean difference ($SRM = (\overline{X}_2 - \overline{X}_1)/sd(\overline{X}_2 - \overline{X}_1)$), the responsiveness index ($RI = (\overline{X}_2 - \overline{X}_1)/sd(\overline{X}_2 - \overline{X}_{1_{stable\ subjects}})$).[119 120] Cohen's effect size estimates of 0.20, 0.50, and 0.80 were considered small, moderate, and large, respectively.[119]

External responsiveness reflects how measured changes relate to corresponding changes in the GROC. The Pearson correlation coefficient determined the association of change in a PRO and the GROC. Anchor-based assessments evaluate the mean change in a PRO score for individuals reporting significant improvement (GROC +4 or +5) and slight improvement (GROC +2 or +3) using a "between-subjects", and "within-subjects" approach, respectively.[79 82] Between group differences were assessed using a one way ANOVA and post hoc Tukey HSD. Receiver operating characteristic (ROC) curves were constructed for clinically meaningful GROC criteria (-2, +2, +4) in a sensitivity and specificity-based approach for external responsiveness. The area under curve (AUC) was compared and the threshold for meaningful UWRI change was identified.[121 122]

Responsiveness calculations were performed for all PROs; the UWRI responsiveness was evaluated using the following a priori hypothesis:

(H5) The mean UWRI change for those reporting significant improvement, slight improvement, and no change will be significantly different.

Interpretability

For individual change, the minimal detectable change at 95% confidence (MDC_{95}) was estimated using the equation $MDC_{95}=1.96 * \sqrt{2} * SEM$.[76] For group-based changes, the MDC_{95} calculation included sample size, $MDC_{95}=1.96 * \sqrt{2} * SEM/\sqrt{n}$.[76] Minimal clinically important difference (MCID) is 'the smallest difference in score which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management.[80] The MCID was estimated by calculating the mean change in UWRI score for individuals reporting significant change and by the optimal threshold value from the ROC constructed using a threshold of GROC ≥+4. A similar strategy was used to determine the smallest amount of change patients perceive as important, minimal important change (MIC). The MIC was estimated as the mean UWRI change score for individuals reporting slight improvement and by the optimal threshold value from the ROC constructed using a threshold of GROC ≥+2. Floor and ceiling effects were considered present if at least 15% of respondents scored the lowest or highest possible scores, respectively.[76]

(H6) The MCID will exceed the MDC_{95}, and the responsiveness index will be greater than 1.96.

Missing data were analyzed for detectable patterns. UWRI item responses deemed to be missing completely at random were imputed as the median value of the respondents available items; whereas, data not missing at random were reported as missing. Missing VR-12 items were imputed using the published scoring algorithm.[123] Censoring occurred for new injuries unrelated to running and loss to follow-up. Subjects with a baseline UWRI collected <8 days from enrollment and not censored were considered complete. Demographic data from complete data were compared to the overall sample using Mann-Whitney U tests for continuous data and Pearson Chi Squared tests for categorical data. All analyses were performed using R software (Vienna, Austria), using the psych, irr, psychometric, and pROC packages.[83-87 124]

Results

This prospective cohort study enrolled 396 participants; demographic data is presented in TABLE 2.1. Due to administrative error at a single center, the initial UWRI was not collected in 122(30.1%) instances. Censoring occurred for new injuries unrelated to running (n=5) and loss to follow-up (n=167). Complete data was available from 157(39.6%) of participants; race/ethnicity (*P*=0.03) was the only significantly different demographic characteristic. Baseline

TABLE 2.1. Participant demographics			
	Overall (n=396)	Complete Cases (n=157)	p-value
Gender, n(%)			0.67
Male	145(37.6)	57(36.3)	
Female	246(62.1)	93(59.2)	
Unanswered	5(1.3)	1(0.6)	
Status, n(%)			0.33
Novice	9(2.3)	2(1.3)	
Recreational	256(65.2)	94(59.9)	
Competitive	130(33.3)	54(34.4)	
Professional	1(0.2)	1(0.6)	
Age (yrs)	34.5 ± 12.4	35.3 ± 12.5	0.27
Ethnicity, n(%)			0.03
White, not Hispanic or Latino	340(85.9)	132(84.1)	
White, Hispanic or Latino	26(6.6)	9(5.7)	
Other, not white	22(5.6)	5(3.2)	
Not reported	6(1.5)	5(3.2)	
Body mass index (kg/m^2)	23.3 ± 3.9	23.4 ± 3.5	0.83
Total diagnoses	445	194	0.54
Injury location, n(%)			0.33
Low back or pelvis	24(5.4)	14(8.9)	
Hip	66(14.8)	29(18.5)	
Thigh or hamstring	30(6.7)	5(3.2)	
Knee	132(29.7)	43(27.4)	
Calf or shin	71(16)	31(19.7)	
Ankle	61(13.7)	28(17.8)	
Foot	61(13.7)	26(16.6)	
Injury duration, n(%)			0.42
<1 month	28(7.1)	10(6.4)	
1-2 months	66(16.7)	20(12.8)	
3-6 months	135(34.1)	60(38.2)	
7-11 months	57(14.4)	22(14.0)	
1-2 years	70(17.7)	22(14.0)	
3-5 years	22(5.6)	9(5.7)	
6+ years	17(4.3)	8(5.1)	
Coached, n(%)	77(19.4)	31(19.7)	0.74
On a team, n(%)	117(29.5)	45(28.7)	1
Prior running injury, n(%)	214(54.0)	79(19.2)	0.67
Experience (yrs)	12.7 ± 10.4	13.4 ± 10.5	0.20

PRO scores for complete data did not differ from the overall sample. Physical therapists provided 495 RRI diagnoses with 86(19.3%) participants having two or more diagnoses. TABLE 2.2 shows RRIs reduced training load; running increased at the follow-up, but still remained below pre-injury levels.

TABLE 2.2. Training load metrics

	Pre-injury	Baseline	Follow-up
Weekly volume (km/week)*	38.1 ± 23.8	20.2 ± 18.9	24.8 ± 22.0
Pace (m/s)*	0.32 ± 0.05	0.33 ± 0.09	0.33 ± 0.1
Typical distance (km)*	10.5 ± 10.2	6.8 ± 8.8	7.6 ± 9.1
Long run distance (km) *	-	9.7 ± 10.3	10.9 ± 9.7
Frequency (days/week)†	4(0-7)	3(0-7)	3(0-7)

* Values are mean ± SD
† Values are median(range)

The PRO composite scores are presented in TABLE 2.3. Using the GROC, 110(48.2%) individuals reported significant improvement, 85(37.3%) slight improvement, 26(11.4%) unchanged, and 7(3.1%) worsened. Due to the low occurrence of worsened condition, slight and significant worsening were not differentiated during analysis. The mean(SD) change in UWRI score was 7.1(9.0). Combining baseline and follow-up assessments, the UWRI proportion of missing items was 0.1%. The proportion of missing items on the other PRO measures was 0.1-1.8%.

Structural validity

The change in UWRI score is explained by 2 correlated factors (r=0.79) with items 6-8 loading on factor representing load management and the remaining items representing symptom surveillance. The results of the item and factor analyses are available (SUPPLEMENT 2.2).

TABLE 2.3. Patient-reported outcome measure scores							
	Baseline		Follow-up		Change		
	n	Composite score*	n	Composite score*	n	Mean difference*	*p*-value†
UWRI	274	16.5 ± 6.7 (0, 34)	207	24.9 ± 8.0 (0, 36)	152	7.7± 8.9 (-23, 28)	<0.001
VR-12 PCS	281	50.7 ± 7.0 (24.0, 66.1)	207	53.8 ± 5.5 (18.5, 63.3)	157	2.0 ± 7.5 (-38.9, 22.1)	0.001
VR-12 MCS	281	51.8 ± 9.7 (14.6, 69.6)	207	51.7 ± 48.3 (16.1, 66.7)	157	-0.7 ± 7.7 (-21.7, 24.9)	0.24
VR-6D	281	0.76 ± 0.09 (0.37, 0.97)	207	0.79 ± 0.08 (0.51, 0.97)	157	0.01 ± 0.08 (-0.30, 0.27)	0.03
ODI	35	82.8 ± 8.8 (52, 100)	18	84.4 ± 4.4 (76, 90)	18	-2.0 ± 5.1 (-11.8-8.0)	0.11
iHOT-12	115	68.8 ± 17.9 (32.6, 97.6)	67	78.1 ± 18.5(28.6, 100)	65	9.0 ± 19.0 (-39.4, 50.4)	0.001
AKPS	150	81.9 ± 10.6 (45, 100)	83	82.2 ± 8.4 (48, 90)	83	-1.7 ± 9.9 (-45, 34)	0.13
FAAM Sport	182	75.2 ± 20.8 (0, 100)	111	84.1 ± 20.5 (0, 100)	108	8.5 ± 26.7 (-100, 90)	0.001
GROC	NA				223	3.1 ± 1.7 (-5, +5)	

UWRI, University of Wisconsin Running Injury and Recovery Index; VR-12 PCS, Veterans Rand 12 item physical composite score; VR-12 MCS, Veterans Rand 12 item mental composite score; VR-6D, Veterans Rand 6D; ODI, Oswestry Disability Index 2.0; iHOT-12, International Hip Outcome Tool 12 item; AKPS, Anterior Knee Pain Scale; FAAM Sport, Foot and Ankle Ability Measure sport subscale; GROC, global rating of change; NA, not applicable.

* Values are mean ± SD (min, max)

†*P*-value of paired t-test

Reliability

The UWRI had good internal consistency with an overall Cronbach's α of 0.86, as well as the load management (α =0.91) and for symptom surveillance (α =0.79) factors. The UWRI standard error of measurement (SEM) was 1.7; whereas the SEM for the ODI was 2.6, iHOT-12 was 5.9, AKPS was 3.2, and $FAAM_{sport}$ was 7.5.

Construct validity

(H1) The change in UWRI score demonstrated a moderate, positive correlation with the GROC (r = 0.67; 95% CI [0.57, 0.75]; P<0.001).

(H2) The change in UWRI score demonstrated a moderate, positive correlation with the change score of the iHOT-12 (r=0.75; 95% CI [0.55, 0.87]; P<0.001) and AKPS (r=0.48; 95% CI [0.21, 0.68]; P=0.001), and $FAAM_{sport}$ (r=0.75; 95% CI [0.63, 0.84]; P<0.001). The UWRI and ODI change scores were not associated (r=0.05; 95% CI [-0.57, 0.63]; P=0.89).

(H3) As hypothesized, the change in UWRI score demonstrated a moderate correlation with the change in PCS (r=0.54; 95% CI [0.42, 0.64], P<0.001).

(H4) The change in UWRI score demonstrated a fair correlation with the change in MCS (r=0.31; 95% CI [0.16,0.44], P<0.001).

Responsiveness

The mean difference from the baseline to follow-up assessment is statistically significant for all PRO's, except the MCS, AKPS, and ODI (TABLE 2.4). The UWRI had a large effect size (Cohen's d=1.15); the standardized response mean difference was 0.87. The iHOT-12 had a moderate effect size, but small effect sizes were observed for the other PROs (TABLE 2.4).

(H6) The responsiveness index for the UWRI was 1.33, lower than the hypothesized value. TABLE 4 provides the responsiveness indices for the comparator PROs.

(H5) As hypothesized, participants reporting clinically significant improvement (GROC +4 or +5) had a mean(SD) UWRI change of 12.2(5.9); whereas, the UWRI change for participants reporting no change (GROC -1 to +1) was 0.0(9.1) (P<0.001) (TABLE 4). The mean change in UWRI score was 7.1(6.6) for participants reporting slight improvement (GROC +2 or +3) and was significantly different from those reporting significant improvement (P<0.001) or no change (P=0.002). The UWRI change for participants reporting worsening (GROC ≤-2) was -14.6(7.4), and significantly different than those reporting no change (P=0.006). TABLE 2.4 presents the anchor-based responsiveness assessment comparing all PROs to the GROC.

The ROC curve set to define significant clinical improvement (GROC +4) determined 7.5 points of change in the UWRI score had 61.0% specificity and 78.6% sensitivity to properly classify improvement (AUC 77.5%; 95% CI [70.2, 84.8]). The ROC curve set to define slight improvement (GROC +2) determined 2.5 points of change in the UWRI score had 88.0% sensitivity and 70.4% specificity (AUC of 85.0; 95% CI [75.6, 94.3]) to determine improvement. AUC differences between the levels of improvement were not statistically significant; however, significant differences in AUC were present for slight improvement (GROC +2) vs. worsening (GROC -2) (P=0.005).

TABLE 2.4. Longitudinal validity and responsiveness									
	Effect size (Cohen's *d*)	Standardized response mean difference	Responsiveness index	Correlation of PRO change to GROC (*r* [95% CI])	Significant improvement[*] (Mean difference ± SD)	Slight improvement[*] (Mean difference ± SD)	No change[*] (Mean difference ± SD)	Worsening[*] (Mean difference ± SD)	Main effect[†] (*p*-value)
UWRI	1.15	0.87	1.33	0.67 [0.57, 0.75]	12.2 ± 5.9	7.1 ± 6.55	0.0 ± 9.1	-14.6 ± 7.4	<0.001
VR-12 PCS	0.28	0.26	0.45	0.40 [0.26, 0.52]	3.0 ± 5.5	3.6 ± 7.0	-2.0 ± 6.7	-13.8 ± 15.8	<0.001
VR-12 MCS	-0.07	-0.09	0.08	0.20 [0.05, 0.35]	0.5 ± 46.7	-1.5 ± 9.1	-1.7 ± 6.2	-5.8 ± 8.4	0.17
VR-6D	0.16	0.17	0.50	0.49 [0.37, 0.60]	0.04 ± 0.06	0.02 ± 0.08	-0.04 ± 0.07	-0.16 ± 0.13	<0.001
ODI	0.23	0.4	0.07	-0.06 [-0.51, 0.42]	0.1 ± 5.4	3.8 ± 5.0	-1.0 ± 1.4	none	0.36
iHOT-12	0.50	0.47	0.76	0.42 [0.20, 0.60]	15.5 ± 13.8	9.1 ± 19.3	-11.2 ± 20.4	-14.6 (n=1)	0.007
AKPS	-0.16	-0.17	0.03	0.47 [0.28, 0.62]	0.2 ± 9.8	-0.9 ± 5.8	-7.0 ± 6.1	-25.3 ± 18.6	<0.001
FAAM$_{Sport}$	0.41	0.32	0.34	0.53 [0.39, 0.66]	14. 7± 13.8	12.2 ± 21.0	-5.7 ± 43.1	-64.1 ± 23.2	<0.001

UWRI, University of Wisconsin Running Injury and Recovery Index; VR-12 PCS, Veterans Rand 12 item physical composite score; VR-12 MCS, Veterans Rand 12 item mental composite score; VR-6D, Veterans Rand 6D; ODI, Oswestry Disability Index 2.0; iHOT-12, International Hip Outcome Tool 12 item; AKPS, Anterior Knee Pain Scale; FAAM$_{Sport}$, Foot and Ankle Ability Measure sport subscale.

