3-14-2016

Manual Unloading of the Lumbar Spine: Can It Identify Immediate Responders to Mechanical Traction in a Low Back Pain Population? A Study of Reliability and Criterion Referenced Predictive Validity

John Carlos Jr.
Andrews University, carlosjj@andrews.edu

Brian T. Swanson
Texas Woman's Univrsity

Sean P. Riley
University of Connecticut Health Center

Mark P. Cote
University of Connecticut Health Center

Robin L. Leger
Salem State University

See next page for additional authors

Recommended Citation

Authors

John Carlos Jr., Brian T. Swanson, Sean P. Riley, Mark P. Cote, Robin L. Leger, and Isaac L. Moss

Manual unloading of the lumbar spine: can it identify immediate responders to mechanical traction in a low back pain population? A study of reliability and criterion referenced predictive validity

Brian T. Swanson[1], Sean P. Riley[2], Mark P. Cote[2], Robin R. Leger[3], Isaac L. Moss[2], John Carlos, Jr[4]

[1]Texas Woman's University, Houston, TX, USA, [2]University of Connecticut Health Center, Farmington, CT, USA, [3]Salem State University, MA, USA, [4]Andrews University, Berrien Springs, MI, USA

Background: To date, no research has examined the reliability or predictive validity of manual unloading tests of the lumbar spine to identify potential responders to lumbar mechanical traction.

Purpose: To determine: (1) the intra and inter-rater reliability of a manual unloading test of the lumbar spine and (2) the criterion referenced predictive validity for the manual unloading test.

Methods: Ten volunteers with low back pain (LBP) underwent a manual unloading test to establish reliability. In a separate procedure, 30 consecutive patients with LBP (age 50.86 ± 11.51) were assessed for pain in their most provocative standing position (visual analog scale (VAS) 49.53 ± 25.52 mm). Patients were assessed with a manual unloading test in their most provocative position followed by a single application of intermittent mechanical traction. Post traction, pain in the provocative position was reassessed and utilized as the outcome criterion.

Results: The test of unloading demonstrated substantial intra and inter-rater reliability $K=1.00$, $P=0.002$, $K=0.737$, $P=0.001$, respectively. There were statistically significant within group differences for pain response following traction for patients with a positive manual unloading test ($P<0.001$), while patients with a negative manual unloading test did not demonstrate a statistically significant change ($P>0.05$). There were significant between group differences for proportion of responders to traction based on manual unloading response ($P=0.031$), and manual unloading response demonstrated a moderate to strong relationship with traction response Phi$=0.443$, $P=0.015$.

Discussion and conclusion: The manual unloading test appears to be a reliable test and has a moderate to strong correlation with pain relief that exceeds minimal clinically important difference (MCID) following traction supporting the validity of this test.

Keywords: Low back pain, Traction, Unloading, Reliability, Validity

Introduction

Low back pain (LBP) is one of the most common musculoskeletal conditions treated by physical therapists.[1] Although lumbar traction is frequently used by physical therapists in the treatment of patients with LBP, there is limited evidence to support its use.[2] Reports have suggested that the use of lumbar traction is based primarily on anecdotal evidence,[3] with review articles[4] and clinical practice guidelines[2,5] questioning its effectiveness. Despite this absence of evidence, lumbar traction continues to be a frequently utilized modality.[6–8] Research identifying who will benefit from mechanical traction and its efficacy is therefore needed.

One challenge in evaluating the efficacy of treatments for LBP is the lack of clear diagnostic subgroups based on pathology or presentation. It has been reported that, in patients with LBP, a specific anatomic structure may be implicated in as few as 10% of cases.[9] Given the uncertainty of a pathoanatomic diagnosis, grouping patients on the basis of their mechanical behaviors may lead to a more accurate treatment model.

Correspondence to: Brian T. Swanson, Texas Woman's University, Houston, TX, USA. Email: bswanson@twu.edu; swanson996@aol.com

One possible means of identifying patients with LBP who may benefit from traction, via assessment of mechanical behavior, is through manual unloading tests. Global and segmental unloading tests are frequently applied as part of the provocation/alleviation scheme as described by Kaltenborn,[10] as well as Evjenth and Gloeck.[11] The manual unloading test is used as a means to assess load-sensitive structures of the lumbar spine and to determine if a patient may benefit from traction. A case study by Corkery[12] reported successful treatment of a patient with LBP utilizing standing traction, which was selected based upon the results of manual unloading tests of the lumbar spine. A recent study by Holtzman *et al.*[13] reported the use of various unloading techniques to alleviate symptoms in a cohort of patients with chronic LBP. Their study found significant decreases in pain during unloading techniques applied in positions which reproduced the patient's lumbar spine complaints.[13]

Although the manual unloading tests have been used clinically to determine when to use traction as a treatment, we are unaware of any studies that verify the reliability or validity of these tests. The purpose of this study was to assess the reliability and criterion referenced predictive validity of a manual unloading test of the lumbar spine. We expected that: (1) two clinicians could reliably identify a positive or negative response to manual unloading of the lumbar spine both within and between sessions; (2) there would be differences in response to traction (pain relief) between groups, based on the results of the manual unloading test; and (3) there would be a significant relationship between the results of manual unloading and response to mechanical traction (predictive validity).

