Quantification of Walking Ability in Participants with Neurogenic Claudication from Lumbar Spinal Stenosis – A Comparative Study

James Rainville, MD[a,b], Lisa Childs, PT[b], Enrique Peña, MD[c], Pradeep Suri, MD[a,b,d,e],

Janet Limke, MD[a,b], Cristin Jouve, MD[a,b], David J Hunter, MD[f]

[a] Department of Physical Medicine and Rehabilitation, Harvard Medical School, Boston,

MA; [b] New England Baptist Hospital, 125 Parker Hill Avenue, Boston, MA 02120;

[c] Seton Spine & Scoliosis Center 1600 West 38th St Austin, TX 78731; [d] VA Boston

Healthcare System, Boston, Ma; [e] Spaulding Rehabilitation Hospital, Boston, MA;

[f] Northern Clinical School, University of Sydney, Sydney, Australia

Corresponding author: James Rainville, MD,

New England Baptist Hospital,

125 Parker Hill Avenue,

Boston, MA 02120

617-754-5246, Fax 617-754-6332, jrainvil@caregroup.harvard.edu

Abstract :

Background Context: Walking limitations caused by neurogenic claudication (NC) are typically assessed with self-reported measures, though objective evaluation of walking utilizing motorized treadmill test (MTT) or self-paced walking tests (SPWT) have periodically appeared in the lumbar spinal stenosis (LSS) literature.

Purpose: This study compared the validity and responsiveness of MTT and SPWT for assessing walking ability before and after common treatments for NC.

Study Design: Prospective, observational cohort study

Patient Sample: 50 adults were recruited from an urban spine center if they had LSS, substantial walking limitations from NC, and were scheduled to undergo surgery (20%) or conservative treatment (80%).

Outcome Measures: Walking times, distances and speeds along with characteristics of NC symptoms were recorded for MTT and SPWT. Self-reported measures included back and leg pain intensity assessed with $0-10$ numeric pain scale, disability assessed with Oswestry Disability Index (ODI), walking ability assessed with estimated walking times and distances, and subscales from the Spinal Stenosis Questionnaires (SSQ).

Methods: MTT used a level track, and SPWT were conducted in a rectangular hallway. Walking speeds were self-selected and test endpoints were: a) NC, b) fatigue, or c) completed the 30-minute test protocol. Results from MTT and SPWT were compared with each other and with self-reported measures. Internal responsiveness was assessed by comparing changes in initial to post-treatments results, and external responsiveness by comparing walking test results between those that improved with those that did not improve by self-report criteria.

Results: Mean age was 68 years. 58% were male. NC included leg pain (88%), and buttock(s) pain (12%). 5 participants could not safely perform MTT. Walking speeds were faster and distances were greater with SPWT, although results from both tests correlated with each other and with self-reported measures. 72% of participants reported improvement following treatment which was confirmed by significant mean differences in self-reported measures. MTT results did not demonstrate internal responsiveness to change in clinical status following treatment, but SPWT did with increased mean walking times (6 min) and distances (387 m). When responsiveness was assessed against external criterion, both SPWT and MTT demonstrated substantial divergence with self-reported changes in clinical status and alternative outcome measures.

Conclusions: Both MTT and SPWT can quantify walking abilities in NC. As outcome tools, SPWT demonstrated better internal responsive than MTT, but neither test demonstrated external responsiveness. Neither test should be considered as a meaningful substitution for disease specific measures of function.

INTRODUCTION

With advancing age, spinal degeneration can cause substantial structural changes in the lumbar spine. In some adults these changes result in progressive narrowing of the spinal canal and compression of the lumbar nerve roots - a condition referred to as lumbar spinal stenosis (LSS) [1-3]. In population-based studies, moderate LSS is noted in up to 40 percent of adults over the age of 60, though most do not report symptoms [4].

For symptomatic adults with imaging studies demonstrating LSS, neurogenic claudication (NC) is the symptom complex that most spine experts agree is attributable to LSS [5]. NC is defined as intermittent pain radiating to the buttock(s), thigh(s) and/or lower leg(s) that is induced with standing, walking and/or lumbar extension, and relieved with sitting, lying down or lumbar flexion [5]. When of severe intensity, NC causes considerable limitations for walking [6,7]. These limitations drive many elderly adults to seek medical care, and are the most frequent reasons for lumbar spine surgery in Medicare recipients [8].

Because limited walking is the most relevant area of impaired function for patients with NC, improvement of walking abilities is the primary goal of most treatments [5]. For this reason, assessment of patients' ability to walk is a standard part of the evaluation of NC. Typically, in both clinical practice and research, walking limitations are assessed through direct questioning (example: "How far or long can you walk on even ground without a

break?) [9,10], or through disease-specific disability questionnaires [11-13]. Although self-reported measures of walking are certainly meaningful, actual quantification of walking ability would seem to be a useful endeavor for assessment of NC, and may aid in the objective evaluation of the effects of treatments. To date, motorized treadmill tests (MTT) are the most frequently cited methods for quantifying walking ability of patients with NC [10,14-24].

Recently, several observational studies have reported on self-paced walking tests (SPWT) that directly observe walking time and distance as patients with NC walk on level walking courses [25-29]. SPWT has obvious validity, as it directly observes the function of interest under conditions representative of real world settings, such as walking in a shopping mall, and therefore meets the first requirement of an outcome measure for clinical trials [25, 26]. Because of its validity, SPWT results have been used as a criterion standard to assess the accuracy of other measures for quantifying walking limitation caused by NC, including MTT [29], and self-reported measures [30]. SPWT also has adequate test-retest reproducibility [29], and thus meet the second requirement of an outcome measure for use in clinical trials.