[*]Significant improvement = GROC +4 or +5; Slight improvement = GROC +2 or +3; No change = GROC -1, 0, +1; Worsening = GROC -2 or lower

[†]1-way analysis of variance (ANOVA) evaluated differences between levels of change

Interpretability

The MDC_{95} (individual) for the UWRI was 4.6, and MDC_{95} (group) was 0.3. The UWRI MIC is 5 points, and the MCID is 8 points. Floor or ceiling effects were not present in the UWRI. TABLE 2.5 provides interpretability data for the comparator PROs.

TABLE 2.5. Interpretability

	MDC_{95} individual	MDC_{95} group	Ceiling (%)	Floor (%)
UWRI	4.7	0.3	3.7%	0%
ODI	7.3	1.2	4.9%	0%
iHOT-12	16.5	1.5	3.9%	0%
AKPS	8.8	0.7	0.5%	0%
FAAM Sport	20.7	0.7	24.4%	0%

UWRI, University of Wisconsin Running Injury and Recovery Index; VR-12 PCS, Veterans Rand 12 item physical composite score; VR-12 MCS, Veterans Rand 12 item mental composite score; VR-6D, Veterans Rand 6D; ODI, Oswestry Disability Index 2.0; iHOT-12, International Hip Outcome Tool 12 item; AKPS, Anterior Knee Pain Scale; FAAM Sport, Foot and Ankle Ability Measure sport subscale; MDC, minimal detectable change

Discussion

The results of this study show the UWRI is a valid measure of running ability that is responsive to longitudinal changes following RRI. The UWRI demonstrated superior measure properties to the comparison PROs used in this study.

Structural validity

At all assessment points, each UWRI item makes a unique and meaningful contribution to the composite score and all UWRI items were responsive to change. The load management and symptom surveillance factors included in this assessment of longitudinal change are consistent with the structure found in development using only baseline assessments. This structure demonstrates face validity with runners' self-assessment of running ability following injury.[125]

Internal consistency

The internal consistency of the UWRI and each factor was acceptable.

Validity

This study applied the recommended strategy to assess longitudinal validity of a novel PRO; the UWRI demonstrated a moderate, rather than good, association with the GROC.[76 80 82] The *a priori* hypothesis was established using pilot data acquired with a shorter follow-up duration. Temporal instability of the GROC may explain the lower strength of association because retrospectively evaluating the entire duration of recovery biases responses to represent the most recent status, especially over longer follow-up periods.[101 126] This consideration is present in the longitudinal validity assessment for all PROs and the change UWRI score was more strongly correlated with overall change.

The UWRI demonstrated construct-related validity by satisfying the *a priori* hypotheses for convergent and divergent validity. Despite the overlap in content between PROs, moderate associations of convergent validity were expected because the UWRI is intended to be a more critical assessment of running ability. The comparator PROs contain items assessing activities of reduced physical demand or unrelated to running, reducing the application to the construct of running ability. Very few RRI studies include general HRQoL measures, but recently Messier et al found the SF-12 MCS, an equivalent tool to the VR-12, was predictive of RRI.[10] Based on these findings, the MCS does not appear to be a valid measure for determining treatment response or assessing recovery following RRI.

Responsiveness

Distribution-based assessments show the UWRI was the most responsive to changes in running ability. The effect sizes of the comparator PROs were smaller than previously reported.[61 62 116 117] External responsiveness anchored the UWRI change against the GROC to provide more meaningful responsiveness estimates.[119] It is critical for an evaluative PRO measure to detect change exceeding measurement error and differentiate small amounts of improvement from clinically important improvement.[112] The UWRI was the only PRO to have statistical significant differences at all levels of change (e.g., significant improvement, slight improvement, no change, and worsening) and the mean UWRI score showed no difference in running ability in the group reporting no change in the external criterion. Score plateau despite improved running ability, decreased PRO score with unchanged running ability, and failure to detect change summarize the responsiveness limitations of the comparison PROs.

Interpretability

The MCID is a critical value for clinical implementation, which is the subject of debate and uncertainty within the clinical research community.[127 128] In this study, the sensitivity- and specificity-based approach was used to establish the MCID. Individuals exceeding the desired minimum standard are included in the within-subjects approach because the change of all people reporting significant improvement is averaged.[80 82] The physical therapists involved with this study are clinical experts working at specialized running clinics who may achieve better patient outcomes; greater improvement in running ability would further increase the anchor-based MCID estimate. Alignment of the MDC_{95} and between-subjects approach indicates a change ≥5 points in UWRI score should be considered meaningful, as it reflects a change in the individual's running ability exceeding measurement error.

Limitations

The most significant limitation is the large proportion of people who were lost to follow-up or had missed PRO measures; however, this problem is common in longitudinal PRO studies and a recent PRO validation study included longitudinal data from 51% of the sample.[129 130] The primary recruitment center implemented a process for collecting PROs in the electronic health record coinciding with the start of enrollment. Capturing study and clinical information using parallel systems created significant barriers to data collection.[131] Self-reported running metrics are inaccurate and future work should incorporate objective assessments of training load that may be used to assess the longitudinal change in running ability.[132]

Conclusion

The UWRI is a valid and reliable measure of running ability following RRI. The UWRI measured the recovery of injured runners receiving physical therapy care better than PROs not specifically designed to assess running ability.

SUPPLEMENT 2.1. Description of outcome measures

The Veterans RAND 12-item Health Survey (VR-12) was developed from the Veterans RAND 36 Item Health Survey (VR-36), which was developed from the MOS RAND SF-36 Version 1.0. The 12-items in the VR-12 evaluate eight domains of health-related quality of life: general health, physical functioning, role physical, role emotional, bodily pain, vitality, mental health, social functioning. The domains are used to calculate physical (PCS) and mental (MCS) component scores with higher values representing greater health-related quality of life.[115] Component scores are normalized to a mean±SD of 50 ± 10 and US population standards are 39.8 and 50.2 for the PCS and MCS, respectively.[133] Individual change of greater than 6.5 or 7.9 units in the PCS or MCS, respectively, are considered clinically relevant.[134]

The VR-6D is a preference-based utility measure derived from select VR-12 items, assessing 6 domains: physical functioning, role limitations (physical and emotional), social functioning, pain, mental health, vitality.[54] The VR-6D utility value reflects preferences for health states on a scale where 0 equates to death and 1.0 is perfect health.[54]

The Oswestry Disability Index 2.0 (ODI) is a 10-item PRO validated for use with a wide array of patients with lumbar conditions. A percentage index is the sum of patient responses divided by the maximum potential score.[116] Reverse item scoring was used in this study so higher scores indicate greater ability. This study did not include item 8 because the respondent’s sexual activity is not related to the study content and is potentially offensive to participants. ODI has good reliability ($ICC_{2,1} = 0.91$) and composite score changes exceeding 9-15 points have been considered significant individual change in previous studies.[116]

The International Hip Outcome Tool 12-item (iHOT-12) was developed to assess clinical status in patients with hip pathology, especially younger, more active patients for whom other hip-specific tools had limited responsiveness.[117] Item responses are recorded on a 0-100 slider bar and the composite score is created from the arithmetic average of the responses. Higher scores indicate greater functional ability with changes greater than 13 points indicating significant change in patients undergoing hip arthroscopy for intra-articular pathology.[135] Item 9 evaluates sexual activity and was not included the study questionnaires. The ihot-12 test-retest reliability is good ($ICC_{2,1}$ = 0.89).[117]

The anterior knee pain scale (AKPS) is a 13-item self-reported questionnaire validated for use in people with knee disorders and demonstrates comparable measurement ability to the Lower Extremity Functional Scale.[111] The AKPS has good reliability ($ICC_{2,1}$ = 0.91) and is an accepted measure for the most common type of running-related injury, anterior knee pain.[61] Item responses have disparate values that when added in a simple sum produce a weighted composite score. The range of potential scores is 0-100 with higher scores indicating greater ability and a minimal detectable change of 13 points.[61]

The Foot and Ankle Ability Measure includes a sport-specific subscale ($FAAM_{Sport}$) validated for use without the complementary activities of daily living subscale.[62] The $FAAM_{Sport}$ is an 8-item self-reported questionnaire that has good test-retest reliability ($ICC_{2,1}$ = 0.87) and is validated for a broad range of leg, ankle, and foot conditions.[62] Individual item scores (0-4) are used to construct an index composite score (0-100%) with higher scores indicating greater sport function without regard for a specific sport and a 9-point increase indicates significant clinical improvement.[62]

SUPPLEMENT 2.2. Structural validity

Methods

Structural validity

The item examination included examining distribution of UWRI item responses at baseline, follow-up and the change in item response. Classical test theory methods were used, measuring the item-total score Pearson correlation and item-rest Pearson correlation for the item-specific change.

An exploratory factor analysis was performed to assess the unidimensionality of the change score. A maximum likelihood regression analysis with promax rotation was performed. A sample size of 10 per item is recommended for factor analysis with a total sample of at least 100 responses.[76] A linear model using the identified factors was created to predict the GROC score.

Results

Item responses are distributed across the range of response options and demonstrate change (TABLE S2.2.1. Item analysis revealed all items demonstrate a strong association with the UWRI composite score at the initial (*r*=0.45-0.79) and final (*r*=0.68-0.88) assessments. Strong item-rest correlations are present for all items at all assessment points including change scores (*r*=0.63-0.79). TABLE S2.2.2 presents the results of the factor analysis. In univariate models, load management (β=0.31; P<0.001; adj. R^2=0.40) and symptom surveillance (β=0.22; P<0.001; adj. R^2=0.34) were significant predictors of GROC response. In the multivariate model, load management (β=0.26; p<0.001) was significant while symptom surveillance (β=0.07; P=0.1) was not. The adjusted R^2 estimated the multivariate model explained 55% of the variance in GROC responses.

TABLE S2.2.1. Item analysis					
UWRI item	Baseline (mean ± SD)	Follow-up (mean ± SD)	Mean difference ± SD	Item-total correlation (*r*)	Item-rest correlation (*r*)
1: Injury impact on daily tasks	3.0 ± 0.9	3.3 ± 0.8	0.2 ± 1.1	0.67	0.63
2: Frustration due to injury	1.6 ± 1.1	2.7 ± 1.3	1.1 ± 1.4	0.75	0.71
3: Perceived recovery	1.7 ± 1.0	2.7 ± 1.0	1.0 ± 1.2	0.71	0.68
4: Pain during running	2.2 ± 1.0	3.0 ± 0.9	0.7 ± 1.1	0.73	0.70
5: Pain after running	2.4 ± 1.1	3.0 ± 0.9	0.5 ± 1.2	0.73	0.70
6: Weekly running volume	1.4 ± 1.1	2.5 ± 1.3	1.2 ± 1.4	0.82	0.79
7: Typical run distance	1.3 ± 1.1	2.5 ± 1.3	1.2 ± 1.4	0.81	0.78
8: Running pace	1.8 ± 1.2	2.8 ± 1.2	1.0 ± 1.4	0.76	0.72
9: Confidence to increase training	1.3 ± 1.1	2.3 ± 1.3	1.0 ± 1.6	0.75	0.71

TABLE S2.2.2. Factor loadings		
UWRI item	Load Management	Symptom Surveillance
1: Injury impact on daily tasks		0.71
2: Frustration due to injury	0.38	0.36
3: Perceived recovery		0.64
4: Pain during running	-0.19	0.97
5: Pain after running		0.87
6: Weekly running volume	0.97	
7: Typical run distance	1.08	-0.17
8: Running pace	0.58	0.18
9: Confidence to increase training	0.25	0.46
Eigenvalue	4.78	0.75
Proportion of Variance	0.34	0.30
Correlation	*r*=0.79	

Chapter 3. Mapping the University of Wisconsin Running Injury and Recovery Index to the VR-12 Using Data from a Prospective Cohort Study

Evan O. Nelson, David Vanness, Bryan Heiderscheit

Abstract

Background: Physical activity, including running, increases health-related quality of life (HRQoL) but the effect of running-related injury (RRI) is unknown. The purpose of this study was to explore the relationship between RRIs and HRQoL by developing a protocol to map the University of Wisconsin Running Injury and Recovery Index (UWRI) item responses to the Veterans RAND 12-item (VR-12) component scores.

Methods: Data were collected in a prospective cohort study from patients receiving physical therapy care for RRI. Spearman correlations evaluated the association between the UWRI, VR-12 and covariates. Multivariate ordinary least squares regression mapped the UWRI item responses, as ordered factors, to the VR-12 physical (PCS) and mental (MCS) component scores.