Methods

The protocol had two parts: (1) reliability and (2) predictive validity. The protocol was reviewed and approved by the Institutional Review Boards of The University of Connecticut Health Center and Andrews University before initiation of the study. This study was registered through ClinicalTrials.gov, study #NCT02026076.

Participants

Two separate groups were recruited for this study. The first group was a sample of convenience, consisting of 10 individuals with LBP meeting the inclusion/exclusion criteria. This sample participated in the reliability portion of the study only. The second group consisted of a consecutive sample of patients with LBP (*n*=30). This sample was recruited to examine the predictive validity of manual unloading.

Inclusion criteria

Patients between the ages of 18 and 75 years, with complaints of non-radicular LBP were eligible for inclusion into the study. Non-radicular LBP was defined as pain in the lumbar area that did not extend below the knee.[14]

Exclusion criteria

Patients were excluded from this study if they presented with advanced pathology including tumor, fracture, infectious disorder, central nervous system involvement, presence of medical red flags, absence of LBP, radicular leg pain (below the knee), pregnancy, epidural steroid injection within 4 weeks before study involvement, previous back surgery, workers compensation involvement, or active litigation. These criteria are similar to those established by previous traction studies.[15,16]

Test procedure (manual unloading)

Patients were tested via manual unloading as described by Kaltenborn.[10] This test involves the therapist applying a low-grade lifting force to the patient in a standing position. A positive test consists of a decrease in the patient's symptoms. The test was performed with one of two variations, depending upon the lumbar position where the patient was symptomatic. For individuals with pain at rest, the patient stood with their arms crossed across their chest, and the therapist stood on the patient's least painful side. The therapist grasped the patient around the lower aspect of the ribcage, and gradually applied a low-grade, vertically oriented, lifting force (Fig. 1). The unloading force was gradually increased until the patient's upper body began to lift. Care was taken not to lift the patient from the ground. The therapist then asked the patient if there was any change in his/her symptoms by asking, 'Are your symptoms 'better', 'worse', or 'the same'? A response of 'better' was considered a positive (+) result, while a response of 'the same' or 'worse' was considered a negative (−) result.

For patients who had no pain at rest and presented with lumbar positional pain only, the test was modified as follows: if the pain was provoked by a side bending motion, the therapist stood on the side opposite of the painful direction of sidebending (Fig. 2). For flexion or extension pain, the therapist again stood at the patient's side of least pain. The therapist grasped around the ribcage as previously described. The patient was then asked to move into their pain provoking direction until pain was slightly reproduced. Once the patient had reproduced their pain, the therapist then applied the vertical unloading force, as previously described.

Before the initiation of the study, both therapists completed a training session with the assistance of a

Figure 1 Unloading in neutral standing; testing position, force applied in direction of arrow.

third colleague, who acted as a mock patient. One therapist (BTS) had previously been trained in the use of the technique in an Orthopaedic Manual Therapy (OMT) fellowship program. This therapist provided training to the second therapist (SPR) through direct instruction. The mock patient was tested by unloading, assessing both the resting and provocative positions (flexion, extension, side bending). Feedback was provided by the mock patient regarding the direction and amplitude of forces, to help ensure consistency of the technique. The training was considered complete when the colleague reported that the forces and directions were consistent between examiners.