The third requirement of an outcome measure is responsiveness, and this has not been explored for SPWT. Responsiveness has two major aspects [31]. The first aspect is internal responsiveness, which assesses the ability of a measure to change over a pre-specified time frame, such as before and after an intervention. Since internal

responsiveness evaluates the ability of a measure to detect change, it is dependent on the effectiveness of the intervention utilized to induce change. Of importance, internal responsiveness is assessed independent of changes in other outcome measures, and therefore it does not necessarily speak to the relevancy of the detected changes compared to overall changes in clinical status. The second aspect of responsiveness, external responsiveness, reflects the relevance of changes detected by a measure by evaluating its relationship to corresponding changes in a reference standard of clinical status. External responsiveness is dependent solely on the choice of the reference standard, and not on the treatment under investigation. External responsiveness allows the assessment of whether a change in the measure can be viewed as an accepted indicator of a meaningful change in the condition of a patient, and the appropriateness of using the measure as a substitute for alternative reference standards [31].

The purpose of this study was to explore the utility of SPWT and MTT for evaluating walking limitations produced by NC in a group of patients undergoing common treatments for this condition. Specific factors under study included comparison of each tests ability to quantify walking limitation produced by NC and evaluation of internal and external responsiveness.

MATERIALS AND METHODS

Participants

This study received Investigational Review Board approval and all participants signed an informed consent.

Participants were recruited from a spine center of an urban, academic hospital by 4 physiatrists with extensive experience in treating spinal disorders. Inclusion criteria were: 1) degenerative lumbar spinal stenosis documented by lumbar MRI or CT scan; 2) walking induced NC with or without concurrent neurological symptoms of weakness, sensory loss or impaired balance [5]; 3) self-reported walking ability limited by NC to 30 minutes or less; 4) duration of NC of at least 3 months; and 5) scheduled to undergo physical therapy, spinal injections or surgery as treatment for NC. Exclusion criteria were: 1) symptoms of buttocks and leg pain not aggravated by walking; 2) concurrent acute disc herniation as the probable cause of symptoms; 3) spinal stenosis caused by non-degenerative spinal disorders (neoplasm, metabolic bone disease, vertebral fracture); 4) non-palpable dorsalis pedis and posterior tibial pulses in the symptomatic leg(s) suggesting possible peripheral vascular insufficiency; 5) symptomatic arthritis of the hip, knee, ankle or foot causing limitation in walking; 6) neurological disease affecting ambulation (Parkinson's disease, myelopathy, stroke); 7) cardiac or pulmonary disease that limits walking; 8) severe cognitive difficulties; 9) general frailty that would make participating in the walking test unsafe.

Assessments
The enrolling physiatrists complete a study questionnaire that recorded demographics, duration, location and intensity of symptoms, physical examination findings, burden of

co-morbidities (range 0-42 with higher scores for increasing burdens of co-morbidities) [32], and the level(s) of spinal stenosis. Spinal stenosis of the central canal was defined as reduction of the CSF signal in MRI T2 axial and sagittal images, and the degree of stenosis defined as Grade 1 for 1/3 reduction of canal area, Grade 2 for 1/3 to 2/3 reduction of canal area, and Grade 3 for greater than 2/3 reduction of canal area [33].

All participants completed the following paper and pencil measures. NC induced limitations in walking were assessed by questions that asked participants to estimate the 1) distance (in feet, yards or miles) and 2) time (in minutes or hours) that they could walk without a break on even ground in a typical community setting before symptoms become intolerable [10]. Participants were also asked to estimate the intensity of NC back and leg pain induced by walking in community settings over the last week utilizing an 11-point numerical pain scale anchored with 0 (no pain) and 10 (worst possible pain) [34]. NC symptoms were also assessed using the validated Spinal Stenosis Questionnaire (SSQ) Symptom Severity and Physical Function Subscales [12,13,35,36]. These subscales are scored between 1 and 5 with higher values indicate greater symptoms and physical limitations. Finally, total scores from Oswestry Disability Index Version 2.0 (ODI) (range 0-100% disabled), and scores from the ODI item assessing walking ability (range 0-no limitations to 5- can walk only a few steps) were recorded [11].

Changes in clinical status were assessed using two methods. Global change in clinical status was assessed with the following question. "Since treatment, how have your

symptoms with walking changed? (Choose 1) Completely resolved; Improved but are still present; No changed; Worsened" Additionally, at the time of final testing, participants completed the numeric pain scales, self-reported walking distance and walking time questions, SSQ symptom severity, SSQ physical function and ODI. These results were contrasted with initial results and changes in scores were calculated. Participants also completed the SSQ satisfaction subscale with range of scores from1 – most satisfied to 4 – very dissatisfied.

Walking tests

Walking tests were performed at a hospital-affiliated physical therapy facility located in a medical office building. Testing sessions were scheduled within one week of study entry, and before any treatments for NC were administered. SPWT and MTT were administered by the same research physical therapist on the same day separated by at least a 5 minute rest or until all symptoms from the first test had resolved. The sequence of walking tests during all test sessions was determined by a random number table that pre-assigned the first test (SPWT or MTT) based on the order of subject enrollment.