Results: 396 participants (62.1% female; mean(SD), age 34.5(12.4) years) provided UWRI and VR-12 data from 499 assessments (281 baseline; 218 follow-up). The mean(SD; range) UWRI score was 20.3(8.3; 0-36), PCS 52.1(46.6; 18.5-66.1), and MCS 51.7(9.2; 14.6-69.6). The UWRI score was associated with the PCS (ρ=0.34) and MCS (ρ=0.27). Performance of the selected model was poor with an adjusted R^2 of 0.36, 0.05 and mean absolute percentage error of 7.9%, 15.6% for the PCS and MCS models, respectively. The PCS and MCS models were correlated (r=0.34). Covariates were not associated with the VR-12 and were not significant model parameters.

Conclusion:

RRIs produce measureable changes in physical, and to a lesser degree mental, HRQoL. Mental HRQoL may decline with greater levels of psychological distress due to RRI. Future studies should use preference-based measures to assess runners' HRQoL, such as the VR-6D.

Keywords: running-related injuries, running, health-related quality of life, patient-reported outcome measure, University of Wisconsin Running Injury and Recovery Index, VR-12

Introduction

Health-related quality of life (HRQoL) is a multidimensional concept describing the interaction between an individual health state and quality of life. Athletes, and individuals participating in regular physical exercise have higher HRQoL than sedentary peers, and disparate HRQoL scores are associated with higher levels of sport or physical activity in student-athletes.[136-138] However, musculoskeletal injuries are known to reduce HRQoL and reductions may persist until well after an athlete returns to sport.[139 140] Lower quality of life is associated with worse functional outcomes, more depressive symptoms, and increased body mass index in athletes undergoing anterior cruciate ligament reconstruction.[141] Less significant injuries requiring nonsurgical treatment can cause significant reductions in young female athletes' physical and mental HRQoL that gradually improve during the 12-months following injury.[139] While athletic injuries can negatively affect HRQoL, sport-specific demands or distinctive athlete characteristics further influence HRQoL measurement. To accurately understand how HRQoL is affected by athletic injuries individual sports must be studied.

Running is a very popular form of physical activity with established benefit for multiple health indices.[142 143] Compared to non-runners, runners have greater life satisfaction, HRQoL, and reduced disability and all-cause mortality.[106 144] However, RRIs are very common.[8 11 14 17] RRIs include a heterogeneous collection of musculoskeletal injuries typically causing transient limitations, but injury is the primary reason many people permanently stop running.[11 12 14 31] Engagement with vigorous physical activity produces prompt HRQoL increases, but it is unknown if the reduction in physical activity caused by RRI is associated with changes in HRQoL.[145 146] Approximately 30% of RRIs require medical care, and this subgroup has more significant symptoms and greater participation limitations in running, work, and daily life.[11 14 49] Recent research shows the uncertainty regarding the health impact of RRIs. RRI incidence was found to be associated with mental, but not physical, HRQoL, and physical, but not mental, HRQoL measures were responsive to changes in function during recovery.[10](chapter 2) Understanding the relationship between RRIs and HRQoL is necessary to determine the medical justification for RRI care.

HRQoL is typically measured using generic patient-reported outcome measures (PROs) designed to span the spectrum of health. However, patients and clinicians prefer clinically-oriented PROs specifically measuring patients' primary problems. [147 148] The University of Wisconsin Running Injury and Recovery Index (UWRI) is a unidimensional, running-specific, evaluative PRO assessing perceived running ability following injury.[109] The UWRI is more responsive to RRI recovery than the Veterans Rand 12-item Health Survey (VR-12) HRQoL measure, but the construct-specific focus may disregard many HRQoL attributes.[109] In clinical settings, multiple barriers limit PRO collection (e.g., time, provider disincentives, patient burden, and technological limitations), and mapping is used to predict HRQoL values from other health measures.[147 149] Mapping requires a developmental dataset with both sets of measures, and regression techniques are used to map a clinically-oriented measure to a HRQoL measure. Premodelling recommendations include a detailed examination of the relationship between the two PRO measures.

The purpose of this study was to explore the relationship between RRIs and HRQoL. Specifically, we plan to develop a protocol to map UWRI item responses to the VR-12 physical and mental component scores.

Methods

Source of data

This is a secondary data analysis of a prospective, multicenter cohort study evaluating the psychometric properties of multiple patient-reported outcome measures for running-related injuries. Participants were recruited from specialized physical therapy running clinics. Inclusion criteria were age greater than or equal to 14 years, English literacy, and seeking care for at least 1 running-related injury. The treating physical therapist confirmed the presence of a RRI, which was defined as a musculoskeletal problem resulting from running that required the individual to prematurely stop running while training or during a competitive event or to miss a training session or competitive event.[73] Exclusion criteria were surgery at the site of injury in the preceding 12 months. The minimal risk institutional review board at the University of Wisconsin-Madison (submission ID #2016-1544) approved the study. Informed consent was obtained from all individual participants included in the study. Participants electronically completed study questionnaires at the onset of physical therapy care for a RRI (baseline) and 12 weeks later (follow-up) using Research Electronic Data Capture (REDCap™) hosted at the University of Wisconsin-Madison.[114] Baseline and follow-up data were pooled for mapping. Further description of the data collection methods and cohort is described elsewhere.(chapter 2)

Primary Measures for Analysis

University of Wisconsin Running Injury & Recovery Index (UWRI)

The UWRI is a valid, reliable, and responsive 9-item evaluative patient-reported outcome measure assessing running ability following running-related injury.(chapter 2) Respondents consider the preceding 7-days when responding to the individual items that are scored 0-4. The sum of item scores is the composite score. Higher scores indicating greater running function and the maximum score of 36 indicates a return to pre-injury running ability. In running-related injury populations, the UWRI demonstrates superior psychometric properties than other clinically-oriented, patient-reported outcome measures. The 95% minimal detectable change is 4.75 and 0.28 for individual and group assessments, respectively. The minimal clinically important difference is 8 points.

Veterans RAND 12-item Health Survey (VR-12)

The VR-12 is a 12-item generic health-related quality of life measure assessing 8 health domains: physical functioning, role limitations due to physical problems, bodily pain, general health, vitality, social functioning, and role limitations due to emotional problems. The VR-12 explains 90% of the variance of the Veterans RAND 36-item Health Survey (VR-36) from which it was derived. Physical (PCS) and mental (MCS) composite scores are scaled from 0-100 with higher scores indicating greater HRQoL. Extensive validity and reliability evaluations occurred ambulatory care settings and the VR-12 composite scores have a mean(SD) of 50(10) in the general US population.[133]

Statistical Analysis

Descriptive statistics were calculated for age, sex, race, height, weight, body mass index, injury location, injury duration, patient expectation for recovery time, and item and composite scores for the UWRI, PCS, and MCS. We then estimated the association between the UWRI and VR-12 by examining the pairwise comparisons of the UWRI, VR-12 and covariates at the item and composite score level using Spearman's rho. Associations were considered good $|r| > 0.75$, moderate $0.50 < |r| \leq 0.75$, fair $0.25 < |r| \leq 0.50$ and no association $|r| \leq 0.25$.[81]

Model Development

We estimated the relationship between PRO measures by constructing a multivariate, multiple ordinary least squares regression model. The PCS and MCS were included as multivariate dependent variables and all UWRI item responses were included in the initial model as ordered factor independent variables. For each item, the reference level was the response indicating the highest level of function. Statistical significance of the model parameters was determined using an analysis of variance (ANOVA). The modelling approach proceeded by sequentially adding sex, age, race, and injury duration as independent variables. Injury duration was an ordered factor variable. Models were compared using adjusted R^2, root mean square error (RMSE), mean absolute area (MAE), and mean percentage absolute area (MPSE) comparing observed and predicted scores. The effect size of significant models was determined using Cohen's f^2. The variance-covariance matrix produced by the multivariate model was used to determine the strength of association between the MCS and PCS models using the Pearson correlation coefficient. All analyses were conducted in R (R Project Core, Vienna, Austria), including the Metrics package.[150]

Results

396 participants enrolled in the study (246(62.1%) female; 145(37.6%) male; 5(1.3%) prefer not to answer). UWRI and VR-12 data were available from 281 baseline and 218 follow-up assessments. Censoring occurred for loss to follow-up or sustaining a new injury unrelated to running. The mean(SD) age was 34.5(12.4) years, BMI 23.3(3.9) kg/m^2, with 12.7(10.4) years of running experience. The knee 132(29.7%) was the most common site of injury (TABLE 3.1) followed by the calf 71(16%), hip 66(14.8%), ankle 61(13.7%), and foot 61(13.7%). Relative to pre-injury, self-reported weekly running volume reduced 47.0%, typical distance of a single run reduced 35.8%, and the median number of days per week was reduced by 1 day per week. The mean(SD; range) UWRI score was 16.8(6.7; 0-34) and 24.7(8.0; 0-36) at baseline and follow-up, respectively. The mean(SD; range) of observed VR-12 scores was 52.1(46.6; 18.5-66.1) and 51.7(9.2; 14.6-69.6) for the PCS and MCS, respectively.

TABLE 3.1. Sample characteristics*				
	Total			Total
Subjects, n(%)	396(100)		Injury location, n(%)	
Age (yrs)	34.5(12.4)		Low back or pelvis	24(5.4)
Status, n(%)			Hip	66(14.8)
Novice	9(2.3)		Thigh or hamstring	30(6.7)
Recreational	256(65.2)		Knee	132(29.7)
Competitive	130(33.3)		Calf or shin	71(16)
Professional	1(0.2)		Ankle	61(13.7)
Experience (yrs)	12.7(10.4)		Foot	61(13.7)
Ethnicity, n(%)			Injury duration, n(%)	
White, not Hispanic or Latino	340(85.9)		<1 month	28(7.1)
White, Hispanic or Latino	26(6.6)		1-2 months	66(16.7)
Other, not white	22(5.6)		3-6 months	135(34.1)
Not reported	6(1.5)		7-11 months	57(14.4)
Body mass index (kg/m^2)	23.3(3.9)		1-2 years	70(17.7)
Prior running injury, n(%)	214(54.0)		3-5 years	22(5.6)
Total diagnoses, n(%)	445(100)		6+ years	17(4.3)
*Values are mean(SD) unless otherwise indicated				

TABLE 3.2 presents Spearman correlation coefficients describing the association between the UWRI, selected covariates and the VR-12 component scores at baseline. The UWRI composite score demonstrated a fair association with the PCS (ρ=0.34) and MCS (ρ=0.27). Fair associations were also present between the PCS and UWRI item 1 (ρ=0.46), item 4 (ρ=0.28), item 5 (ρ=0.37), and item 6 (ρ=0.29). The remaining UWRI items were not associated with the PCS and no UWRI items demonstrated an association with the MCS. Demographic variables were not associated with the MCS or PCS. The association of the UWRI to VR-12 items (TABLE 3.3) were below the recommended threshold (ρ>0.80) to perform item linking.[151] UWRI item 1 evaluating the impact the RRI has on daily tasks, demonstrated a fair association with 7 VR-12 items (ρ=0.29-0.48). Fair associations were observed between the pain experienced in the 24

hours following a run (UWRI item 5) and 5 VR-12 items (ρ=0.28-0.37) while pain experienced while running (UWRI item 4) was associated with 3 VR-12 items (ρ=0.27-0.32). The runner's confidence to increase training (UWRI item 9) was also associated with 7 VR-12 items (ρ=0.29-0.48). The VR-12 vitality role physical domain was associated with 7 UWRI items as 6 items were associated with vitality role physical item 2 (ρ=0.26-0.41) and 5 items were associated with vitality role physical item 3 (ρ=0.34-0.41). The single item representing the VR-12 bodily pain domain was associated with 5 UWRI items (ρ=0.28-0.48).

TABLE 3.2. Spearman's Rho Correlation of Baseline Data (n=283)

Variables	PCS Correlation (ρ)	MCS Correlation (ρ)
UWRI Score	0.34	0.27
UWRI item 1	0.46	0.17
UWRI item 2	0.22	0.19
UWRI item 3	0.13	0.18
UWRI item 4	0.28	0.14
UWRI item 5	0.37	0.13
UWRI item 6	0.29	0.17
UWRI item 7	0.24	0.21
UWRI item 8	0.18	0.21
UWRI item 9	0.17	0.23
Sex	0.00	-0.01
Age	0.03	0.00
Injury duration	0.03	0.01
Race or ethnicity	0.06	-0.11

UWRI, University of Wisconsin Running Injury and Recovery Index; PCS, Veterans RAND 12-item health survey physical component score; MCS, Veterans RAND 12-item health survey mental component score

TABLE 3.3. Spearman's correlation coefficients of baseline UWRI items and VR-12 items (n=283)

		UWRI item								
		1	2	3	4	5	6	7	8	9
VR-12 item	GHx	0.11	0.05	0.05	0.09	0.08	0.04	0.03	0.02	0.11
	PF02x	0.32	0.19	0.14	0.16	0.32	0.21	0.22	0.22	0.32
	PF04x	0.41	0.18	0.16	0.31	0.28	0.22	0.22	0.21	0.41
	VRP2x	0.41	0.24	0.03	0.27	0.34	0.28	0.26	0.23	0.41
	VRP3x	0.41	0.31	0.01	0.18	0.34	0.35	0.31	0.34	0.41
	VRE2x	0.32	0.25	0.08	0.19	0.24	0.25	0.25	0.26	0.32
	VRE3x	0.29	0.19	0.12	0.18	0.20	0.17	0.18	0.19	0.29
	BP2x	0.48	0.20	0.20	0.32	0.37	0.28	0.24	0.16	0.48
	MH3x	0.13	0.16	0.16	0.17	0.18	0.18	0.21	0.17	0.13
	VT2x	0.23	0.17	0.20	0.19	0.18	0.21	0.23	0.17	0.23
	MH4x	0.22	0.17	0.20	0.24	0.20	0.20	0.24	0.29	0.22
	SF2x	0.19	0.24	0.20	0.13	0.17	0.16	0.18	0.16	0.19

Veterans RAND 12-item Health Survey (VR-12) scores are transformed to a 0-100 scale (denoted by the x). GH, general health; PF, physical functioning; VRP, vitality role physical; VRE, vitality role emotional; BP, bodily pain; VT, vitality; MH, mental health; SF, social functioning; UWRI, University of Wisconsin Running Injury and Recovery Index.