Part 1: reliability

Ten participants, with LBP, who met the inclusion/exclusion criteria, were recruited for the reliability portion of this study. Participants were tested for response to unloading by the two examiners. Both examiners performed the test procedure on the day of enrollment for all 10 participants, and were blinded to the results of the other examiner. This procedure was repeated a second time by each examiner, at least 7 days and no more than 28 days after initial testing. The testing order was not standardized and was based on the availability of the participant and examiner. Response to unloading was recorded as positive (+) relief or negative (−) relief for both sessions.
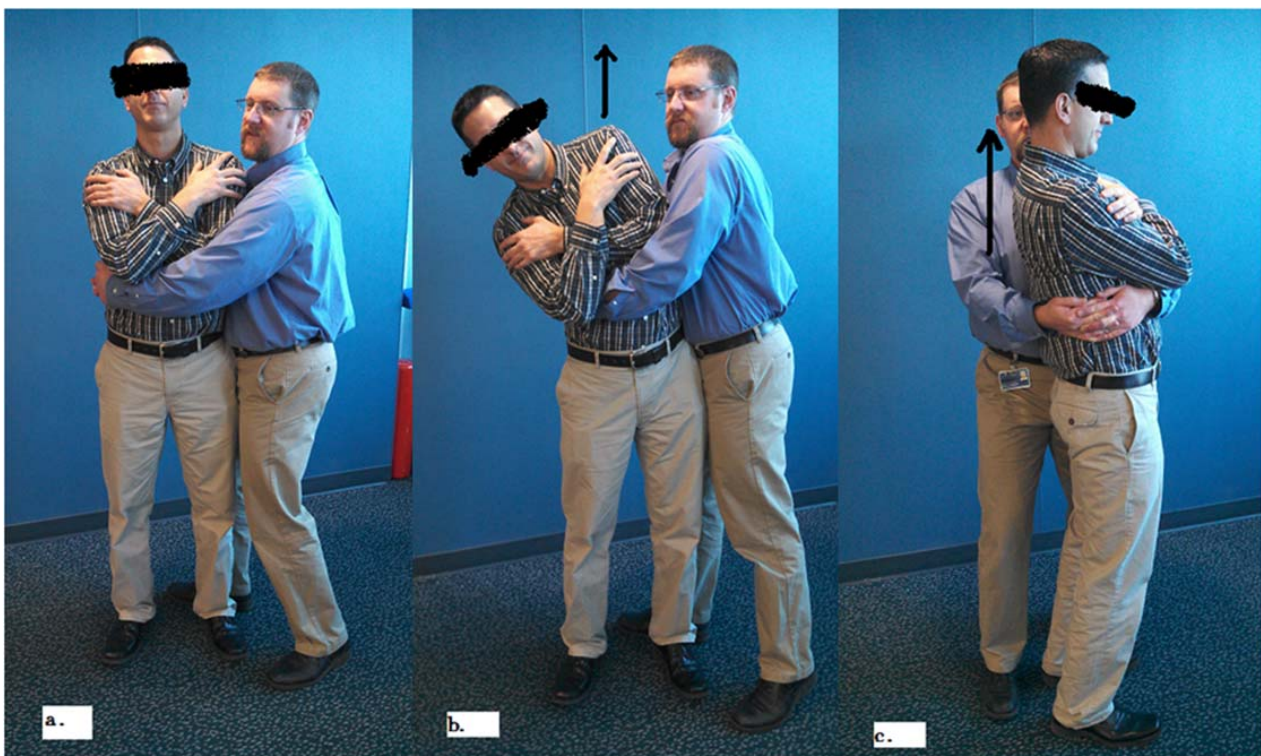

Figure 2 Unloading in sidebending/extended postures; (A) Starting position; (B) patient in sidebend, with test force applied; (C) patient in extension, with test force applied.

Part 2: predictive validity of manual unloading

A consecutive sample of 30 patients who met all study criteria participated in this portion of the study. All patients completed a screening questionnaire and underwent a standardized, comprehensive examination (standard of care), which included screening for exclusion criteria. Additionally, patients completed the modified Oswestry questionnaire and a 100 mm visual analog scale (VAS) to indicate pain in their most provocative test motion (flexion, extension, or side flexion of the lumbar spine). Following collection of these measures, one of the two examiners performed manual unloading as previously described. Response to manual unloading was recorded as either positive ($+$) or negative ($-$).

Before the application of mechanical traction, patients were weighed on a calibrated digital scale to determine traction force. Height was measured using a standard tape measure to allow body mass index (BMI) to be calculated. The reference test consisted of a single application of intermittent mechanical traction (Chattanooga, model TX-1), of 15 minutes duration, 30 seconds on/10 seconds off,[17] at up to 50% of body weight,[18] in a supine hook lying neutral posture with the belts orientated in the mid-position to provide a neutral pulling force. A split table in the open position was used to minimize the effect of friction. These positions were selected in an effort to minimize any confounding effects from sustained flexion or extension postures. At the completion of the traction session, the patients were allowed to rest on the traction table for up to 5 minutes before returning to a sitting then standing position. Post traction, all patients completed the VAS a second time, immediately following a retest of movement into the previously identified provocative position.

