

Association between clinician-assessed lifting ability and workplace tolerance and patient self-reported pain and disability following work conditioning

Dennis L. Hart^{a,*}, Marian Kirk^b, Julie Howar^b and Sheila Mongeon^b

^aPresident, Hart Ergonomic Consulting, 551 Yopps Cove Road, White Stone, VA, 22578, USA

^bREHAB AT WORK, Rockville, MD, USA

Received 28 September 2005

Accepted 10 November 2005

Abstract. *Objective:* We investigated the association between clinician-assessed performance-based measures of improvement in lifting ability and workplace tolerance and patient self-reported improvement in pain and perceived disability following work conditioning (WC).

Methods: A sample of 76 patients (42 ± 9 yrs, 21 to 60 yrs, 74% male) was selected from a retrospective database because they had lumbar spine impairments and received treatment in a WC program. Patients completed self-report surveys for perceived disability (Oswestry), pain intensity (visual analog pain scale – VAS), and pain concerns (McGill short form) before initial functional capacity evaluation (FCE) and after termination of the WC program. During the FCE and during the WC, therapists assessed patient workplace tolerance (WT) and ability to lift from floor to waist (PDL).

Results: Over the WC program that averaged 19 (6 SD) days, Oswestry and VAS scores improved ($P < 0.05$), but the McGill scores did not ($P = 0.334$). 72% of patients improved their PDL, and 64% met their WT goals. None of the associations between patient self-report scores and performance-based measures were significant ($P > 0.05$).

Conclusions: In spite of continuing pain complaints, patients decreased their perceived disability and pain intensity, increased their lifting ability, and improved their workplace tolerance while participating in a work-conditioning program.

Keywords: Functional capacity evaluation, work conditioning, perceived disability, pain

1. Introduction

Workers who sustain work-related complaints that do not subside commonly are assessed with functional capacity evaluations (FCE) [22] and treated in work conditioning programs (WC) [6] for the specific purpose of improving the worker's functional abilities, so

they can safely return to work. For the purpose of this paper, we operationally define an FCE as a test designed to determine the level of function of a client with the purpose of providing an objective measure of a client's safe functional abilities compared to the physical demands of their work [2,17]. We operationally define a WC program as a work-related, intensive, goal-oriented treatment program specifically designed to restore a client's systemic, neuromusculoskeletal (strength, endurance, movement, flexibility, and motor control), and cardiopulmonary functions with the purpose of restoring the client's physical capacity and function, so the

*Address for correspondence: Dennis L. Hart, PhD, PT, Hart Ergonomic Consulting, 551 Yopps Cove Road, White Stone, VA 22578, USA. Tel.: +1 804 436 9727; Fax: +1 804 436 9328; E-mail: dsailhart@verizon.net.

client can safely return to work [1].

Because both the FCE and WC are designed to return the client to safe work, clinicians should collect and use objective performance-based measures with demonstrated predictive validity during the FCE and WC when making clinical decisions concerning return to work for the benefit of all the stakeholders involved in the case [13,14,16]. Fortunately, an increasing number of recent studies have shed some light on the reliability [4,25,30,35,36,40,41,43,47] and validity [13,14,28,32,34,40,42,44,49] of specific portions of the FCE. Because of mixed results of the psychometric testing and concerns for whether FCEs should be considered behavioral tests influenced by multiple factors including physical ability, beliefs, and perceptions [11,12], caution has been urged for interpretation of the measures collected during FCEs, and more future work is recommended [10,19–21,29]. Unfortunately, only limited data were found describing the psychometrics of measures collected during WC programs designed for clinical decision-making [31]. The measures commonly collected during WC programs include clinician assessments of how the client is functioning, such as the amount of weight the client can safely lift and the client's workplace tolerance. We operationally define workplace tolerance as the client's ability to successfully sustain work tasks, so that the client can perform those tasks during the performance of routine reimbursable employment.

Another type of measure common in health services research but uncommon in research concerning FCEs and WC programs is patient self-report of functional status. Patient self-report surveys present functional questions to patients who describe how they perceive their functional abilities by answering the questions. Responses are transformed into measures of functional status. For example, self-report of perceived disability, the converse of functional status, in patients with lumbar spine impairments has been shown to be reliable, valid and responsive [9] using the Oswestry Low Back Pain Disability Questionnaire originally described by Fairbanks [8]. Other areas of patient self-report are pain intensity [26,27] and concerns related to pain complaints [37,38], both of which have been shown to have adequate reliability and validity. Although the relation between pain intensity, pain complaints, and functional status are complex and not well understood [50], measures of each are readily available to the clinician from which the clinician can use to assist in their clinical decision-making.