Based on the observed correlation strengths, model development proceeded by mapping the UWRI items as predictors of the VR-12 component scores. TABLE 3.4 presents the model coefficients, standard errors, and statistical significance. The PCS (P<0.001) and MCS (P=0.009) models were statistically significant and items 1-5 were significant parameters. The PCS developmental model had a large effect size (f^2= 0.55) while the MCS developmental model had a small effect size (f^2= 0.05).[152] All response levels of UWRI item 1 were significant predictors in the PCS model. "Unable to run" for UWRI item 8 was a significant predictor in the PCS model, but the same response was not significant for UWRI item 4-7. The response "Unable to run" was selected with variable frequency for item 4-8, 29(5.8%), 25(5.0%), 65(13.0%), 70(14.0%), 60(13.2%), respectively. The MCS model included 2 significant coefficients. A 6.9

point increase in MCS score was associated with the absence of perceived recovery (UWRI item 3) (*P*=0.01), and a lack of confidence to increase training without causing regression in injury (UWRI item 9) was associated with 4.8 point decrease in MCS score (*P*=0.04). The mean(SD; range) of predicted scores was 52.2(4.1; 36.4-57.9) and 51.7(3.2; 41.5-62.2) for the PCS and MCS, respectively. The adjusted R^2 estimates that this model explains 35.6% of the PCS variance and 4.8% of the MCS variance. The MAE for this model is 3.71 and 6.58 for the PCS and MCS, respectively. The model includes considerable prediction error as the MAPE was 7.9% and 15.6% for the PCS and MCS, respectively. The PCS and MCS models were correlated (*r*=0.34). Stepwise addition of demographic parameters did not improve model performance (TABLE 3.5).

TABLE 3.4. Developmental model

UWRI		PCS			MCS			Parameter significance
Item	Response	Coefficient	SE	*p*-value	Coefficient	SE	*p*-value	*p*-value
	Intercept	56.44	0.89	<0.001	53.61	1.53	<0.001	<0.001
1: Injury impact on daily tasks	No impact	Reference						<0.001
	Slight impact	**-2.37**	**0.89**	**<0.001**	-0.38	1.04	0.71	
	Moderate impact	**-5.47**	**0.60**	**<0.001**	-2.14	1.53	0.16	
	Significant impact	**-7.77**	**1.49**	**<0.001**	-2.41	2.58	0.35	
	Unable to perform	**-11.72**	**2.28**	**<0.001**	-6.22	3.94	0.12	
2: Frustration due to injury	Not frustrated	Reference						0.003
	Mildly frustrated	-1.10	0.96	0.26	0.81	1.66	0.63	
	Moderately frustrated	-0.72	1.07	0.50	-0.62	1.85	0.74	
	Significantly frustrated	0.86	1.17	0.46	-2.29	2.02	0.26	
	Extremely frustrated	0.08	1.26	0.95	-3.12	2.17	0.15	
3: Perceived recovery	Complete recovery	Reference						0.015
	Significant recovery	0.11	1.15	0.92	2.27	1.99	0.25	
	Moderate recovery	-0.44	1.31	0.74	4.30	2.27	0.06	
	Minimal recovery	1.73	1.40	0.22	1.03	2.42	0.67	
	No recovery	-1.58	1.53	0.30	**6.91**	**2.63**	**0.01**	

4: Pain during running	No pain	Reference						<0.001
	Minimal pain	0.003	0.87	0.99	-2.01	1.50	0.18	
	Moderate pain	-0.41	1.06	0.70	-3.59	1.83	0.05	
	Significant pain	-3.37	1.42	0.02	-1.67	2.45	0.50	
	Unable to run	-5.1	2.55	0.05	-1.58	4.41	0.72	
5: Pain after running	No pain	Reference						0.03
	Minimal pain	-0.85	0.75	0.26	0.14	1.30	0.92	
	Moderate pain	-1.61	0.95	0.09	-0.72	1.64	0.66	
	Significant pain	-5.22	1.40	<0.001	1.38	2.42	0.57	
	Unable to run	-0.40	2.58	0.88	-2.66	4.46	0.55	
6: Weekly running volume	Same as before injury	Reference						0.48
	Minimally reduced	0.28	1.20	0.81	-1.41	2.08	0.50	
	Moderately reduced	-0.40	1.45	0.78	0.50	2.50	0.84	
	Significantly reduced	-1.57	1.63	0.34	4.02	2.82	0.15	
	Unable to run	-6.25	3.37	0.06	2.84	5.81	0.63	
7: Typical run distance	Same as before injury	Reference						0.52
	Minimally reduced	-0.61	1.22	0.62	2.04	2.11	0.33	
	Moderately reduced	-0.16	1.50	0.92	1.56	2.58	0.55	
	Significantly reduced	0.33	1.58	0.83	-1.76	2.73	0.52	
	Unable to run	-2.06	2.59	0.43	-3.35	4.47	0.45	
8: Running pace	Same as before injury	Reference						0.18
	Minimally reduced	-0.05	0.80	0.95	0.60	1.39	0.67	
	Moderately reduced	0.20	0.91	0.83	-1.87	1.58	0.24	
	Significantly reduced	-0.05	1.01	0.96	-0.83	1.74	0.64	
	Unable to run	**5.98**	**2.75**	**0.03**	2.04	4.75	0.67	
9: Confidence to increase training	Confident to increase	Reference						0.43
	If I increase I might be fine	1.44	0.95	0.13	-2.39	1.64	0.15	
	Neutral	0.47	1.11	0.67	-2.71	1.92	0.16	
	If I increase I might get worse	0.64	0.99	0.52	-1.82	1.72	0.29	
	I cannot increase	1.74	1.32	0.19	**-4.80**	**2.29**	**0.04**	

UWRI, University of Wisconsin Running Injury and Recovery Index; PCS, Veterans Rand 12-item physical component score; MCS, Veterans Rand 12-item mental component score; SE, standard error.

TABLE 3.5. Model performance										
	Model 1: UWRI items		Model 2: Model 1 + sex		Model 3: Model 2 + age		Model 4: Model 3 + race		Model 5: Model 4 + injury duration	
	PCS	MCS	PCS	MCS	PCS	MCS	PCS	MCS	PCS	MCS
Adjusted R^2	0.36	0.05	0.35	0.05	0.36	0.05	0.35	0.04	0.35	0.03
MAE	3.71	6.58	3.71	6.57	3.71	6.58	3.71	6.56	3.71	6.55
RMSE	5.02	8.67	5.02	8.67	5.01	8.66	5.01	8.65	5.00	8.64
MAPE (%)	7.92	15.56	7.92	15.55	7.93	15.54	7.93	15.53	7.93	15.49
p-value[a]	-		0.86		0.28		0.70		0.98	

MAE, mean absolute error; RMSE, root mean square error; MAPE, mean absolute percentage error; PCS, Veterans Rand 12-item physical component score; MCS, Veterans Rand 12-item mental component score.
[a]ANOVA comparing model 1 to the subsequent model

Discussion

This is the first study to investigate the relationship between RRIs and HRQoL by mapping the UWRI to the VR-12. Direct, prospective collection of preference-based measures is preferred, but this type of data is not always available.[149] Mapping is commonly employed to overcome this limitation, and mapping algorithms have been successfully developed to estimate preference-based measures from low back pain condition-specific measures.[56] Developmental models were computed using accepted techniques, but model performance was not consistent with accepted mapping studies.[153] A broad sample is recommended in mapping studies and running inherently requires a high level of physical function, shifting the sample in this population to represent a narrow section on the HRQoL continuum.[149] Even though the mapping algorithms demonstrated unsatisfactory performance, these results provide unique information about the impact RRIs have on HRQoL.

Prior studies found that runners have better HRQoL compared to peers, and participation in vigorous forms of physical activity, such as running, improves HRQoL and may prevent age-related declines.[145 154 155] The current study provides evidence that RRIs produce measureable reductions in physical, and to a lesser degree mental, HRQoL. A meta-analysis conducted by Houston et al found injuries reduced HRQoL in athletes with a strong effect (Hedge's g=0.85) for the PCS and a weak effect (Hedge's g=0.15) for the MCS, similar to the effect size of the developed models.[137] In the current study, the majority of participants were running in a reduced capacity, and the observed HRQoL reductions are consistent with other studies showing reductions persist, even after the return to sport.[139 156] The focus of this study provides benefit to sports medicine because RRIs include a heterogeneous collection of primarily overuse, lower extremity musculoskeletal injuries that may not have the same HRQoL impact as the traumatic injuries occurring in other sports. This study included longitudinal assessments, but the effect of RRIs on HRQoL over the course of recovery remains unknown.

Very few RRI studies include HRQoL measures, but Messier et al assessed HRQoL using the 12-item short form health survey (SF-12), a comparable measure to the VR-12, in a 2-year, prospective study evaluating the etiological factors associated with overuse RRIs. SF-12 MCS was predictive of injury with a mean(SD) of 47.8(5.9) and 49.5(3.2) for injured and uninjured runners, respectively (*P*=0.01), whereas the mean(SD) PCS was 56.0(2.4) and 55.8(1.7) for injured and uninjured runners, respectively (*P*=0.37).[10] Differences in demographic characteristics are present with the current study recruiting a younger sample running fewer miles at a faster pace with a lower average PCS and higher average MCS than the study by

Messier and colleagues. Given the current results, the utility of the MCS for determining injury risk or prognosis remains uncertain because of the substantial error in the MCS model and potential mediating effect of sociodemographic factors, fitness, or self-efficacy.[157 158]

This study offers granular detail regarding the association between HRQoL and running ability following RRI. Function, representing 92% of SF-12 items, is a predominant construct in many HRQoL measures overlapping the primary limitation imposed by RRIs.[159] In the developmental model, multiple UWRI items demonstrated dose-response relationships with VR-12 component scores. Physical HRQoL declines as the RRI imposes progressively greater limitations in daily tasks. Model coefficients indicate HRQoL may be affected by incremental changes in other elements contributing to perceived running ability. In general, mental HRQoL is a unique construct inconsistently affected by RRIs. However, the psychological response to RRI can impose a detrimental toll, and mental HRQoL may decline with progressively greater levels of psychological distress. The VR-12 component scores contrast physical and mental contributions to overall health, and the direct association between HRQoL components may explain the observed model correlation.[115 160]

A key strength of this study is the use of multivariate multiple regression modelling enabling comparison of the two models. The results of this study represent clinical care very well as all data were collected from patients receiving physical therapy care. Multiple limitations exist that influence the generalizability of these results. Participants represent a fairly narrow portion of the HRQoL spectrum and factors influencing HRQoL such as medical co-morbidities, education level, socioeconomic status or health behavior in addition to running were unavailable and could not be included in the model.[161] Conducting a regression with ordered factors distributed the study's relatively small sample size. This may contribute to coefficient variability within the model, including statistical significance in the absence of face validity. This occurred in the PCS model; the response "Unable to run" demonstrated statistical significance for the assessment of running pace (UWRI item 8), but the same response was not significant in items 5-7.

This study used a multisite data set to empirically estimate the relationship between the UWRI and VR-12. Running ability following RRI is associated with HRQoL and further work is necessary to determine the toll injuries impose. Future studies should use direct assessment with preference-based measures to investigate the relationship between RRIs and HRQoL and evaluate how other measures of health status affect the observed relationship.

Conclusion

Summary of findings

The objective of this thesis was to develop a running-specific evaluative PRO to measure longitudinal change in running ability for use in clinical practice and research. The UWRI demonstrates acceptable content validity, reliability, construct validity, and responsiveness when measuring longitudinal change in running ability following injury. In a sample of runners with RRIs, the UWRI demonstrates superior psychometric integrity compared to the other PROs included in this thesis. Research describing RRI recovery is scarce, and this thesis makes a number of other contributions. This is the first study to conduct a psychometric assessment of the included PROs in a population of runners. Consensus statements promote recurrent assessment of the psychological response to athletic injury.[34] Runners involved in the development process illuminated this important concept well before it was established in published literature. Items which evaluate the most common psychological responses to injuries enable efficient, comprehensive assessment of the runner's recovery. Finally, the most detailed investigation to date of the relationship between RRIs and HRQoL was performed.

Assessment of impact

Remarkably, the UWRI translated into routine practice even before its dissemination in refereed journals. The UWRI is a standard measure in multiple organizations throughout the United States and is used internationally as well. At a national physical therapy conference, a peer-reviewed case presentation describing a running-specific intervention included the UWRI as an assessment of treatment outcome. Early adopters anecdotally report that the ease of administration, application to the construct, and heightened responsiveness promote clinical integration. A study, conducted by a different research group, is currently underway to validate the Spanish translation of the UWRI. These examples indicate the UWRI is an asset to patient-oriented clinical outcomes measurement with promising potential for widespread clinical adoption. The use of the UWRI for clinical research is growing as well. TRIA Orthopaedic Center is using the UWRI in ongoing research to investigate the relationship between RRI and psychological distress. Early results from this work were presented as a poster at a national physical therapy conference. The UWRI is adopted by external researchers conducting doctoral thesis projects investigating the management and treatment of running injuries in the postpartum population at the University of Newcastle and running recovery following anterior cruciate ligament reconstruction at The United States Military Academy at West Point.

Future directions

The conclusion of this thesis generates a range of research questions. Additional psychometric evaluation of the UWRI in specific subgroups is warranted. A single professional runner was represented in the current study, and novice runners were underrepresented. Novice runners represent a unique subgroup with a higher RRI incidence and greater likelihood to discontinue running due to injury.[8 9] Novice runners' limited running experience may influence perceived recovery. Elite competitive and professional runners have unique training demands, requiring significant commitment to the sport. Running may partially define these athletes' self-identity, and injuries limiting running ability may have unique psychological consequences.[162] Financial motivators associated with professional running could affect recovery. Continued psychometric evaluation of the UWRI in elite competitive and professional runners is justified.