Reference criterion

The selected reference criterion was based upon patient response to mechanical traction (positive or negative). A positive response was defined as pain relief in the provocative test motion which met or exceeded the minimal clinically important difference (MCID) on the VAS. A negative result was defined as no change in pain or a change that did not meet MCID. Our operational definition of MCID was an improvement of at least 15 mm, or a change of at least 30% if the change was less than 15 mm. This definition was based on the recommendations of an expert panel formed at the VIII International Forum on Primary Care Research on Low Back Pain.[19]

Sample size

A power analysis was conducted to determine an adequate sample size for comparisons of response to mechanical traction with the result of the manual unloading test (positive/negative relief) as the grouping variable. As previously described, an improvement of at least 15 mm, or a change of at least 30%,[19] was considered the threshold for a clinically meaningful difference on the VAS. Standard deviations for the VAS have been reported to range from 5.7 to 22.[20] We calculated the average of these reported standard deviations (16.14) to serve as an estimate of variability. Using an alpha value set at 0.05 and beta set at 0.2 (power of 0.80), we estimated a total sample of 25 patients. To compensate for patients who chose to withdraw, as well as the uncertain distribution of the test response, the total sample size was inflated to 30.

Statistical analysis

Reliability

The kappa statistic was used to assess intra and inter-rater reliability for the manual unloading test, using the dichotomous variable of positive or negative relief.

Patient demographics

Descriptive statistics, including means and standard deviations, were used to characterize study participants where appropriate. The variables of BMI and chronicity were transformed into grouping variables. The BMI grade was categorized as underweight, normal, overweight, or obese.[21] Chronicity was defined as acute (0–42 days), subacute (43–84 days), or chronic (greater than 84 days).[22] Data were assessed for normality using the Shapiro–Wilk test and the Levene statistic for homogeneity of variance. Data were then assessed for between group differences at baseline as follows: A chi-square was used to assess age, gender, chronicity, extent of pain, and BMI grade. An independent sample *t*-test was used to assess weight, Oswestry score, and pain in the provocative position. A Mann–Whitney *U* test was used to assess BMI since this variable did not meet parametric assumptions of normality.

Criterion referenced predictive validity

To determine validity, it was first determined if there was a difference in response to the reference criterion (traction) between patients differentiated by results of the manual unloading test. A Wilcoxon signed-rank test was used to determine if a within groups statistical difference existed for those with a positive and a negative unloading response for change in pain pre–post traction. This non-parametric test was used due to the unbalanced group sizes.

Pain scores were then collapsed to form the dichotomous groups of responder/non-responder to traction in order to determine if there were clinically significant differences between the groups based on the result of the manual unloading test. This dichotomization was based on patient meeting/not meeting the MCID for pain response as previously
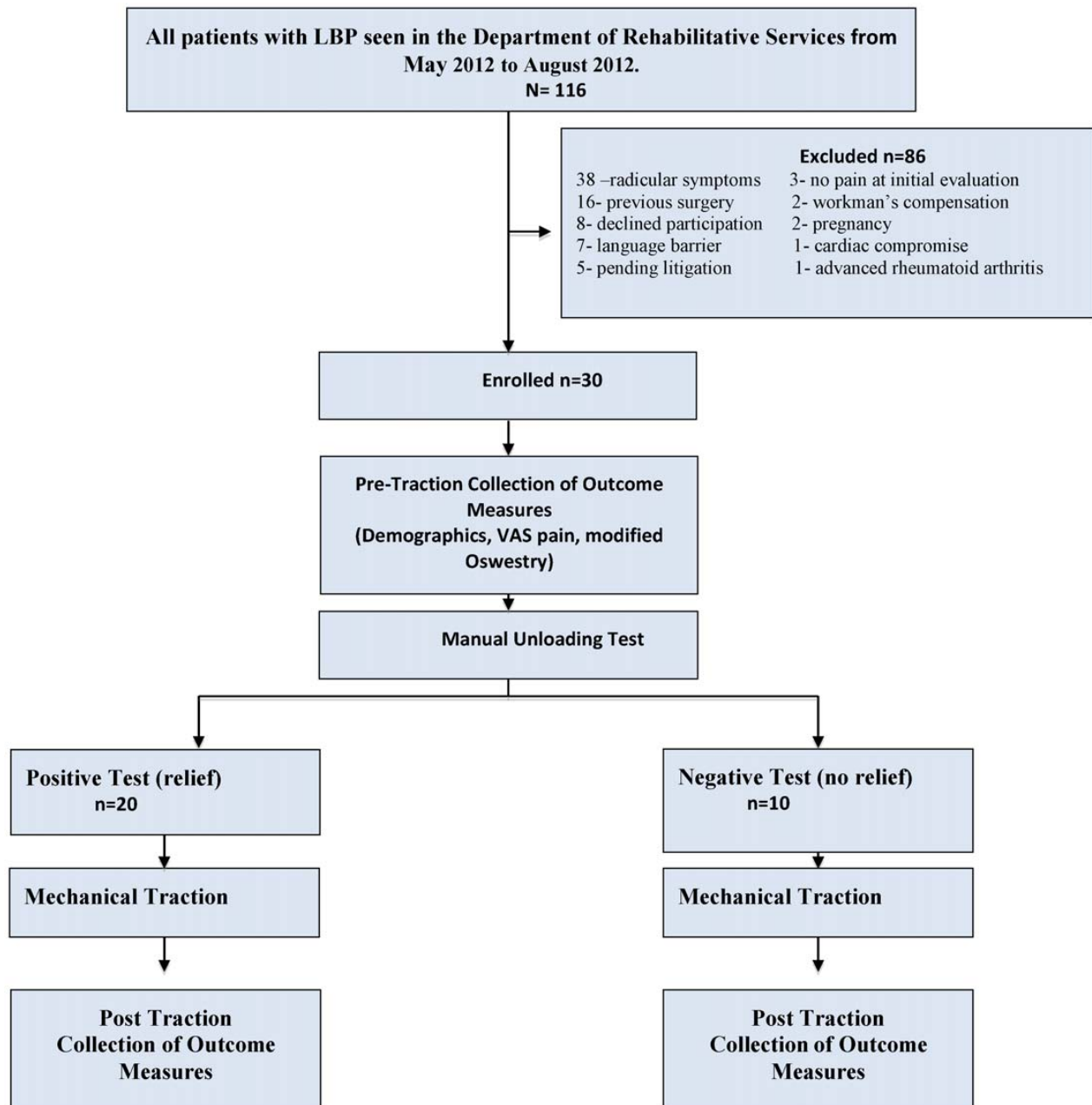
All patients with LBP seen in the Department of Rehabilitative Services from May 2012 to August 2012.
N= 116