The SPWT was conducted on a 52 meter (170 feet) rectangular course in the carpeted corridors of the medical building in which the physical therapy site was located. Chairs were positioned at three locations along the walking course so that participants would always be within 10 meters of a seat in case they felt unable to stand or walk. Study participants sat in a chair at the starting line, and the test began when they stood and

began walking around the course at a self-selected pace. The research physical therapist walked approximately one meter behind the participants during the entire test, and recorded the walking time with a stopwatch and laps with a hand held mechanical counter. Distances of partial laps were estimated from pre-placed marks every 10 meters along the course.

MTT was conducted with the incline of the treadmill set at zero degrees. In order to prevent participants from potentially improving their walking by bending forward [14], participants were not allowed to place both hands on the handrails for support, but were allowed to use one hand on the handrail for balance, if needed. Walking speed was determined during a pre-test walk on the treadmill during which the speed was adjusted until the subject reported that they were walking at their desired speed. After the brief pre-test, participants sat for at least two minutes before the actual test began. The time of walking was measured with a stopwatch. For repeat testing, the original walking speed was selected at beginning of the pre-test walk on the treadmill and adjusted per request of participants during the test, until the desired walking speed was found.

For both SPWT and MTT, the ability of the walking test to produce NC was recorded as a "positive" result. The test was terminated when their symptoms reached the intensity at which participants would usually stop walking in a community setting. The tests were stopped immediately by the participants for reasons of fatigue, and the results were recorded as "negative - fatigue". For participants that completed the 30 minute protocol

without developing symptoms, results were recorded as "negative – completed without symptoms". Results for each test were recorded for walking time in minutes (min.), speed in kilometers per hour (km/h), and distance in meters (m). Following completion of each walking test, participants completed a questionnaire inquiring about the characteristics of NC symptoms (pain location, pain intensity, presence of neurological symptoms).

Interventions

This study utilized typical treatments for neurogenic claudication as the interventions to produce change of status against which both internal and external responsiveness could be assessed. Because our goal was to assess responsiveness of SPWT and MTT, and not treatment, we included participants undergoing non-surgical (exercise-oriented physical therapy, lumbar spine injections with corticosteroids, or both) and surgical treatments (decompression with or without fusion) for neurogenic claudication. It was assumed that this would lead to a wide range of changes in clinical status, which when measured, would be adequate for assessing responsiveness of SPWT and MTT. Assessment of the effectiveness of any particular treatments was not an objective of this study, and therefore not performed. For participants undergoing physical therapy, post-treatment testing was done at discharge from therapy (average therapy time of 6 weeks).). For participants undergoing spinal injections only, reassessment was done at 4 weeks after injections. For participants undergoing spine surgery, reassessment was done 3 months after surgery.

Statistical Methods

Data analysis was performed with SPSS 14.0 (SPSS, Chicago, IL).

To explore their ability to assess NC, results of initial walking times, distances, and walking speeds were compared between SPWT and MTT using paired-sample t-tests and intraclass correlation coefficients (ICC). ICC was chosen for this analysis as it assesses the conformity of observations by two different measures of the same variable expressed in the same units [37]. The κ (kappa) coefficient was used to measure the agreement in the ability of the two walking tests to reproduce NC symptoms [38]. Pearson product moment correlations were used to compare results from SPWT and MTT with self-reported walking measures.

This study utilized two statistical methods to evaluate internal responsiveness [34]. First, paired sample *t*-test were calculated to assess change between initial and final SPWT and MTT scores. Next, standardized effect size (ES) were used to assess the magnitude of changes in SPWT and MTT compared to the variance (standard deviation) in those measures at baseline assessment [39]. The values for ES are commonly interpreted as .20, .50 and .80 or greater as indication small, moderate or large internal responsiveness, respectively.

External responsiveness was assessed using 2 statistical methods. Receiver operating characteristics curves (ROC) were plotted and area under the curve (AUC) were calculated to explore the degree to which specific values of change scores for SPWT and MTT walking times and distances correctly reflect the reference standard of improved or not improved clinical status [40]. This reference standard was derived from lumping responses to the change in clinical status question into 2 categories - improved (combined 'completely resolved'; 'improved but are still present') and not improved (combined 'unchanged' and 'worsened'). The validity of this reference standard was first confirmed by comparing the change scores of low back pain, leg pain, ODI, SSQ symptom, and SSQ function, and raw scores of SSQ satisfaction between participants classified as improved and not improved using independent sample t-test. The external responsiveness of the walking test related to alternative outcomes was examined by comparing change scores of SPWT and MTT with change scores for self-report walking, SSQ physical function and ODI walking score using Pearson product moment correlations.

RESULTS

Fifty participants participated in this study. Their average age was 68 years old (standard deviation [SD] 7.9, range 48-86) and the average duration of NC was 18 months (SD 21). The majority of participants were white (92%), male (58%), retired (54%), and highly educated (16.7 years, SD 3.4). Mean co-morbidity score was 5.1 (SD 4.5). Lumbar spine imaging was by MRI in 94% of participants and the remainder had CT. L3-4 and L4-5 were the most common levels of spinal stenosis (56% and 82 % respectively) with ≥2

stenotic levels noted in 42% of participants. Stenosis was rated as severe in 54% and moderate in 38% of participants. Neurogenic claudication symptoms included leg pain in 88% of participants (bilateral in 54%), with buttocks pain in the remaining 12%. Walking induced low back pain was reported by 82%. Mean self-reported intensity of NC leg pain was 5.7 (SD 2.6) and back pain was 6.0 (SD 2.5). Eighty-two percent of participants reported that walking produced one or more neurological symptoms. These included paresthesias (58%), leg weakness (38%) and unsteadiness (38%).