Many people seek care when injuries unrelated to running or a surgery consequently limit running ability. To enhance generalizability, a psychometric assessment of the UWRI should be performed with patients whose running ability is limited by alternative causes. UWRI items evaluate limitations due to “injury”, without specifically identifying RRIs. This allows limitations in running ability from other causes to be adequately assessed. This study was challenged by missing data and loss to follow-up, reducing the net sample size. A larger study which includes a robust set of covariates would provide a more accurate assessment of item performance and UWRI psychometric performance in runner or injury subgroups.

Comparing new patient-reported outcomes to global assessments of perceived change is preferred when a gold standard does not exist.[76 112] However, global measures requiring patients to accurately compare their current status to a historical state are temporally unstable.[94 126] Runners and running injury researchers are rapidly adopting unobtrusive, wearable technology which provides granular assessments of training load and behavior.[163] Wearable technology provides accurate, detailed assessments of running in the normal training environment, but this technology cannot include the runner’s perception. Including the UWRI and wearable technology in a prospective study would assess the validity of the UWRI against a comparator representing the primary construct of interest, running ability.

The advanced techniques necessary to collect, analyze, and interpret wearable technology data are currently used in RRI research. In a worldwide prospective study, the "Garmin-RUNSAFE Running Health Study" is gathering large-scale data from wearable devices to explore how changes in running affect various health-related factors and lead to RRI.[164] In the DANO-RUN study of RRI risk, the same research group used GPS watches to measure running exposure, but the time to recovery analysis arbitrarily defined recovery as "no pain in the affected anatomical location following two consecutive running sessions of at least 500m."[14] These recent examples demonstrate research partnerships to further investigate the UWRI psychometric integrity are viable, but also show the UWRI is a valuable asset to incorporate the patient-perspective in RRI research. Defining recovery using arbitrary distance criteria, or even complicated data from wearable technology, fails to incorporate the multiple internal and external factors influencing recovery. Failing to accurately assess the multifactorial process risks inaccurately estimating RRI consequences or incorrectly assessing treatment benefits.

Health-related quality of life

The resources required to care for RRI motivates the need to understand the impact RRIs have on quality of life. This thesis demonstrated meaningful changes in the VR-12 PCS and VR-6D occur during RRI recovery. Logically, the predominantly physical nature of RRIs was associated with physical HRQoL, but not mental HRQoL. The similarity in VR-12 PCS and VR-6D responsiveness was unexpected because the VR-6D combines elements of the PCS and MCS. The association between the UWRI and each HRQoL measure was stable across the initial, final, and change assessments in this longitudinal project. The mapping analysis provided additional insight into the association between different levels of patient-perceived impairment and HRQoL. Considering the known presence of direct associations between HRQoL domains, it appears runners experiencing more significant RRIs have distressing symptoms that may affect multiple domains of health. This thesis demonstrated HRQoL improves in runners receiving care for a RRI. Future studies of HRQoL involving runners should include a preference-based utility measure.

Previous studies demonstrated athletes have better HRQoL; but injuries reduce HRQoL for an extended period following injury.[136 137 165] However, these studies did not have the heterogeneity of age, injury, physical performance, training behavior, or overall health present in the running population. This thesis was unable to determine if injuries reduce HRQoL and it is unclear how the current sample compares with the general population. A large longitudinal study that performs a baseline assessment prior to injury, has a follow-up period of at least 12-months, and robust set of covariates would provide better insight into changes RRIs impose on HRQoL. Alternatively, a cross-sectional study could test for HRQoL differences in runners with and without injuries, and compare runners' HRQoL to matched controls in the general population. It is recommended this type of cross-sectional study collect sufficient data to represent the variability observed in the patient population and within different health states.[53] Either design would build upon the current result, showing the impact RRIs have on health and providing appropriate context for HRQoL in a unique population that regularly participates in a positive health activity.

Running is an extremely accessible primary or secondary prevention strategy for some of the leading causes of death. High RRI incidence, the functional consequences of injury, and costs associated with RRI care justify a deeper understanding of RRI recovery. Ultimately, the proposed research progressions provide foundational knowledge supporting the critical inquiries related to resource allocation. The UWRI provides an essential tool for assessing the effectiveness of accepted running-specific interventions, such as gait retraining, but it is not an appropriate surrogate measure of HRQoL.[2] The VR-6D or VR-12 PCS can be used to develop a deeper understanding of the relationship between RRIs and overall health.

References

1. Barton CJ. Managing RISK when treating the injured runner with running retraining, load management and exercise therapy. *Physical therapy in sport : official journal of the Association of Chartered Physiotherapists in Sports Medicine* 2018;29:79-83. doi: 10.1016/j.ptsp.2017.10.002 [published Online First: 2017/10/17]
2. Barton CJ, Bonanno DR, Carr J, et al. Running retraining to treat lower limb injuries: a mixed-methods study of current evidence synthesised with expert opinion. *British journal of sports medicine* 2016;50(9):513-26. doi: 10.1136/bjsports-2015-095278 [published Online First: 2016/02/18]
3. Andersen JJ. The State of Running 2019: RunRepeat.com; 2019 [updated 7/10/2019. accessed 7/12/2019 2019.
4. The National Association of State High School Athletic Associations. 2017-18 High School Athletics Participation Survey. Online, 2018.
5. Rauh MJ, Koepsell TD, Rivara FP, et al. Epidemiology of musculoskeletal injuries among high school cross-country runners. *American journal of epidemiology* 2006;163(2):151-9. doi: 10.1093/aje/kwj022 [published Online First: 2005/11/25]
6. Taunton JE, Ryan MB, Clement DB, et al. A retrospective case-control analysis of 2002 running injuries. *British journal of sports medicine* 2002;36(2):95-101. [published Online First: 2002/03/28]
7. Van Middelkoop M, Kolkman J, Van Ochten J, et al. Risk factors for lower extremity injuries among male marathon runners. *Scandinavian journal of medicine & science in sports* 2008;18(6):691-7. doi: 10.1111/j.1600-0838.2007.00768.x [published Online First: 2008/02/13]
8. Videbaek S, Bueno AM, Nielsen RO, et al. Incidence of Running-Related Injuries Per 1000 h of running in Different Types of Runners: A Systematic Review and Meta-Analysis. *Sports medicine (Auckland, NZ)* 2015;45(7):1017-26. doi: 10.1007/s40279-015-0333-8 [published Online First: 2015/05/09]
9. Buist I, Bredeweg SW, Bessem B, et al. Incidence and risk factors of running-related injuries during preparation for a 4-mile recreational running event. *British journal of sports medicine* 2010;44(8):598-604. doi: 10.1136/bjsm.2007.044677 [published Online First: 2008/05/20]
10. Messier SP, Martin DF, Mihalko SL, et al. A 2-Year Prospective Cohort Study of Overuse Running Injuries: The Runners and Injury Longitudinal Study (TRAILS). *Am J Sports Med* 2018;46(9):2211-21. doi: 10.1177/0363546518773755 [published Online First: 2018/05/24]
11. van Poppel D, Scholten-Peeters GG, van Middelkoop M, et al. Prevalence, incidence and course of lower extremity injuries in runners during a 12-month follow-up period. *Scandinavian journal of medicine & science in sports* 2014;24(6):943-9. doi: 10.1111/sms.12110 [published Online First: 2013/08/21]
12. Linton L, Valentin S. Running with injury: A study of UK novice and recreational runners and factors associated with running related injury. *Journal of science and medicine in sport* 2018;21(12):1221-25. doi: 10.1016/j.jsams.2018.05.021 [published Online First: 2018/06/02]
13. Rauh MJ, Margherita AJ, Rice SG, et al. High school cross country running injuries: a longitudinal study. *Clin J Sport Med* 2000;10(2):110-6. [published Online First: 2000/05/08]
14. Nielsen RO, Ronnow L, Rasmussen S, et al. A prospective study on time to recovery in 254 injured novice runners. *PloS one* 2014;9(6):e99877. doi: 10.1371/journal.pone.0099877 [published Online First: 2014/06/14]
15. Ramskov D, Rasmussen S, Sorensen H, et al. Progression in Running Intensity or Running Volume and the Development of Specific Injuries in Recreational Runners: Run Clever, a Randomized Trial

Using Competing Risks. *The Journal of orthopaedic and sports physical therapy* 2018;48(10):740-48. doi: 10.2519/jospt.2018.8062 [published Online First: 2018/06/14]
16. Saragiotto BT, Yamato TP, Hespanhol Junior LC, et al. What are the main risk factors for running-related injuries? *Sports Med* 2014;44(8):1153-63. doi: 10.1007/s40279-014-0194-6 [published Online First: 2014/05/09]
17. Nielsen RO, Buist I, Sorensen H, et al. Training errors and running related injuries: a systematic review. *International journal of sports physical therapy* 2012;7(1):58-75. [published Online First: 2012/03/06]
18. Johansen KK, Hulme A, Damsted C, et al. Running Injury Development: The Attitudes of Middle- and Long-Distance Runners and Their Coaches. *International journal of sports physical therapy* 2017;12(4):634-41. [published Online First: 2017/09/14]
19. Buist I, Bredeweg SW, van Mechelen W, et al. No effect of a graded training program on the number of running-related injuries in novice runners: a randomized controlled trial. *Am J Sports Med* 2008;36(1):33-9. doi: 10.1177/0363546507307505 [published Online First: 2007/10/18]
20. Bertelsen ML, Hulme A, Petersen J, et al. A framework for the etiology of running-related injuries. *Scandinavian journal of medicine & science in sports* 2017;27(11):1170-80. doi: 10.1111/sms.12883 [published Online First: 2017/03/23]
21. Heiderscheit BC, Chumanov ES, Michalski MP, et al. Effects of step rate manipulation on joint mechanics during running. *Med Sci Sports Exerc* 2011;43(2):296-302. doi: 10.1249/MSS.0b013e3181ebedf4 [published Online First: 2010/06/29]
22. Wille CM, Lenhart RL, Wang S, et al. Ability of sagittal kinematic variables to estimate ground reaction forces and joint kinetics in running. *The Journal of orthopaedic and sports physical therapy* 2014;44(10):825-30. doi: 10.2519/jospt.2014.5367 [published Online First: 2014/08/27]
23. Finch CF, Cook J. Categorising sports injuries in epidemiological studies: the subsequent injury categorisation (SIC) model to address multiple, recurrent and exacerbation of injuries. *British journal of sports medicine* 2014;48(17):1276-80. doi: 10.1136/bjsports-2012-091729 [published Online First: 2013/03/19]
24. Timpka T, Alonso JM, Jacobsson J, et al. Injury and illness definitions and data collection procedures for use in epidemiological studies in Athletics (track and field): consensus statement. *British journal of sports medicine* 2014;48(7):483-90. doi: 10.1136/bjsports-2013-093241 [published Online First: 2014/03/13]
25. Bahr R. No injuries, but plenty of pain? On the methodology for recording overuse symptoms in sports. *British journal of sports medicine* 2009;43(13):966-72. doi: 10.1136/bjsm.2009.066936 [published Online First: 2009/12/01]
26. Clarsen B, Myklebust G, Bahr R. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) overuse injury questionnaire. *British journal of sports medicine* 2013;47(8):495-502. doi: 10.1136/bjsports-2012-091524 [published Online First: 2012/10/06]
27. Kluitenberg B, van Middelkoop M, Smits DW, et al. The NLstart2run study: Incidence and risk factors of running-related injuries in novice runners. *Scandinavian journal of medicine & science in sports* 2015;25(5):e515-23. doi: 10.1111/sms.12346 [published Online First: 2014/12/03]
28. Ryan MB, Valiant GA, McDonald K, et al. The effect of three different levels of footwear stability on pain outcomes in women runners: a randomised control trial. *British journal of sports medicine* 2011;45(9):715-21. doi: 10.1136/bjsm.2009.069849 [published Online First: 2010/06/30]
29. Hespanhol Junior LC, Huisstede BM, Smits DW, et al. The NLstart2run study: Economic burden of running-related injuries in novice runners participating in a novice running program. *Journal of*

science and medicine in sport 2016;19(10):800-4. doi: 10.1016/j.jsams.2015.12.004 [published Online First: 2016/01/05]