Excluded n=86
38 –radicular symptoms   3- no pain at initial evaluation
16- previous surgery   2- workman's compensation
8- declined participation   2- pregnancy
7- language barrier   1- cardiac compromise
5- pending litigation   1- advanced rheumatoid arthritis

Enrolled n=30

Pre-Traction Collection of Outcome Measures
(Demographics, VAS pain, modified Oswestry)

Manual Unloading Test

Positive Test (relief)
n=20

Negative Test (no relief)
n=10

Mechanical Traction

Mechanical Traction

Post Traction Collection of Outcome Measures

Post Traction Collection of Outcome Measures

Figure 3   Patient flow.

described. Fisher's exact test was utilized to determine the significance of the differences in proportion for traction response based upon the dichotomized manual unloading result.

As a measure of predictive validity, the Phi coefficient was calculated to examine the strength of the relationship between the results of the manual unloading test and the results of the mechanical traction intervention. Strength of association was interpreted as follows: 0–0.1=weak association, 0.1–0.2=weak to moderate association, 0.25–0.35=moderate association, 0.4=moderate to strong, and >0.5=strong association.[23] All inferential statistical analyses were performed using PASW 18.0 (SPSS Inc. Released 2009. PASW Statistics for Windows, Version 18.0. Chicago, IL, USA: SPSS Inc.).

Results
Patient demographics
Between 17 May 2012 and 10 August 2012, a total of 116 patients were screened for potential inclusion. A total of 30 patients (mean $\pm$ SD age, $50.86 \pm$ 11.51 years, range 26–71 years; 63% female) met the eligibility criteria, were enrolled, and underwent a single session of traction at up to 50% body weight (46.21 ± 3.95). Figure 3 depicts the flow of study participants and reasons for exclusion. Of the 30 patients, 20 had a positive (relief) test result with manual unloading and 10 had a negative (no relief) test result. There were no statistically significant differences between groups at baseline for BMI, BMI grade, duration of pain, extent of pain, gender, age, Oswestry score, or pain in the provocative position.

Means and standard deviations of demographic and outcome variables at baseline for both groups are reported in Table 1.

Part 1: reliability

For intra-rater reliability, agreement within examiners was excellent at 100% agreement for the raw data. This corresponded to a kappa value=1.00 (P=0.002) indicating perfect agreement for each examiner. The inter-rater reliability presented with substantial agreement,[24] K=0.737, P=0.001. For our sample, all inter-rater disagreement occurred while testing the second enrolled participant. We used this individual's response to clarify our test procedure, as the participant spontaneously reported a difference in technique (the lifting vector) to both testers. Following correction of this difference, both testers elicited a positive test result. While we considered this difference to be an issue of technique application rather than a difference in the participant's condition, we left the discrepant values uncorrected in our calculation of inter-rater reliability. Data were blinded on all other reliability participants.