Initial Walking Test Results

Twenty seven participants were randomly assigned to first undergo MTT, and the remaining 23 to first undergo SPWT. Five participants could not walk safely on the treadmill without both hands on the handrail, and therefore did not undergo MTT. These 5 participants were older, more disabled, and had higher comorbidity scores than the participants that could safely walk on a treadmill. Of the 45 participants undergoing MTT, 12 were able to walk the full 30 minutes, with 5 completing the test without symptoms. All 50 participants underwent SPWT, with 12 walking the full 30 minutes, and 2 completing the test without symptoms.

For the 45 participants that completed both walking tests, results demonstrated that both SPWT and MTT have similar abilities for quantifying walking and reproducing symptoms in patients with NC (Table 1). Walking times were similar for both tests and

showed a high correlation (ICC .79, $P \leq .001$). Walking speeds were faster and walking distances greater during SPWT, though results showed high correlations between the two tests (walking speed ICC .50, $P \leq .001$, walking distance ICC .84, $P \leq .001$). Both test had similar abilities to reproduced NC symptoms with modest agreement between test for walking induced back pain ($\kappa = .43$, $P \leq .01$), leg pain ($\kappa = .45$, $P \leq .01$), paresthesias ($\kappa = .77$, $P \leq .001$), leg weakness ($\kappa = .69$, $P \leq .001$), and unsteadiness ($\kappa = .79$, $P \leq .001$). Self-reported estimated walking times (mean 16.5 minutes) and distances (mean 778 meters) were statistically similar (t-test P not significant) to actual results from MTT and SPWT. Results form both test demonstrated similar correlations with self-reported walking measures (Table 2).

Responsiveness

Eleven of the 50 participants did not return for retesting, with 6 not willing to return, 3 not showing for 2 scheduled retesting sessions, and 2 developing unrelated medical problems that disqualified them from undergoing final walking tests. Six of these 11 participants had undergone spine surgery. Of the 39 participants that were reassessed after treatment(s), 33% received physical therapy alone, 26% had spinal injections only, 31% had both physical therapy and spinal injection, and 10% underwent spine surgery. Only 32 participants underwent the final MTT, as 5 participants remained unable to undergo testing for safety reasons and 2 participants refused to undergo MTT during the final test session. All 39 participants underwent the final SPWT.

Comparisons of initial and post-treatment results from self-reported measures suggested that modest change in pain and function had occurred (Table 3), making assessment of responsiveness plausible. Change in global clinical status was also observed, as 28 participants reported that they had *improved* (6 symptoms resolved and 22 improved), and 11 participants reported they had *not improved* (9 unchanged and 2 worsened). Validation of this grouping was confirmed by comparing change scores for self-reported measures using independent sample t-test , and significant greater improvements in scores were noted between the *improved* and *not improved* groups for self-reported walking time, back pain, leg pain, ODI walking and SSQ physical function (all $P \leq .01$). Differences between *improved* and *not improved* groups approached but did not reach statistical significance for self-reported walking distance, ODI total scores and SSQ symptom scores. SSQ satisfaction scores were more favorable for the *improved* than the *not improved* participants ($P < .01$).

Table 4 presents internal responsiveness results. MTT walking time and distance did not significantly change which suggests that internal responsiveness was negligible under the treatment conditions of this study. In contrast, SPWT demonstrated modest internal responsiveness, as walking time and distance did significantly change following treatment. When SPWT results were analyzed for just the 32 participants that also completed MTT, the results were similar (data not presented). Walking speed lacked internal responsiveness for both test.

Evaluation of external responsiveness revealed that neither SPWT nor MTT demonstrated meaningful capabilities. Figure 1 presents ROC for SPWT and clearly demonstrates that changes in SPWT walking time (AUC = .545) or distance (AUC = .564) did not differentiate between *improved* and *not improved* participants. Figure 2 presents ROC for MTT, which performed slightly better than SPWT for walking time (AUC =.717) and distance (AUC =.702), though these results did not reach statistical significance.

Correlations between changes in MTT and SPWT walking times and distances with changes in self-reported measures are reported in Table 5. Changes in MTT results correlated with change in self-report walking time and distance, whereas SPWT did not. When compared to changes for other self-reported measures of walking, MTT walking time correlated with changes in ODI walking item. Changes in SPWT and MTT walking distance showed significant correlations only with SSQ physical function.

DISCUSSION

Validation of Walking Tests

This study supports the findings of others that limitations of walking caused by NC can be quantified using MTT [10,14-24] and SPWT [25-29]. Results from both MTT and

SPWT showed strong correlations with each other and with self-reported walking abilities, offering further validation that both walking test assess limitation that are characteristic of NC to a similar degree [24,30].

Both MTT and SPWT produced comparable results in terms of walking times, confirming the observation of Tomkins et al [29]. Walking speeds were slower for MTT than SPWT, and therefore the walking distance covered during similar walking times until onset of significant symptoms were shorter for MTT, a result also noted by Tomkins et al [29]. Of interest, similar observations by these two studies suggest that the duration of walking may be the most limited factor for production of NC, and slower pace of walking (as noted during treadmill walking) may afford little benefit in terms of total distance walked.