30. Vadeboncoeur TF, Silvers SM, Taylor WC, et al. Impact of a high body mass index on lower extremity injury in marathon/half-marathon participants. *J Phys Act Health* 2012;9(1):96-103. [published Online First: 2012/01/11]
31. Koplan JP, Rothenberg RB, Jones EL. The natural history of exercise: a 10-yr follow-up of a cohort of runners. *Med Sci Sports Exerc* 1995;27(8):1180-4. [published Online First: 1995/08/01]
32. Marti B, Vader JP, Minder CE, et al. On the epidemiology of running injuries. The 1984 Bern Grand-Prix study. *Am J Sports Med* 1988;16(3):285-94. doi: 10.1177/036354658801600316 [published Online First: 1988/05/01]
33. Ekegren CL, Gabbe BJ, Finch CF. Sports Injury Surveillance Systems: A Review of Methods and Data Quality. *Sports medicine (Auckland, NZ)* 2016;46(1):49-65. doi: 10.1007/s40279-015-0410-z [published Online First: 2015/10/02]
34. Psychological Issues Related to Illness and Injury in Athletes and the Team Physician: a Consensus Statement-2016 Update. *Curr Sports Med Rep* 2017;16(3):189-201. doi: 10.1249/JSR.0000000000000359 [published Online First: 2017/05/13]
35. Ardern CL, Taylor NF, Feller JA, et al. A systematic review of the psychological factors associated with returning to sport following injury. *British journal of sports medicine* 2013;47(17):1120-6. doi: 10.1136/bjsports-2012-091203 [published Online First: 2012/10/16]
36. Forsdyke D, Smith A, Jones M, et al. Psychosocial factors associated with outcomes of sports injury rehabilitation in competitive athletes: a mixed studies systematic review. *British journal of sports medicine* 2016;50(9):537-44. doi: 10.1136/bjsports-2015-094850 [published Online First: 2016/02/19]
37. Morrey MA, Stuart MJ, Smith AM, et al. A longitudinal examination of athletes' emotional and cognitive responses to anterior cruciate ligament injury. *Clinical journal of sport medicine : official journal of the Canadian Academy of Sport Medicine* 1999;9(2):63-9. [published Online First: 1999/08/12]
38. Podlog L, Heil J, Schulte S. Psychosocial factors in sports injury rehabilitation and return to play. *Phys Med Rehabil Clin N Am* 2014;25(4):915-30. doi: 10.1016/j.pmr.2014.06.011 [published Online First: 2014/12/03]
39. Kellmann M, Bertollo M, Bosquet L, et al. Recovery and Performance in Sport: Consensus Statement. *Int J Sports Physiol Perform* 2018;13(2):240-45. doi: 10.1123/ijspp.2017-0759 [published Online First: 2018/01/19]
40. Almeida MO, Davis IS, Lopes AD. Biomechanical Differences of Foot-Strike Patterns During Running: A Systematic Review With Meta-analysis. *The Journal of orthopaedic and sports physical therapy* 2015;45(10):738-55. doi: 10.2519/jospt.2015.6019 [published Online First: 2015/08/26]
41. Brindle RA, Milner CE, Zhang S, et al. Changing step width alters lower extremity biomechanics during running. *Gait Posture* 2014;39(1):124-8. doi: 10.1016/j.gaitpost.2013.06.010 [published Online First: 2013/07/09]
42. Tate JJ, Milner CE. Sound-Intensity Feedback During Running Reduces Loading Rates and Impact Peak. *The Journal of orthopaedic and sports physical therapy* 2017;47(8):565-69. doi: 10.2519/jospt.2017.7275 [published Online First: 2017/07/07]
43. Willy RW, Scholz JP, Davis IS. Mirror gait retraining for the treatment of patellofemoral pain in female runners. *Clin Biomech (Bristol, Avon)* 2012;27(10):1045-51. doi: 10.1016/j.clinbiomech.2012.07.011 [published Online First: 2012/08/25]

44. Roper JL, Harding EM, Doerfler D, et al. The effects of gait retraining in runners with patellofemoral pain: A randomized trial. *Clin Biomech (Bristol, Avon)* 2016;35:14-22. doi: 10.1016/j.clinbiomech.2016.03.010 [published Online First: 2016/04/26]
45. Esculier JF, Bouyer LJ, Dubois B, et al. Is combining gait retraining or an exercise programme with education better than education alone in treating runners with patellofemoral pain?A randomised clinical trial. *British journal of sports medicine* 2018;52(10):659-66. doi: 10.1136/bjsports-2016-096988 [published Online First: 2017/05/10]
46. Davis I. Optimising the efficacy of gait retraining. *British journal of sports medicine* 2018;52(10):624-25. doi: 10.1136/bjsports-2017-098297 [published Online First: 2017/09/21]
47. Esculier JF, Dubois B, Roy JS. Response to: 'Optimising the efficacy of gait retraining'. *British journal of sports medicine* 2019;53(9):530-31. doi: 10.1136/bjsports-2017-098609 [published Online First: 2018/03/09]
48. Walker N, Thatcher J, Lavallee D. Psychological responses to injury in competitive sport: a critical review. *The journal of the Royal Society for the Promotion of Health* 2007;127(4):174-80. [published Online First: 2007/08/23]
49. Hespanhol Junior LC, van Mechelen W, Postuma E, et al. Health and economic burden of running-related injuries in runners training for an event: A prospective cohort study. *Scandinavian journal of medicine & science in sports* 2016;26(9):1091-9. doi: 10.1111/sms.12541 [published Online First: 2015/08/19]
50. World Health Organization. Constitution of the World Health Organization, 2005.
51. Karimi M, Brazier J. Health, Health-Related Quality of Life, and Quality of Life: What is the Difference? *Pharmacoeconomics* 2016;34(7):645-9. doi: 10.1007/s40273-016-0389-9 [published Online First: 2016/02/20]
52. Gold MR. Cost-effectiveness in health and medicine. New York: Oxford University Press 1996.
53. Wolowacz SE, Briggs A, Belozeroff V, et al. Estimating Health-State Utility for Economic Models in Clinical Studies: An ISPOR Good Research Practices Task Force Report. *Value Health* 2016;19(6):704-19. doi: 10.1016/j.jval.2016.06.001 [published Online First: 2016/10/08]
54. Selim AJ, Rogers W, Qian SX, et al. A preference-based measure of health: the VR-6D derived from the veterans RAND 12-Item Health Survey. *Qual Life Res* 2011;20(8):1337-47. doi: 10.1007/s11136-011-9866-y [published Online First: 2011/02/22]
55. Iqbal SU RW, Selim A, Qian S, Lee A, Ren XS, Rothendler J, Miller D and Kazis LE. The Veterans RAND 12 Item Health Survey (VR-12): What It Is and How It Is Used. Bedford, MA: VA Medical Center, 2007.
56. Rundell SD, Bresnahan BW, Heagerty PJ, et al. Mapping a patient-reported functional outcome measure to a utility measure for comparative effectiveness and economic evaluations in older adults with low back pain. *Med Decis Making* 2014;34(7):873-83. doi: 10.1177/0272989X14533995 [published Online First: 2014/05/16]
57. Griggs CL, Schneider JC, Kazis LE, et al. Patient-reported Outcome Measures: A Stethoscope for the Patient History. *Annals of surgery* 2017;265(6):1066-67. doi: 10.1097/sla.0000000000002165 [published Online First: 2017/01/31]
58. Jette DU, Halbert J, Iverson C, et al. Use of standardized outcome measures in physical therapist practice: perceptions and applications. *Physical therapy* 2009;89(2):125-35. doi: 10.2522/ptj.20080234 [published Online First: 2008/12/17]
59. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an

international Delphi study. *Qual Life Res* 2010;19(4):539-49. doi: 10.1007/s11136-010-9606-8 [published Online First: 2010/02/20]

60. Binkley JM, Stratford PW, Lott SA, et al. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Physical therapy* 1999;79(4):371-83. [published Online First: 1999/04/14]
61. Watson CJ, Propps M, Ratner J, et al. Reliability and responsiveness of the lower extremity functional scale and the anterior knee pain scale in patients with anterior knee pain. *The Journal of orthopaedic and sports physical therapy* 2005;35(3):136-46. doi: 10.2519/jospt.2005.35.3.136 [published Online First: 2005/04/21]
62. Martin RL, Irrgang JJ, Burdett RG, et al. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int* 2005;26(11):968-83. doi: 10.1177/107110070502601113 [published Online First: 2005/11/29]
63. Mohtadi NG, Griffin DR, Pedersen ME, et al. The Development and validation of a self-administered quality-of-life outcome measure for young, active patients with symptomatic hip disease: the International Hip Outcome Tool (iHOT-33). *Arthroscopy* 2012;28(5):595-605; quiz 06-10 e1. doi: 10.1016/j.arthro.2012.03.013 [published Online First: 2012/05/01]
64. 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]
65. Hernandez-Sanchez S, Hidalgo MD, Gomez A. Responsiveness of the VISA-P scale for patellar tendinopathy in athletes. *British journal of sports medicine* 2014;48(6):453-7. doi: 10.1136/bjsports-2012-091163 [published Online First: 2012/09/27]
66. Macdermid JC, Silbernagel KG. Outcome Evaluation in Tendinopathy: Foundations of Assessment and a Summary of Selected Measures. *The Journal of orthopaedic and sports physical therapy* 2015;45(11):950-64. doi: 10.2519/jospt.2015.6054 [published Online First: 2015/10/17]
67. Gallagher J, Needleman I, Ashley P, et al. Self-Reported Outcome Measures of the Impact of Injury and Illness on Athlete Performance: A Systematic Review. *Sports medicine (Auckland, NZ)* 2017;47(7):1335-48. doi: 10.1007/s40279-016-0651-5 [published Online First: 2016/12/21]
68. Mallows A, Littlewood C, Malliaras P. Measuring patient-reported outcomes (PROs/PROMs) in people with Achilles tendinopathy: how useful is the VISA-A? *British journal of sports medicine* 2018;52(19):1221. doi: 10.1136/bjsports-2017-097531 [published Online First: 2017/06/24]
69. Damsted C, Parner ET, Sorensen H, et al. Design of ProjectRun21: a 14-week prospective cohort study of the influence of running experience and running pace on running-related injury in half-marathoners. *Inj Epidemiol* 2017;4(1):30. doi: 10.1186/s40621-017-0124-9 [published Online First: 2017/11/07]
70. Kluitenberg B, van Middelkoop M, Diercks RL, et al. The NLstart2run study: health effects of a running promotion program in novice runners, design of a prospective cohort study. *BMC Public Health* 2013;13:685. doi: 10.1186/1471-2458-13-685 [published Online First: 2013/07/31]
71. Ramskov D, Nielsen RO, Sorensen H, et al. The design of the run Clever randomized trial: running volume, -intensity and running-related injuries. *BMC Musculoskelet Disord* 2016;17:177. doi: 10.1186/s12891-016-1020-0 [published Online First: 2016/04/25]
72. van Gent RN, Siem D, van Middelkoop M, et al. Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review. *British journal of sports medicine* 2007;41(8):469-80; discussion 80. doi: 10.1136/bjsm.2006.033548 [published Online First: 2007/05/03]

73. Kluitenberg B, van Middelkoop M, Verhagen E, et al. The impact of injury definition on injury surveillance in novice runners. *Journal of science and medicine in sport* 2016;19(6):470-5. doi: 10.1016/j.jsams.2015.07.003 [published Online First: 2015/07/25]
74. Pipkin A, Kotecki K, Hetzel S, et al. Reliability of a Qualitative Video Analysis for Running. *The Journal of orthopaedic and sports physical therapy* 2016;46(7):556-61. doi: 10.2519/jospt.2016.6280 [published Online First: 2016/06/09]
75. Juniper EF, Guyatt GH, Streiner DL, et al. Clinical impact versus factor analysis for quality of life questionnaire construction. *Journal of clinical epidemiology* 1997;50(3):233-8. [published Online First: 1997/03/01]
76. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of clinical epidemiology* 2007;60(1):34-42. doi: 10.1016/j.jclinepi.2006.03.012 [published Online First: 2006/12/13]
77. 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: 2016/06/23]
78. Bland JM, Altman DG. Statistics Notes: Cronbach's Alpha. *BMJ: British Medical Journal* 1997;314(7080):572. doi: 10.2307/25173851
79. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 2005;30(11):1331-4. [published Online First: 2005/06/02]
80. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Controlled clinical trials* 1989;10(4):407-15. [published Online First: 1989/12/01]
81. Portney LG, Watkins MP. Foundations of clinical research : applications to practice. 3rd ed. Upper Saddle River, N.J.: Pearson/Prentice Hall 2009.
82. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7(5):541-6. doi: 10.1016/j.spinee.2007.01.008 [published Online First: 2007/04/24]
83. irr: Various Coefficients of Interrater Reliability and Agreement [program]. R package version 0.84 version, 2012.
84. R: A language and environment for statistical computing [program]. Vienna, Austria: R Foundation for Statistical Computing, 2016.
85. psych: Procedures for Personality and Psychological Research [program]. R package version 1.8.4 version. Evansville, IL, USA: Northwestern University, 2018.
86. xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files [program]. R package version 0.6.1 version, 2018.
87. psychometric: Applied Psychometric Theory [program]. R package version 2.2 version, 2010.
88. Lissitz R, Green S. Effect of the Number of Scale Points on Reliability: A Monte Carlo Approach. *Journal of Applied Psychology* 1975;60(1):10-13.
89. Guyatt GH, Townsend M, Berman LB, et al. A comparison of Likert and visual analogue scales for measuring change in function. *Journal of chronic diseases* 1987;40(12):1129-33. [published Online First: 1987/01/01]
90. Chatman AB, Hyams SP, Neel JM, et al. The Patient-Specific Functional Scale: measurement properties in patients with knee dysfunction. *Phys Ther* 1997;77(8):820-9. [published Online First: 1997/08/01]
91. McPoil TG, Martin RL, Cornwall MW, et al. Heel pain--plantar fasciitis: clinical practice guildelines linked to the international classification of function, disability, and health from the orthopaedic

section of the American Physical Therapy Association. *The Journal of orthopaedic and sports physical therapy* 2008;38(4):A1-A18. doi: 10.2519/jospt.2008.0302 [published Online First: 2008/04/25]