Part 2: predictive validity

Statistical between group differences

A Wilcoxon signed-rank test revealed a statistically significant change in pain, pre–post intervention, in the provocative position for the positive manual unloading group (P<0.001), mean 24.70 ± 16.55 mm. The negative manual unloading group did not demonstrate a statistically significant change in pain in the provocative position (P=0.059), mean 13.00 ± 19.34 mm.

Clinical between group differences

There was a significant difference in the proportion of responders whose pain relief crossed the MCID for the VAS following mechanical traction in the positive manual unloading group compared to the negative manual unloading group (Fisher's exact test, P=0.031), supporting the clinical significance of the manual unloading test.

Predictive value

The Phi coefficient demonstrated a statistically significant relationship between response to manual unloading and response to mechanical traction. The strength of this relationship was determined to be moderate to strong, Phi=0.443, P=0.015.

Those patients with a positive manual unloading test and a positive result to mechanical traction (true positive, n=19, −25.53 mm, 54.3%), as well as those with a negative manual unloading test and a positive response to mechanical traction (false negative, n=6, −26.33 mm, 48.9%) had large improvements in pain. One patient with a positive unloading result demonstrated small improvements in pain not reaching our definition of success (false positive, n=1, −9.0 mm, 22.5%). Patients with a negative unloading response and a negative response to mechanical traction demonstrated increased levels of pain (true negative, n=4, +7.22 mm, +26.1%). The relationship between manual unloading and response to mechanical traction is graphically displayed in Fig. 4.

Discussion

The goal of this study was to evaluate the reliability and predictive validity of the manual unloading test for response to mechanical traction in patients with LBP. The manual unloading test demonstrated acceptable levels of both intra and inter-examiner reliability. The manual unloading test is designed to discriminate between patients who will and will not

Table 1 Patient demographics

	Positive unloading test (n=20)	Negative unloading test (n=10)	P value
Demographics			
Age (years)	52 ± 12	49 ± 10	0.553
Weight (lbs)	199.70 ± 63.90	180.50 ± 40.30	0.819
Gender	11 Females, 9 males	8 Females, 2 males	0.180
BMI	30.13 ± 8.33	29.88 ± 7.52	0.846
BMI grade			
Normal	6 (30%)	3 (30%)	
Overweight	5 (25%)	4 (40%)	0.646
Obese	9 (45%)	3 (30%)	
Duration of symptoms			
Acute	7 (35%)	1 (10%)	
Subacute	2 (10%)	0 (0%)	0.151
Chronic	11 (55%)	9 (90%)	
Extent of pain			
LBP	12 (60%)	7 (70%)	
Buttocks	5 (25%)	2 (20%)	0.860
Thighs	3 (15%)	1 (10%)	
Outcome measures			
MODI	36.20 ± 18.63	30.60 ± 15.75	0.422
VAS pain (mm)	53.30 ± 25.91	42.00 ± 24.22	0.260

Data presented as means and standard deviations or portions where appropriate. BMI: body mass index, MODI: modified Oswestry disability index, VAS pain: visual analog pain scale in the provocative position, LBP: low back pain.
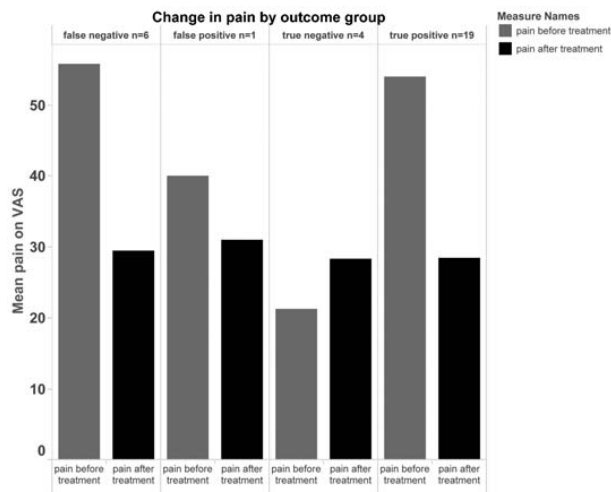
Figure 4 Change in pain by outcome group; true positive positive unloading result, pain relief exceeds minimal clinically important difference (MCID); false positive positive unloading result, pain relief less than MCID; true negative negative unloading result, pain relief less than MCID; false negative negative unloading result, pain relief exceeds MCID.

benefit from traction as an intervention. Significant statistical and clinical differences were observed for response to mechanical traction between those with a positive manual unloading test response and those with a negative manual unloading test response, supporting the discriminative ability of the manual unloading test and criterion referenced validity. A moderate-to-strong correlation was demonstrated between response to manual unloading and response to mechanical traction, demonstrating predictive validity.