The reason for slower walking during MTT might related to caution while on this equipment that resulted from participants' unfamiliarity or unease with treadmill walking. If this is correct, MTT might systematically underestimate walking distances in NC similar to what was noted in studies of pulmonary diseases [41]. One must also consider that the shared walking experience between the test subject and the examiner during SPWT may have inadvertently resulted in pacing of the subject by the examiner. If true, pacing might have been an unanticipated, but important source of bias that may have elevated participants' walking speed and distances above the levels that would have occurred with unaccompanied walking.

It is of some concern that our MTT protocol, which disallowed holding the handrail with both hand, disqualified 10% of our participants from performing this test. Moon et al also reported that some participants with LSS cannot safely walk on a treadmill [21]. As these participants were older and generally more impaired, MTT may be less plausible as an outcome measure in clinical trials that focus on more frail individuals with NC. In contract, SPWT which mimicked real world walking situations could be performed by all study participants.

Even with the 30 minute protocols for SPWT and MMT and selection of participants with self-reported walking ability of 30 minutes or less, we found significant ceiling effect in this study. As one might expect, the influence of ceiling effect increased during the study with nearly one half of our participants completing the SPWT and one third completing MTT without symptoms at final testing. Though this improvement offers objective evidence of improved walking abilities, the completion of walking test protocols before symptoms were induced caps the measured walking times and distances below their true levels. This under measurement of walking abilities produced an underestimation of treatment response by these measures. This may be particularly problematic when using walking tests to assess outcomes in those with only modest limitations in initial walking (and therefore more likely to undergo conservative care), along with those who have the most dramatic improvements in walking abilities in response to treatment. This influence

of ceiling effect on measurement of walking ability would obviously be far greater when walking test time utilize protocols of less than 30 minutes [10].

A clear disadvantage of SPWT is its requirement for an appropriate space for the walking course. We were fortunate to have access to a rectangular public corridor in a medical office building that was similar to the walking courses used by others [27,28]. The availability of an appropriate space for a walking course may be a significant barrier to the adoption of SPWT as a widely used outcome measure for clinical trials, especially when compared to the space requirements for MTT, where only several square meters of space are needed. However, when an appropriate space is available, the SPWT is a much cheaper alternative to the MTT, as the cost of a motorized treadmill can be prohibitive.

Responsiveness of Walking Test

Overall, the study cohort demonstrated evidence of clinical improvement between initial and final completion of self-reported measures, making the assessment of responsiveness of SPWT and MTT plausible. Improvement was reflected by change scores that reached published improvement thresholds for LSS treatments for SSQ symptoms and physical function subscales, and ODI [35,36]. Back pain, leg pain and ODI walking item scores also improved, but clinically important threshold values for improvements of these measures are not widely agreed upon for LSS. Furthermore, the proposed criteria for use of global change in clinical status as a reference standard for assessment of external responsiveness was met, as global improvement was reported by a majority of

participants, and improved clinical status was confirmed by concurrent changes in most measures of pain and function.

Examination of responsiveness produced several important findings. MTT did not demonstrate adequate internal responsiveness for our group of participants undergoing a variety of treatment, as MTT time and distance did not significantly change between initial and final testing, and effect sizes as measured by MTT were insignificant. These findings are in contrast to the observations noted in several studies of lumbar spine surgery for LSS where improved treadmill walking time and/or distances were documented [18,19,43-45]. As internal responsiveness is dependent on the magnitude of change induced by treatment, it is possible that the changes in walking produced during this study were less substantial than those that result from cohort of patients treated with spine surgeries, though this would not be supported by results published by Malmivaara et al [10]. Addition, the ceiling effect may have limiting the ability of MTT to fully measure walking capacities in some participants, thus underestimating actual improvement in walking. This would be most true for subject that were only moderately limited by NC, and therefore more likely to reach the test's ceiling. MTT may best be able to detect changes in walking status in the most disabled individuals who are more likely to consider surgery [46].

Modest internal responsiveness was demonstrated for SPWT, as mean walking time, and distance did improve between initial and final testing. This occurred despite similar

ceiling effects, and offers evidence that SPWT may be responsive for detecting change in walking abilities following treatments of NC.

Results from evaluation of external responsiveness are the most noteworthy findings of this study, as changes in scores for neither MTT nor SPWT closely corresponded with other measures of change in clinical status. These results are surprising for this group of participants with NC, as it was our assumption that actual limitations in walking ability were the major concern for these patients, and this would be reflected in strong relationships between changes in walking abilities and other outcomes dimensions. However, at least for SPWT, where modest improvements were documented, these improvements in measured walking were not significantly different for participants rating themselves as *improved* or *not improved*. These findings suggest that on an individual level, there is a substantial divergence between objective and subjective outcome dimensions. These findings are similar to those noted for chronic low back pain treatments, where improvements in objective measurements of back flexibility, strength and lifting ability show weak (if any) relationships with changes in pain and function [47,48]. This does not suggest that changes in objective measures of function are irrelevant to the outcome of patients with spinal disorders, but instead that the magnitude of changes in objective measures has limited predictive value concerning subjective changes in clinical status.

CONCLUSION

Both MTT and SPWT can assessment the walking limitations that result from NC. In terms of outcomes measures for clinical trials, SPWT demonstrates greater internal responsiveness than MTT for the modest change in clinical status noted in this study. MTT demonstrated slightly better external responsiveness than SPWT, but external responsiveness was generally insufficient for both test, as changes in objective walking test show little concordance with patients' perceptions of change in clinical status. We conclude that objective measures of walking ability such as SPWT and MTT can be used as a distinct outcome measure following treatment of NC, but are not superior to self-reported walking abilities, not are they a substitute for disease specific measures of pain and function.