92. Carcia CR, Martin RL, Houck J, et al. Achilles pain, stiffness, and muscle power deficits: achilles tendinitis. *The Journal of orthopaedic and sports physical therapy* 2010;40(9):A1-26. doi: 10.2519/jospt.2010.0305 [published Online First: 2010/09/02]
93. Dolak KL, Silkman C, Medina McKeon J, et al. Hip strengthening prior to functional exercises reduces pain sooner than quadriceps strengthening in females with patellofemoral pain syndrome: a randomized clinical trial. *The Journal of orthopaedic and sports physical therapy* 2011;41(8):560-70. doi: 10.2519/jospt.2011.3499 [published Online First: 2011/06/10]
94. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *The Journal of manual & manipulative therapy* 2009;17(3):163-70. doi: 10.1179/jmt.2009.17.3.163 [published Online First: 2010/01/05]
95. Louw A, Diener I, Landers MR, et al. Preoperative pain neuroscience education for lumbar radiculopathy: a multicenter randomized controlled trial with 1-year follow-up. *Spine* 2014;39(18):1449-57. doi: 10.1097/BRS.0000000000000444 [published Online First: 2014/05/31]
96. Webster KE, Feller JA, Lambros C. Development and preliminary validation of a scale to measure the psychological impact of returning to sport following anterior cruciate ligament reconstruction surgery. *Physical therapy in sport : official journal of the Association of Chartered Physiotherapists in Sports Medicine* 2008;9(1):9-15. doi: 10.1016/j.ptsp.2007.09.003 [published Online First: 2008/12/17]
97. Johnston LH, Carroll D. The context of emotional responses to athletic injury: A qualitative analysis. *Journal of Sport Rehabilitation* 1998;7:206-20.
98. Marx RG, Bombardier C, Hogg-Johnson S, et al. Clinimetric and psychometric strategies for development of a health measurement scale. *Journal of clinical epidemiology* 1999;52(2):105-11. [published Online First: 1999/04/14]
99. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *Journal of chronic diseases* 1985;38(1):27-36. [published Online First: 1985/01/01]
100. Bland JM, Altman DG. Cronbach's alpha. *BMJ (Clinical research ed)* 1997;314(7080):572. [published Online First: 1997/02/22]
101. Garrison C, Cook C. Clinimetrics corner: the Global Rating of Change Score (GRoC) poorly correlates with functional measures and is not temporally stable. *The Journal of manual & manipulative therapy* 2012;20(4):178-81. doi: 10.1179/1066981712Z.00000000022 [published Online First: 2013/11/02]
102. Schmitt J, Abbott JH. Global ratings of change do not accurately reflect functional change over time in clinical practice. *The Journal of orthopaedic and sports physical therapy* 2015;45(2):106-11, D1-3. doi: 10.2519/jospt.2015.5247 [published Online First: 2015/01/13]
103. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *Journal of clinical epidemiology* 2008;61(2):102-9. doi: 10.1016/j.jclinepi.2007.03.012 [published Online First: 2008/01/08]
104. Melzer K, Kayser B, Pichard C. Physical activity: the health benefits outweigh the risks. *Curr Opin Clin Nutr Metab Care* 2004;7(6):641-7. [published Online First: 2004/11/10]
105. Brene S, Bjornebekk A, Aberg E, et al. Running is rewarding and antidepressive. *Physiol Behav* 2007;92(1-2):136-40. doi: 10.1016/j.physbeh.2007.05.015 [published Online First: 2007/06/15]

106. Chakravarty EF, Hubert HB, Lingala VB, et al. Reduced disability and mortality among aging runners: a 21-year longitudinal study. *Arch Intern Med* 2008;168(15):1638-46. doi: 10.1001/archinte.168.15.1638 [published Online First: 2008/08/13]
107. Kluitenberg B, van Middelkoop M, Diercks R, et al. What are the Differences in Injury Proportions Between Different Populations of Runners? A Systematic Review and Meta-Analysis. *Sports medicine (Auckland, NZ)* 2015;45(8):1143-61. doi: 10.1007/s40279-015-0331-x [published Online First: 2015/04/09]
108. Malisoux L, Nielsen RO, Urhausen A, et al. A step towards understanding the mechanisms of running-related injuries. *Journal of science and medicine in sport* 2015;18(5):523-8. doi: 10.1016/j.jsams.2014.07.014 [published Online First: 2014/09/02]
109. Nelson EO, Ryan MB, Aufderheide E, et al. Development of the University of Wisconsin Running Injury and Recovery Index. *J Orthop Sports Phys Ther* 2019(Epub ahead of print):1-33. doi: 10.2519/jospt.2019.8868 [published Online First: August 3, 2019]
110. Fairbank J. Revised Oswestry Disability questionnaire. *Spine* 2000;25(19):2552. [published Online First: 2000/10/03]
111. Kujala UM, Jaakkola LH, Koskinen SK, et al. Scoring of patellofemoral disorders. *Arthroscopy* 1993;9(2):159-63. [published Online First: 1993/01/01]
112. Guyatt GH, Osoba D, Wu AW, et al. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc* 2002;77(4):371-83. doi: 10.4065/77.4.371 [published Online First: 2002/04/09]
113. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res* 2018;27(5):1171-79. doi: 10.1007/s11136-017-1765-4 [published Online First: 2017/12/21]
114. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics* 2009;42(2):377-81. doi: 10.1016/j.jbi.2008.08.010 [published Online First: 2008/10/22]
115. Kazis LE, Selim A, Rogers W, et al. Dissemination of methods and results from the veterans health study: final comments and implications for future monitoring strategies within and outside the veterans healthcare system. *J Ambul Care Manage* 2006;29(4):310-9. [published Online First: 2006/09/21]
116. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine* 2000;25(22):2940-52; discussion 52. [published Online First: 2000/11/14]
117. Griffin DR, Parsons N, Mohtadi NG, et al. A short version of the International Hip Outcome Tool (iHOT-12) for use in routine clinical practice. *Arthroscopy* 2012;28(5):611-6; quiz 16-8. doi: 10.1016/j.arthro.2012.02.027 [published Online First: 2012/05/01]
118. Beaton DE, Bombardier C, Katz JN, et al. A taxonomy for responsiveness. *Journal of clinical epidemiology* 2001;54(12):1204-17. [published Online First: 2001/12/26]
119. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. *Journal of clinical epidemiology* 2000;53(5):459-68. [published Online First: 2000/05/17]
120. Middel B, van Sonderen E. Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research. *Int J Integr Care* 2002;2:e15. doi: 10.5334/ijic.65 [published Online First: 2006/08/10]

121. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44(3):837-45. [published Online First: 1988/09/01]
122. Baker SG, Kramer BS. Peirce, Youden, and receiver operating characteristic curves. *Am Stat* 2007;61(4):343-46. doi: 10.1198/000313007x247643
123. Spiro A, 3rd, Rogers W, Qian SX, et al. Imputing Physical and Mental Summary Scores (PCS and MCS) for the Veterans SF-12 Health Survey in the Context of Missing Data 2004.
124. pROC [program]. 1.15.0 version: CRAN, 2019.
125. Michalski MP, Gonzalez TA, Metzger MF, et al. Biomechanical Comparison of Achilles Tendon Pullout Strength Following Midline Tendon-Splitting and Endoscopic Approaches for Calcaneoplasty. *Foot Ankle Int* 2019:1071100719856939. doi: 10.1177/1071100719856939 [published Online First: 2019/06/18]
126. Grovle L, Haugen AJ, Hasvik E, et al. Patients' ratings of global perceived change during 2 years were strongly influenced by the current health status. *Journal of clinical epidemiology* 2014;67(5):508-15. doi: 10.1016/j.jclinepi.2013.12.001 [published Online First: 2014/03/07]
127. Wright AA, Cook CE, Baxter GD, et al. A comparison of 3 methodological approaches to defining major clinically important improvement of 4 performance measures in patients with hip osteoarthritis. *The Journal of orthopaedic and sports physical therapy* 2011;41(5):319-27. doi: 10.2519/jospt.2011.3515 [published Online First: 2011/02/22]
128. Gatchel RJ, Mayer TG. Testing minimal clinically important difference: consensus or conundrum? *Spine J* 2010;10(4):321-7. doi: 10.1016/j.spinee.2009.10.015 [published Online First: 2010/04/07]
129. Crossley KM, Macri EM, Cowan SM, et al. The patellofemoral pain and osteoarthritis subscale of the KOOS (KOOS-PF): development and validation using the COSMIN checklist. *British journal of sports medicine* 2018;52(17):1130-36. doi: 10.1136/bjsports-2016-096776 [published Online First: 2017/03/05]
130. Powney M, Williamson P, Kirkham J, et al. A review of the handling of missing longitudinal outcome data in clinical trials. *Trials* 2014;15:237. doi: 10.1186/1745-6215-15-237 [published Online First: 2014/06/21]
131. Valier AR, Jennings AL, Parsons JT, et al. Benefits of and barriers to using patient-rated outcome measures in athletic training. *J Athl Train* 2014;49(5):674-83. doi: 10.4085/1062-6050-49.3.15 [published Online First: 2014/08/08]
132. Dideriksen M, Soegaard C, Nielsen RO. Validity of Self-Reported Running Distance. *J Strength Cond Res* 2016;30(6):1592-6. doi: 10.1519/JSC.0000000000001244 [published Online First: 2015/10/20]
133. Selim AJ, Rogers W, Fleishman JA, et al. Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). *Qual Life Res* 2009;18(1):43-52. doi: 10.1007/s11136-008-9418-2 [published Online First: 2008/12/04]
134. Ware JE, Jr., Bayliss MS, Rogers WH, et al. Differences in 4-year health outcomes for elderly and poor, chronically ill patients treated in HMO and fee-for-service systems. Results from the Medical Outcomes Study. *JAMA* 1996;276(13):1039-47. [published Online First: 1996/10/02]
135. Martin RL, Kivlan BR, Christoforetti JJ, et al. Minimal Clinically Important Difference and Substantial Clinical Benefit Values for the 12-Item International Hip Outcome Tool. *Arthroscopy* 2019;35(2):411-16. doi: 10.1016/j.arthro.2018.09.028 [published Online First: 2019/01/08]

136. Snedden TR, Scerpella J, Kliethermes SA, et al. Sport and Physical Activity Level Impacts Health-Related Quality of Life Among Collegiate Students. *Am J Health Promot* 2019;33(5):675-82. doi: 10.1177/0890117118817715 [published Online First: 2018/12/28]
137. Houston MN, Hoch MC, Hoch JM. Health-Related Quality of Life in Athletes: A Systematic Review With Meta-Analysis. *J Athl Train* 2016;51(6):442-53. doi: 10.4085/1062-6050-51.7.03 [published Online First: 2016/06/04]
138. Snyder AR, Martinez JC, Bay RC, et al. Health-related quality of life differs between adolescent athletes and adolescent nonathletes. *J Sport Rehabil* 2010;19(3):237-48. [published Online First: 2010/09/03]
139. McGuine TA, Winterstein AP, Carr K, et al. Changes in Health-Related Quality of Life and Knee Function After Knee Injury in Young Female Athletes. *Orthop J Sports Med* 2014;2(4):2325967114530988. doi: 10.1177/2325967114530988 [published Online First: 2014/04/01]
140. Valovich McLeod TC, Bay RC, Parsons JT, et al. Recent injury and health-related quality of life in adolescent athletes. *J Athl Train* 2009;44(6):603-10. doi: 10.4085/1062-6050-44.6.603 [published Online First: 2009/11/17]
141. Filbay SR, Ackerman IN, Russell TG, et al. Return to sport matters-longer-term quality of life after ACL reconstruction in people with knee difficulties. *Scandinavian journal of medicine & science in sports* 2017;27(5):514-24. doi: 10.1111/sms.12698 [published Online First: 2016/05/12]
142. Oja P, Titze S, Kokko S, et al. Health benefits of different sport disciplines for adults: systematic review of observational and intervention studies with meta-analysis. *British journal of sports medicine* 2015;49(7):434-40. doi: 10.1136/bjsports-2014-093885 [published Online First: 2015/01/09]
143. Hespanhol Junior LC, Pillay JD, van Mechelen W, et al. Meta-Analyses of the Effects of Habitual Running on Indices of Health in Physically Inactive Adults. *Sports medicine (Auckland, NZ)* 2015;45(10):1455-68. doi: 10.1007/s40279-015-0359-y [published Online First: 2015/07/17]
144. Lee DC, Brellenthin AG, Thompson PD, et al. Running as a Key Lifestyle Medicine for Longevity. *Prog Cardiovasc Dis* 2017;60(1):45-55. doi: 10.1016/j.pcad.2017.03.005 [published Online First: 2017/04/04]
145. Brown DW, Balluz LS, Heath GW, et al. Associations between recommended levels of physical activity and health-related quality of life. Findings from the 2001 Behavioral Risk Factor Surveillance System (BRFSS) survey. *Prev Med* 2003;37(5):520-8. [published Online First: 2003/10/24]
146. Eime R, Harvey J, Payne W. Dose-response of women's health-related quality of life (HRQoL) and life satisfaction to physical activity. *J Phys Act Health* 2014;11(2):330-8. doi: 10.1123/jpah.2012-0073 [published Online First: 2013/02/01]
147. Hsiao CJ, Dymek C, Kim B, et al. Advancing the use of patient-reported outcomes in practice: understanding challenges, opportunities, and the potential of health information technology. *Qual Life Res* 2019;28(6):1575-83. doi: 10.1007/s11136-019-02112-0 [published Online First: 2019/01/27]
148. Philpot LM, Barnes SA, Brown RM, et al. Barriers and Benefits to the Use of Patient-Reported Outcome Measures in Routine Clinical Care: A Qualitative Study. *Am J Med Qual* 2018;33(4):359-64. doi: 10.1177/1062860617745986 [published Online First: 2017/12/21]
149. Wailoo AJ, Hernandez-Alava M, Manca A, et al. Mapping to Estimate Health-State Utility from Non-Preference-Based Outcome Measures: An ISPOR Good Practices for Outcomes Research Task

Force Report. *Value Health* 2017;20(1):18-27. doi: 10.1016/j.jval.2016.11.006 [published Online First: 2017/02/19]