Reliability

The manual unloading test was shown to demonstrate high levels of intra and inter-rater reliability. Additionally, the condition being tested presented with considerable stability, as all participants in the reliability sample reported the same results at follow-up as at initial testing. Response to unloading did not change in our population over a period of 7–28 days. While the long test–retest interval may have presented a potential source of error, we feel that these results show that the mechanical behavior of load-sensitive back pain may be consistent over time, and support the use of the manual unloading test.

Predictive validity

Interpretation of the results of this study should be made in consideration of the methodology. In designing the protocol, attention was paid to methods that would improve clinical applicability. Specifically, the physical exam techniques and methods of statistical analysis used in this study were selected based upon their potential to improve clinical relevance. It has been suggested that to accurately assess the outcome of a trial, statistical and clinical

differences should each be assessed individually,[25] as statistical significance does not always indicate clinical significance. To determine validity of the test, we first assessed whether the criterion referenced results were statistically different between the groups based upon manual unloading results. Despite the presence of clear statistical differences, it is our belief that, for the test to be valid, clinical differences must exist between groups for the manual unloading test to provide the necessary discriminative ability required for clinical decision making. Therefore, we then investigated these differences for clinical significance (relief greater than MCID) via a responder analysis. The strength of the relationship between the test result and criterion result was finally assessed to establish predictive validity.

In assessing the performance of the manual unloading test, all individuals who tested positive with the manual unloading test experienced pain relief following traction. There was only one patient with a positive unloading test result who reported a change in pain (9 mm improvement) that failed to cross the threshold for the MCID. There was a consistent positive response to traction in those with a positive response to manual unloading, with the positive manual unloading group demonstrating improvements of greater than 50%. Based on this result, we suggest that traction is an appropriate treatment option in the presence of a positive manual unloading test in a LBP population. Conversely, the negative response to manual unloading yielded an unpredictable response to traction, with increased pain following mechanical traction observed in 4/10 patients (mean = +7.22 mm). Therefore, we would not recommend the use of mechanical traction in the presence of a negative manual unloading test.

Use of the manual unloading test for the lumbar spine does require consideration based upon the size of both the treating therapist and the patient. There are instances where performance of this test is not practical based upon the physical characteristics of either party. In this study, the patient's weight ranged from 122 to 337 lb and heights ranged from 60–72 inches tall. Both larger and smaller patients required subtle modifications in form to accurately unload the lumbar spine and avoid unwanted force vectors. For shorter patients with taller therapists, it is required to squat significantly to maintain the appropriate forces. Conversely, with shorter therapists and tall patients, it may be necessary to stand on a stool to apply the unloading force. Additionally, in instances where the therapist is unable to fully grasp around the patient due to large girth, it may be required to stand behind the patient and apply the forces either at the lower ribcage or through the patient's elbows. Finally, while we did not observe

differences in our population based on direction of movement, it is possible that the patient's pain does not occur until they have reached the end range of motion. While this typically presents little difficulty for either sidebending or extension, this can present a problem with flexion as the unloading force is not easily applied in this position. An alternative mode of testing may need to be utilized in this instance, for example unloading in a seated posture, as long as symptoms can be reproduced.

In the treatment of LBP, identifying a specific structure at fault may not be possible.[9] We suggest that treatment of mechanical dysfunctions with appropriately selected interventions based on provocation and alleviation testing may be of benefit for patients suffering from LBP. There is most likely a cohort of patients who present with sensitivity to mechanical loading as a primary factor and the manual test of unloading may help to match this group to an appropriate treatment. However, while the manual unloading test predicts pain relief, future study is needed to determine whether it is useful in identifying a subgroup of patients who also achieve functional improvement.

Limitations

There are several limitations of this study. Our reliability sample ($n=10$) may allow for potential recall bias, as it is possible the examiners may have remembered the previous results at the retest interval. Some groups are under-represented, specifically age under 30 years ($n=1$) and subacute LBP ($n=2$) which may limit the ability to generalize these results. A single session design, while appropriate for the intended purpose of establishing reliability and predictive validity, does not allow for follow-up to assess longer-term functional outcomes. The prediction of immediate relief with traction does not indicate that traction is the most effective treatment for these subjects. Further research comparing commonly used treatment approaches among patients with a positive unloading test is needed to identify which treatments are the most effective for this population.

Conclusions

The manual unloading test of the lumbar spine appears to be a reliable measure. A positive result with the manual unloading test was found to be moderately to strongly correlated to the immediate response following a single session of mechanical traction. The use of manual unloading tests should be considered clinically as a tool to determine the appropriateness of mechanical traction as a symptom alleviation tool for patients with LBP. Additionally, the manual unloading test may be a valuable component of future research regarding the efficacy of lumbar traction and may assist in proper

sub-grouping of subjects. Future research is needed to expand upon these findings.