REFERENCES

1. Arnoldi CC, Brodsky AE, Cauchoix J, et al. Lumbar spinal stenosis and nerve root entrapment syndromes. Definition and classification. Clin Orthop Relat Res 1976;115:4-5.

2. Herkowitz HN. Spine update. Degenerative lumbar spondylolisthesis. Spine 1976;20:1084-90.

3. Amundsen T, Weber H, Lilleas F, Nordal HJ, Abdelnoor M, Magnaes B. Lumbar spinal stenosis. Clinical and radiologic features. Spine 1995;20:1178-86.

4. Kalichman L, Cole R, Kim DH, et al. Spinal stenosis prevalence and association with symptoms: the Framingham Study. Spine J 2009;9:545-50.

5. Katz JN, Harris MB. Clinical practice. Lumbar spinal stenosis. N Engl J Med 2008;358:818-25.

6. Goh KJ, Khalifa W, Anslow P, Cadoux-Hudson T, Donaghy M. The clinical syndrome associated with lumbar spinal stenosis. Eur Neurol 2004;52:242-9.

7. Winters CC, Brandes M, Müller C, et al. Walking ability during daily life in patients with osteoarthritis of the knee or the hip and lumbar spinal stenosis: a cross sectional study. BMC Musculoskelet Disord 2010;11:233.

8. Deyo RA, Mirza SK, Martin BI, Kreuter W, Goodman DC, Jarvik JG. Trends, major medical complications, and charges associated with surgery for lumbar spinal stenosis in older adults. JAMA 2010;303:1259-65.

9. Iversen MD, Katz JN. Examination findings and self-reported walking capacity in patients with lumbar spinal stenosis. Phys Ther 2001;81:1296-306.

10. Malmivaara A, Slätis P, Heliövaara M,et al.. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine 2007;32:1-8.

11. Fairbank JCT, Pynsent PB. The Oswestry Disability Index. Spine 2000;25:2940-53.

12. Pratt RK, Fairbank JC, Virra. The reliability of the Shuttle Walking Test, the Swiss Spinal Stenosis Questionnaire, the Oxford Spinal Stenosis Score, and the Oswestry Disability Index in the assessment of patients with lumbar spinal stenosis. Spine 2002;27(1):84-91.

13. Tomkins CC, Battie MC, Hu R. Construct validity of the physical function scale of the Swiss Spinal Stenosis Questionnaire for the measurement of walking capacity. Spine 2007;32:1897-901.

14. Dong GX, Porter RW. Walking and cycling tests in neurogenic and intermittent claudication. Spine 1989;14:965-9.

15. Herno A, Airaksinen O, Saari T. Computed tomography after laminectomy for lumbar spinal stenosis. Patients' pain patterns, walking capacity, and subjective disability had no correlation with computed tomography findings. Spine 1994;19:1975-8.

16. Deen HG, Ximmerman RS, Lyons MK, et al. Measurement of exercise tolerance on the treadmill in patients with symptomatic lumbar spinal stenosis: a useful indicator of functional status and surgical outcome. J Neurosurg 1995;83:27-30.

17. Fritz JM, Erhard RF, Delitto, A, Welch WC, Nowakowski PE. Preliminary results of the use of a two-stage treadmill test as a clinical diagnostic tool in the differential diagnosis of lumbar spinal stenosis. J Spinal Disord 1997;10:410-6.

18. Deen HG, Zimmerman RS, Lyons MK, Mcphee MC, Verheijda JL, Lemens SM. Use of the exercise treadmill to measure baseline functional status and surgical outcome with severe lumbar spinal stenosis. Spine 1998;23:244-8.

19. Herno A, Airaksinen O, Saari T, Pitkanen M, Manninen H, Suomalainen O. Computed tomography findings 4 years after surgical management of lumbar spinal stenosis. No correlation with clinical outcome. Spine 1999;24:2234-9.

20. Whitehurst M, Brown LE, Eidelson SG, D'Angelo A. Functional mobility performance in an elderly population with lumbar spinal stenosis. Arch Phys Med Rehabil 2001;82:464-7.

21. Moon ES, Kim HS, Park JO et al. Comparison of the predictive value of myelography, computer tomography and MRI on the treadmill test in lumbar spinal stenosis. Yonsei Med J 2005;46:806-11.

22. Whitman JM, Flynn TW, Childs JD, et al. A comparison between two physical therapy treatment programs for patients with lumbar spinal stenosis. A randomized clinical trail. Spine 2006;31:2541-9.

23. Barz T, Melloh M, Staub L, et al. The diagnostic value of a treadmill test in predicting lumbar spinal stenosis. Eur Spine J 2008;17:686-90.

24. Zeifang F, Schiltenwolf M, Abel R, Moradi B. Gait analysis does not correlate with clinical and MR imaging parameters in patients with symptomatic lumbar spinal stenosis. BMC Musculoskelet Disord 2008;9:89.

25. Podichetty VK, Segal AM, Lieber M, Mazanec DJ. Effectiveness of salmon calcitonin nasal spray in the treatment of lumbar canal stenosis. A double-blind, randomized, placebo-controlled, parallel group trial. Spine 2004;29:2343-9.

26. Fukusaki M, Kobayashi I, Tetsuya H, Sumikawa K. Symptoms of spinal stenosis do not improve after epidural steroid injection. Clin J Pain 1998;14:148-51.