150. Evaluation Metrics for Machine Learning [program]. 0.1.4 version. CRAN, 2018.
151. Fayers PM, Hays RD. Should linking replace regression when mapping from profile-based measures to preference-based measures? *Value Health* 2014;17(2):261-5. doi: 10.1016/j.jval.2013.12.002 [published Online First: 2014/03/19]
152. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates 1988.
153. Brazier JE, Yang Y, Tsuchiya A, et al. A review of studies mapping (or cross walking) non-preference based measures of health to generic preference-based measures. *Eur J Health Econ* 2010;11(2):215-25. doi: 10.1007/s10198-009-0168-z [published Online First: 2009/07/09]
154. Batmyagmar D, Kundi M, Ponocny-Seliger E, et al. High intensity endurance training is associated with better quality of life, but not with improved cognitive functions in elderly marathon runners. *Sci Rep* 2019;9(1):4629. doi: 10.1038/s41598-019-41010-w [published Online First: 2019/03/16]
155. Acree LS, Longfors J, Fjeldstad AS, et al. Physical activity is related to quality of life in older adults. *Health Qual Life Outcomes* 2006;4:37. doi: 10.1186/1477-7525-4-37 [published Online First: 2006/07/04]
156. Houston MN, Hoch JM, Van Lunen BL, et al. The Impact of Injury on Health-Related Quality of Life in College Athletes. *J Sport Rehabil* 2017;26(5):365-75. doi: 10.1123/jsr.2016-0011 [published Online First: 2016/09/17]
157. Phillips SM, Wojcicki TR, McAuley E. Physical activity and quality of life in older adults: an 18-month panel analysis. *Qual Life Res* 2013;22(7):1647-54. doi: 10.1007/s11136-012-0319-z [published Online First: 2012/11/20]
158. Gobbens RJ, Remmen R. The effects of sociodemographic factors on quality of life among people aged 50 years or older are not unequivocal: comparing SF-12, WHOQOL-BREF, and WHOQOL-OLD. *Clin Interv Aging* 2019;14:231-39. doi: 10.2147/cia.S189560 [published Online First: 2019/02/23]
159. Mayo NE, Moriello C, Asano M, et al. The extent to which common health-related quality of life indices capture constructs beyond symptoms and function. *Qual Life Res* 2011;20(5):621-7. doi: 10.1007/s11136-010-9801-7 [published Online First: 2010/11/26]
160. Farivar SS, Cunningham WE, Hays RD. Correlated physical and mental health summary scores for the SF-36 and SF-12 Health Survey, V.I. *Health Qual Life Outcomes* 2007;5:54. doi: 10.1186/1477-7525-5-54 [published Online First: 2007/09/11]
161. Lubetkin EI, Jia H, Franks P, et al. Relationship among sociodemographic factors, clinical conditions, and health-related quality of life: examining the EQ-5D in the U.S. general population. *Qual Life Res* 2005;14(10):2187-96. doi: 10.1007/s11136-005-8028-5 [published Online First: 2005/12/06]
162. Hilliard RC, Blom L, Hankemeier D, et al. Exploring the Relationship Between Athletic Identity and Beliefs About Rehabilitation Overadherence in College Athletes. *J Sport Rehabil* 2017;26(3):208-20. doi: 10.1123/jsr.2015-0134 [published Online First: 2016/09/17]
163. Willy RW. Innovations and pitfalls in the use of wearable devices in the prevention and rehabilitation of running related injuries. *Physical therapy in sport : official journal of the Association of Chartered Physiotherapists in Sports Medicine* 2018;29:26-33. doi: 10.1016/j.ptsp.2017.10.003 [published Online First: 2017/11/25]
164. Nielsen R. Launching the Garmin-RUNSAFE Running Health Study, 2019.

165. McGuine TA, Winterstein A, Carr K, et al. Changes in self-reported knee function and health-related quality of life after knee injury in female athletes. *Clinical journal of sport medicine : official journal of the Canadian Academy of Sport Medicine* 2012;22(4):334-40. doi: 10.1097/JSM.0b013e318257a40b [published Online First: 2012/05/26]

Appendix 1. The University of Wisconsin Running Injury & Recovery Index

University of Wisconsin Running Injury & Recovery Index

Instructions: Consider your current running injury over the **past 7 days** when answering each question; check (☒) the appropriate box.

1. How does your running injury impact your ability to perform daily activities?	☐ No impact	☐ Slightly impact	☐ Moderately impact	☐ Significantly impact	☐ Unable to perform
2. How frustrated are you by your running injury?	☐ Not frustrated	☐ Mildly frustrated	☐ Moderately frustrated	☐ Significantly frustrated	☐ Extremely frustrated
3. How much recovery have you made from your running injury?	☐ Complete recovery	☐ Significant recovery	☐ Moderate recovery	☐ Minimal recovery	☐ No recovery
4. How much pain do you experience while running?	☐ No pain	☐ Minimal pain	☐ Moderate pain	☐ Significant pain	☐ Unable to run
5. How much pain do you experience during the 24 hours following a run?	☐ No pain	☐ Minimal pain	☐ Moderate pain	☐ Significant pain	☐ Unable to run
6. How has your weekly mileage or weekly running time changed as a result of your injury?	☐ Same or greater than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
7. How has the distance of your longest weekly run changed as a result of your injury?	☐ Same or longer than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
8. How has your running pace or speed changed as a result of your injury?	☐ Same or faster than before my injury	☐ Minimally reduced	☐ Moderately reduced	☐ Significantly reduced	☐ Unable to run
9. How does your injury affect your confidence to increase the duration or intensity of your running?	☐ Confident to increase my running	☐ If I increase I might be fine	☐ Neutral	☐ If I increase I might get worse	☐ I cannot increase my running

For office use:
Add the scores for the 9 questions to create one total score. Total Score: ___________

Scoring Key:	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0

Appendix 2. Abstracts from this thesis

Higher Fear-avoidance Beliefs are Associated with Lower Running Ability in Injured Runners

Maschke B, Nelson EO, Heiderscheit B, Obermeier M, Palmsten A, Yungtum W, Chmielewski TL

(Note: Accepted to the American Physical Therapy Association Combined Sections Meeting 2019)

Purpose/Hypothesis: Psychological distress can negatively influence rehabilitation outcomes. Methods to identify psychological distress have been developed, but have not been examined in injured runners. The purpose was to evaluate psychological distress in injured runners and its association with running ability. The hypothesis was that higher psychological distress would correlate with lower running ability.

Number of Subjects: Twenty-seven distance runners (8 male; mean age, 33.9 years; range, 15-59 years) seen in physical therapy for evaluation of a running-related injury.

Materials and Methods: Questionnaires completed at the initial physical therapy visit included the University of Wisconsin Running Injury and Recovery Index (UWRI), Athletic Fear Avoidance Questionnaire (AFAQ), and Optimal Screening for Prediction of Referral and Outcome-Yellow Flag Screening Tool (OSPRO-YF). The UWRI includes 9 items that quantify an individual's ability to run. The AFAQ includes 10 items related to fear-avoidance beliefs in athletes. The OSPRO-YF assesses pain-related psychological distress (ie, yellow flags) in 11 subscales divided into 3 domains: negative mood, fear avoidance, and positive coping. Scores that fall in the top quartile for negative or bottom quartile for positive psychological questionnaires are defined as yellow flags. Pearson correlation coefficient examined the association of AFAQ and UWRI scores. Spearman rank correlation examined the association of the number of yellow flags on OSPRO-YF with UWRI score.

Results: Mean ± SD scores were UWRI, 14.8 ± 6.0 points and AFAQ, 22.6 ± 7.8 points. The median number of OSPRO-YF yellow flags was 2 (interquartile range, 1-6). Yellow flags were most common on

the State Trait Anxiety Index (23/27, 85% of the sample) and least common on State Trait Anger Expression Inventory, Fear-Avoidance Beliefs Questionnaire-Work, and Pain Catastrophizing Scale (4/27, 14% of the sample). AFAQ score was significantly negatively correlated with the UWRI score ($r = -0.44$, $P = .023$). The number of yellow flags on OSPRO-YF was not correlated with the UWRI score ($r = 0.09$, $P = .651$).

Conclusions: Among injured runners, higher fear avoidance, measured by the AFAQ, was associated with lower self-reported running ability, but yellow flags on the OSPRO-YF were not. Most of the sample met the criteria for a yellow flag related to anxiety.

Clinical Relevance: Addressing high fear-avoidance beliefs in injured runners along with physical impairments may be a necessary part of rehabilitation. The lack of association between current running ability and the number of yellow flags may indicate the OSPRO-YF is more suited to a general population, but it is yet unknown if yellow flags predict rehabilitation outcomes in runners. Psychological distress related to anxiety in this study is consistent with other studies showing anxiety in healthy or injured runners who are restricted from running.

Improved Psychological Distress Is Associated with Improved Running Ability in Injured Runners

Maschke B, Nelson E, Palmsten A, Heiderscheit B, Obermeier M, Russell H, Chmielewski TL

(Note: Submitted to the American Physical Therapy Association Combined Sections Meeting 2020)

Purpose/Hypothesis: Psychological distress can negatively influence rehabilitation outcomes in injured athletes. Few studies have focused on injured runners. The purpose of this study was to examine psychological distress and self-reported running ability over 12 weeks in injured runners. The hypotheses were: 1) running ability and psychological distress would improve over 12 weeks, 2) high baseline psychological distress would be associated with poorer running ability at 12 weeks, and 3) across 12 weeks, decreased psychological distress would be associated with improved running ability.

Number of Subjects: 32 runners attending physical therapy for a running-related injury; 10 males, mean±SD age = 36.7±11.9 years.

Materials and Methods: Questionnaires completed at physical therapy consult (baseline) and 12 weeks later were University of Wisconsin Running Injury and Recovery Index (UWRI), Athletic Fear Avoidance Questionnaire (AFAQ), and Optimal Screening for Prediction of Referral and Outcome – Yellow Flag Screening Tool (OSPRO-YF). UWRI includes nine items (0-36 points), and higher scores mean better running ability. AFAQ includes ten items (10-50 points), and higher scores mean greater fear-avoidance beliefs in athletics. OSPRO-YF includes 17 items that estimate scores on 11 psychological questionnaires. Item scores in the top or bottom quartile for negative or positive psychological questionnaires, respectively, are assigned a yellow flag (0-11 total), and items scores are also summed (6-89 points); more yellow flags or higher score means higher psychological distress. Change scores (12 week – baseline) were computed for all measures. Paired t-tests (UWRI, AFAQ and OSPRO-YF sum) or Wilcoxon test (OSPRO-YF) compared baseline and 12 week scores. Inter-variable associations were examined with

Pearson or Spearman Rank correlation, as appropriate, for baseline AFAQ, OSPRO-YF and OSPRO-YF sum scores with 12 week UWRI score and among change scores.

Results: All questionnaire scores significantly improved from baseline to 12 weeks (UWRI: 15.8±5.5 vs 25.7±7.9; AFAQ: 21.0±8.2 vs 18.7±7.1; OSPRO-YF sum: 31.2±9.8 vs 27.2±9.6; all $P<0.05$). Median yellow flags on OSPRO-YF improved from 2 to 1 ($P=0.02$). Baseline AFAQ, OSPRO-YF and OSPRO-YF sum were not correlated with 12 week UWRI. Change in AFAQ and OSPRO-YF sum score were negatively correlated with change in the UWRI ($r=-0.55$ and $r=-0.41$, $P<0.05$; respectively).

Conclusions: Psychological distress and running ability improved over 12 weeks. Psychological distress at baseline was not associated with running ability 12 weeks. Conversely, improved psychological distress over 12 weeks was associated with improved running ability. Change in OSPRO-YF sum score, not yellow flags, was correlated with change in running ability.

Clinical Relevance: Psychological distress is present in injured runners and improves as running ability improves; thus, recurrent screening of psychological distress and running ability may be necessary. AFAQ and OSPRO-YF sum score may give meaningful insight into psychological distress in injured runners.

Prospective Validation and Responsiveness of the University of Wisconsin Running Injury and Recovery Index

Nelson EO, Kliethermes S, Heiderscheit BC

(Note: Submitted to the American Physical Therapy Association Combined Sections Meeting 2020)

Purpose/Hypothesis: The University of Wisconsin Running Injury and Recovery Index (UWRI) is the first patient-reported outcome measure (PRO) to assess running ability following running-related injury (RRI), demonstrating excellent reliability ($ICC_{agreement}$ =0.93). By incorporating the multiple internal, external, or psychological factors influencing RRI recovery, the UWRI enables the magnitude of recovery or treatment response to be quantified during clinical assessments and research studies. The purpose of this prospective study was to analyze the validity and responsiveness of the UWRI as an evaluative PRO.

Number of Subjects: 396 participants (62.1% female; mean±SD age 34.5±12.4 y; running experience 12.7±10.4 y, 98.5% recreational or competitive runners)

Materials and Methods: Runners seeking care from a physical therapist for a RRI electronically completed PROs at the initial evaluation and 12 weeks later. UWRI change was validated against the global rating of change (GROC), VR-12 change, and change in body region PROs. Responsiveness was evaluated with anchor-based and distribution-based techniques. T-tests evaluated differences in continuous variables and Pearson correlation coefficients were used to assess the strength of association.

Results: UWRI change (7.7±9.0, paired t-test p<0.001) was correlated with GROC (r=0.67; 95% CI [0.57, 0.75]; P<0.001), as well as the changes in VR-12 physical component score (PCS) (r =0.54; 95% CI [0.42, 0.64]; P<0.001) and mental component score (MCS) (r=0.31; 95% CI [0.16, 0.44]; P<0.001). UWRI change was correlated with changes in the sport subscale of Foot and Ankle Ability Measure (r=0.75; 95% CI [0.63, 0.84]; P<0.001), 12-item international hip outcome tool (r=0.75; 95% CI [0.55, 0.87]; p<0.001), and anterior knee pain scale (r=0.48; 95% CI [0.21, 0.68]; P=0.001), but not associated with the Oswestry

Disability Index 2.0 (*r*=0.05; 95% CI [-0.57, 0.63]; *P*=0.89). The UWRI effect size was 1.15, standardized response mean difference was 0.87, responsiveness index was 1.33, minimal detectable change 95% CI was 4.75, and ceiling or floor effects were not present. UWRI change was significantly different in runners reporting significant improvement (12.2±5.9), slight improvement (7.1±6.6), no change (0.0±9.1), and worsening (-14.6±7.4) on the GROC. The UWRI minimal detectable change 95% CI was 5 points and the minimal clinical important difference was 8 points.

Conclusions: The UWRI demonstrated longitudinal validity (GROC), convergent validity (PCS and body regions PROs), and divergent validity (MCS), as well as excellent internal and external responsiveness. The UWRI displayed excellent psychometric integrity and the ability to differentiate graduated levels of running ability during RRI recovery.

Clinical relevance: The UWRI is the only PRO validated in a population of individuals seeking care for a RRI. The UWRI is responsive to changes in running ability, enabling the comparison of patient-perceived ability with objective measures of training load or biomechanical assessments. Novice and professional runners should be targeted during future psychometric assessments.