Disclaimer Statements

References

1 Jette AM, Smith K, Haley SM, Davis KD. Physical therapy episodes of care for patients with low back pain. Phys Ther. 1994;74:101–10.

2 Delitto A, George SZ, Van Dillen L, Whitman JM, Sowa G, Shekelle P, *et al*. Low back pain: clinical practice guidelines linked to the international classification of functioning, disability, and health from the orthopaedic section of the American Physical Therapy Association. J Orthop Sports Phys Ther. 2012;42:A1–57.

3 Krause M, Refshauge K, Dessen M, Boland R. Lumbar spine traction: evaluation of effects and recommended application for treatment. Man Ther. 2000;5:72–81.

4 Clarke J, van Tulder M, Blomberg S, de Vet H, van der Heijden G, Bronfort G. Traction for low back pain with or without sciatica: an updated systematic review within the framework of the Cochrane collaboration. Spine. 2006;31:1591–9.

5 Chou R, Huffman L. Nonpharmacologic therapies for acute and chronic low back pain: a review of the evidence for an American Pain Society/American College of Physicians Clinical Practice Guideline. Ann Intern Med. 2007;147:492–504.

6 Harte AA, Gracey JH, Baxter GD. Current use of lumbar traction in the management of low back pain: results of a survey of physiotherapists in the United Kingdom. Arch Phys Med Rehabil. 2005;86:1164–9.

7 Jette AM, Delitto A. Physical therapy treatment choices for musculoskeletal impairments. Phys Ther. 1997;77:145–54.

8 Li LC, Bombardier C. Physical therapy management of low back pain. An explanatory survey of therapist approaches. Phys Ther. 2001;81:1018–27.

9 Abenheim L, Rossignol AM, Gobeille D, Bonvalot Y. The prognostic consequences in the making of the initial medical diagnosis of work-related back injuries. Spine. 1995;20:791–5.

10 Kaltenborn F. Manual mobilization of the joints: the Kaltenborn method of joint examination and treatment, 4th edn. Oslo, Norway: Norli; 2003.

11 Evjenth O, Gloeck C. Symptom localization in the spine and the extremity joints, Minneapolis, MN: OPTP; 2002.

12 Corkery M. The use of lumbar harness traction to treat a patient with lumbar radicular pain: a case report. J Man Manip Ther. 2001;9:191–7.

13 Holtzman GW, Harris-Hayes M, Van Dillen LR. Clinical exam procedures for unloading the lumbar region: effects in people with chronic low back pain. J Orthop Sports Phys Ther. 2009;39:A32.

14 Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Hanscom B, Skinner J, *et al*. Surgical vs nonoperative treatment for lumbar disk herniation: the spine patient outcomes research trial (SPORT): a randomized trial. JAMA. 2006;296:2441–50.

15 Beattie P, Nelson R, Michener L, Cammarata J, Donley J. Outcomes after a prone lumbar traction protocol for patients with activity limiting low back pain: a prospective case series study. Arch Phys Med Rehabil. 2008;89:269–74.

16 Fritz J, Thackery A, Childs J, Brennan G. A randomized clinical trial of the effectiveness of mechanical lumbar traction

for sub-groups of patients with low back pain: study methods and rationale. BMC Musculoskelet Disord. 2010;11:81.

17 Cai C, Pua YH, Lim KC. A clinical prediction rule for classifying patients with low back pain who demonstrate short term improvement with mechanical traction. Eur Spine J. 2009;18:554–61.

18 Beurskens AJ, de Vet HC, Köke AJ, Regtop W, van der Heijden GJ, Lindeman E, *et al*. Efficacy of traction for nonspecific low back pain: 12-week and 6-month results of a randomized clinical trial. Spine. 1997;22:2756–62.

19 Ostelo RW, Deyo RA, Stratford PW, Waddell G, Croft P, Von Korff M, *et al*. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine. 2008;33:90–4.

20 Mantha S, Thisted R, Foss J, Ellis JE, Roizen MF. A proposal to use confidence intervals for visual analog scale data for pain measurement to determine clinical significance. Anesth Analg. 1993;77:1041–7.

21 Health National Institute of. Available from: http://www.nhlbisupport.com/bmi/

22 Goertz M, Thorson D, Bonsell J, Bonte B, Campbell R, Haake B, *et al*. Institute for Clinical Systems Improvement. Adult Acute and Subacute Low Back Pain. Updated November 2012.

23 Rubin A. Statistics for evidence-based practice and evaluation, 3rd edn. Belmont, CA: Brooks/Cole-Cengage Learning; 2012.

24 Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med. 2005;37:360–3.

25 Snapinn S, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. Trials. 2007;8: 31.