27. Geisser ME, Haig AJ, Tong HC, et al. Spinal canal size and clinical symptoms among persons diagnosed with lumbar spinal stenosis. Clin J Pain 2007;23:780-5.

27. Tong HC, Haig AJ, Geisser ME, Yamakawa KSJ, Miner JA. Comparing pain severity and functional status of older adults without spinal symptoms, with lumbar spinal stenosis and with axial low back pain. Gerontology 2007;53:111-5.

29. Tomkins CC, Battie MC, Rogers T, Jiang H, Petersen S. A criterion measure of walking capacity in lumbar spinal stenosis and its comparison with a treadmill protocol. Spine 2009;34:2444-9.

30. Tomkins-Lane CC, Battié MC. Validity and reproducibility of self-report measures of walking capacity in lumbar spinal stenosis. Spine 2010;35:2097-102.

31. Huster JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-68.

32. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? Med Care 1996;34:73-84.

33. Lurie JD, Tosteson AN, Tosteson TD, et al. Reliability of readings of magnetic resonance imaging features of lumbar spinal stenosis. Spine 2008;33:1605-10.

34. Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. Spine 2005;30:1331-4.

35. Stucki G, DaltroyL, Liang MH, et al. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. Spine 1996;21:796-803.

36. Tuli SK, Yerby SA, Katz JN. Methodological approaches to developing criteria for improvement in lumbar spinal stenosis surgery. Spine 2006;31:1276-80.

37.Müller R, Büttner p. A critical discussion of intraclass correlation coefficients. Statistics in Medicine 1994;13:2465–76.

38. Cohen, Jacob (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20 (1): 37–46.

39. Beaton DF, Hogg-Johnson S, Bombardier C. Evaluating changes in health status measures: reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. J Clin Epidemiol 1997;50:79-93.

40. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change; an analog to diagnostic test performance. J Chron Dis 1986;39:397-906.

41. Swerts PM, Mostert R, Wouters EF. Comparison of corridor and treadmill walking in patients with severe chronic obstructive pulmonary disease. Phys Ther 1990;70:439-42.

42. Deen HG, Zimmerman RS, Lyons MK, et al. Test-retest reproducibility of the exercise treadmill examination in lumbar spinal stenosis. Mayo Clin Proc 2000;75:1002-7.

43. Tenhula J, Lenke LG, Bridwell KH, Guptal P, Riew D. Prospective functional evaluation of the surgical treatment of neurogenic claudication in patients with lumbar spinal stenosis. J Spinal Disord 2000;13:276-82.

44. Papavero L, Thiel M, Fritzsche E, Westphal M, Kothe R. Lumbar spinal stenosis: prognostic factors for bilateral microsurgical decompression using a unilateral approach. Neurosurgery 2009;65:182-7.

45. Yasar B, Simsek S, Er U, et al. Functional and clinical evaluation for the surgical treatment of degenerative stenosis of the lumbar spinal canal, J Neurosurg Spine 2009;11:347-52.

46. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical compared with non-operative treatment for lumbar degenerative spondylolisthesis. Four-year results in the Spine Patient Outcomes Research Trial (SPORT) randomized and observational cohorts. Joint Surg Am 2009;91:1295-304

47. Rainville J, Ahern DK, Phalen L, Childs LA, Sutherland R,. The association of pain with physical activities in chronic low back pain. Spine 1992;17:1060-4.

48. Kernan T, Rainville J. The Influence of Exercise on Kinesiophobia in Chronic Low Back Pain. J Orthop Sports Phys Ther 2007;37:679-87.

Table 1. Comparisons of walking results from self-paced walking test (SPWT) and

motorized treadmill test (MTT) using paired-sample *t*-test. (N = 45 participants that

completed both tests)

	MTT	SPWT	Difference	95% CI	Sig
Time (minutes)	14.2	13.7	0.4	-1.7 – 2.6	.68
Distance (m)	711	872	161	31 – 290	.02
Speed (km/h)	2.6	3.7	1.1	0.7 – 1.4	.01

95 % CI – 95 percent confidence interval for difference; m – meters; km/h – kilometers

per hour

Table 2. Pearson correlations between walking abilities as measured with self-paced walking test (SPWT) and motorized treadmill test (MTT) and self-reported walking measures. (Significance ≤ .01 for all correlations.)

Self-Report Measures	MTT		SPWT	
	Time	**Distance**	**Time**	**Distance**
Estimated Time	.73	.70	.56	.63
Estimated Distance	.66	.72	.58	.65
ODI Walking	-.63	-.54	-.47	-.49
SSQ Phys Function	-.63	-.45	-.58	-.55

SSQ Phys Function – Spinal Stenosis Questionnaire Physical Function Subscale; ODI walking – Oswestry Disability Index walking item

Table 3. Comparisons by paired sample t-test of initial to final results for self-report

measures. **(N=39)**

Self-Report Measure	Initial	Final	Diff	95% CI	Sig.
Walking Time (min)	15.9	27.0	11.1	4.2 – 18.1	.003
Walking Distance (m)	628	1837	1209	240 – 2179	.016
Back Pain	6.1	3.9	2.2	1.2 – 3.1	.001
Leg Pain	5.4	3.4	2.0	1.2 – 3.1	.001
SSQ symptoms	3.0	2.5	0.5	0.3 – 0.7	.001
SSQ Phy Func	2.4	1.9	0.6	0.4 – 0.8	.001
ODI Walking	3.6	2.3	1.3	0.8 – 1.7	.001
ODI Total	35	23	11.7	6.1 – 17.2	.001

Diff – mean difference; 95% CI – 95% confidence interval for mean differences; min –

minutes; m – meters; km/h – kilometers per hour; SSQ symptoms – Spinal Stenosis

Questionnaire Symptoms Subscale; SSQ Phy Func – Spinal Stenosis Questionnaire

Physical Function Subscale; ODI Walking – Oswestry Disability Index Walking Item;

ODI Total – Oswestry Disability Index Total Score

Table 4. Internal responsiveness analyses for motorized treadmill test (MMT) and self-paced walking test (SPWT) using paired sample *t*-test and effect size (ES).

	Initial (SD)	Final (SD)	Mean Diff	95% CI	ES
MTT (N=32)					
Time (min)	14.3	16.2	1.9	-.93 – 4.7	.17
Distance (m)	760	834	74	-90 - 238	.09
Speed (km/h)	2.71	2.83	0.15	-0.2 – 0.1	.11
SPWT (N=39)					
Time (min)	12.7	18.1	6.1	2.9 – 9.2*	.60
Distance (m)	794	1181	387	199 – 575*	.51
Speed (km/h)	3.48	3.56	0.08	-0.3 - 0.2	.07

* *t*-test significance <.01; min – minutes; m – meters; km/h – kilometers per hour

Table 5. Pearson correlations between the changes in scores of self-reported measures with changes in scores for motorized treadmill test (MTT) and self-paced walking test (SPWT).

	MTT		SPWT	
Self-Report	**Time**	**Distance**	**Time**	**Distance**
Est. Walking Time	.48**	.50**	.07	.09
Est. Walk Distance	.37*	.62**	.08	.11
SSQ Phys Funct	.35	.41*	.25	.36*
ODI Walking	.48**	.35	.17	.23

Significance *.05, **.01; Spinal Stenosis Questionnaire Symptoms Subscale; SSQ Phy Funct – Spinal Stenosis Questionnaire Physical Function Subscale; ODI walking – Oswestry Disability Index walking item score

Figure 1. Receiver operating characteristics (ROC) curves for correct classification of improved global clinical status by changes in walking time and distance as measured by self-paced walking test (SPWT).
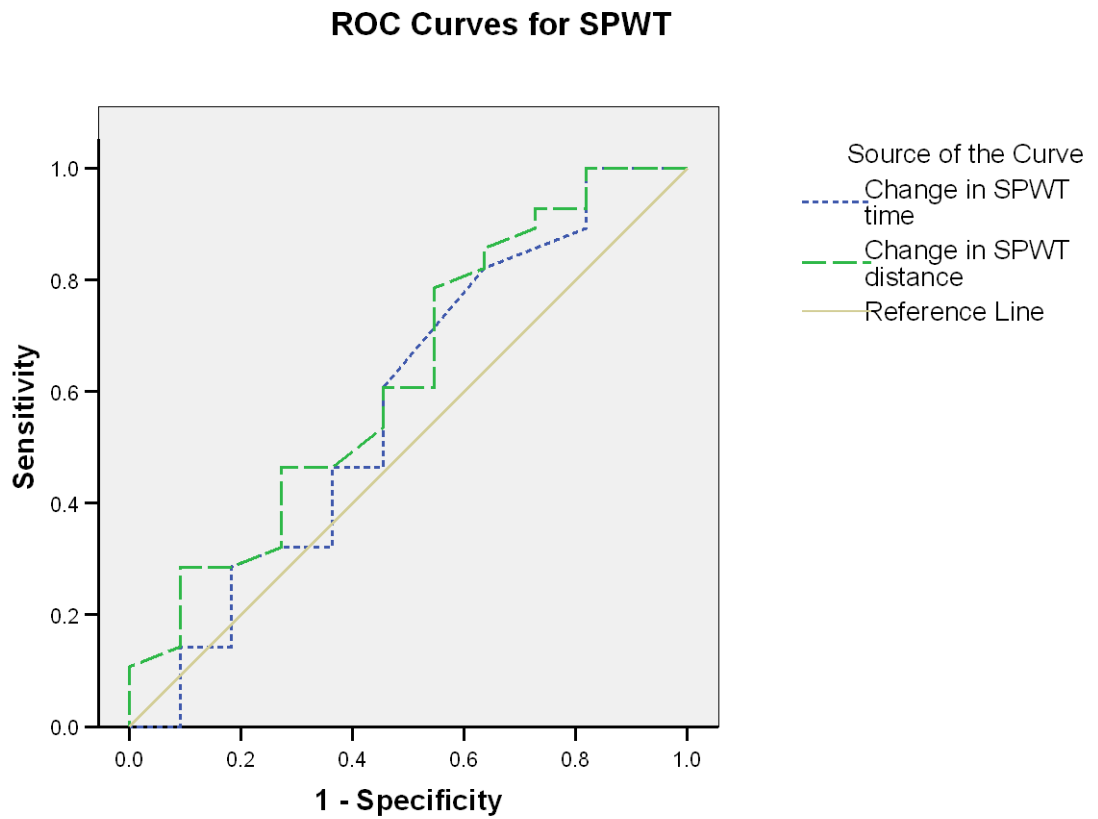
Figure 2. Receiver operating characteristics curves (ROC) for correct classification of improved global clinical status by changes in walking time and distance as measured by motorized treadmill test (MTT).

ROC Curves for MTT