Quantification of both clinician assessed performance-based functional assessments and patient self-

reported measures can be helpful in studies designed to investigate associations between measures for the purpose of developing concurrent validity of measures used to quantify the effectiveness of clinical programs [15,44]. Moderate correlations ($r = 0.46$ and $r = 0.52$) have been reported between the amount of weight lifted during the FCE and the Oswestry score [15,44], and both research teams suggested using both performance-based FCE measures and self-report measures of disability in order to obtain a comprehensive picture of the functional abilities of the client during the FCE because the results suggested that the measures assess different constructs of functioning. However, we found no papers describing the association of performance-based measures of functional ability and patient self-report of functional status during a WC program.

1.1. Purpose

The purpose of this investigation was to assess the association between clinician reported performance-based improvement of lifting ability and workplace tolerance and patient self-reported improvement in pain and perceived disability during a work-conditioning program.

2. Method

2.1. Design

We performed a secondary analysis of a proprietary outcomes database representing a sample of convenience of consecutive patients with lumbar spine impairments referred to a work-conditioning program. The Focus On Therapeutic Outcomes, Inc. Institutional Review Board for the Protection of Human Subjects approved this project. Because the project retrospectively analyzed data, no patient consent forms were used.

2.2. Subjects

Seventy-six adult patients (42 ± 9 years, 21 to 60 years, 74% male), which represent a sample of convenience, were selected from the database because they had lumbar spine impairments and received treatment in the work-conditioning program (Table 1). They represent the entire data set available for analyses. The patients were treated in five clinics by eight clinicians who were employed by a private occupational therapy

Table 1
Patient Characteristics ($n = 76$)

Characteristic	Value*
Age (years)	42 ± 9, 21, 60
Gender (% male)	74%
Multiple body parts affected (%)	22%
Clinical diagnosis (%)	
Lumbar disc pathology	46%
Lumbar sprain/strain	38%
Post lumbar spine fusion surgery	6%
Post lumbar spine fracture	8%
Lumbar contusion	2%
Time from date of incident to date of functional capacity evaluation (months)	18 ± 19, 2, 109
Time from date of functional capacity evaluation to entry into work conditioning program (days)	17 ± 56, 1, 396
Number of days in work conditioning program (days)	19 ± 6, 7, 30

*(Mean ± standard deviation, minimum, maximum).

practice (REHAB AT WORK, Inc., Rockville, MD, USA). All patients were receiving workers' compensation benefits. Patients were referred by physicians, insurance case managers, insurance adjusters, and attorneys for the purpose of reducing the patient's musculoskeletal impairments and improving their functional abilities and feasibility for employment.

2.3. Patient evaluation and treatment

Once a patient was referred for work conditioning, each patient was scheduled for a four-hour one-day functional capacity evaluation (FCE), which was performed prior to entry into the work-conditioning program (WC). The FCE was standardized and was based in part on the philosophies described by Isernhagen [23,24]. The FCE was designed to follow the standards described by Hart et al. [17] and the guidelines described by the American Physical Therapy Association [2], which meant, in part, that the FCE contained the following components: patient history; patient physical impairment examination; worker behavior evaluation; patient's perception of their disability evaluation; functional limitation evaluation; comparison of patient physical abilities and physical demands of the targeted job; a problem list; and recommendations. For the purpose of this study, we report only the patient's perception of their disability and pain, and the clinician's assessment of the patient's functional limitations assessed using performance-based functional tests.

If the therapist determined from the FCE that the patient would benefit from a WC program and the referral source agreed, the patient progressed to the WC

program. The WC program was standardized and designed to follow the standards described by Hart et al. [17] and guidelines described by the American Physical Therapy Association [1], which meant, in part, that the WC program covered the following content: goals were established to directly relate to job skills and job requirements, if appropriate; clinical techniques were designed to improve strength, endurance, movement flexibility, motor control, and cardiopulmonary capacity related to the performance of work tasks; practice, modifications, and instructions in work-related activities; education related to safe job performance and injury prevention; promotion of patient responsibility and self management; and provision in multi-hour sessions up to 5 days a week for a duration of up to 8 weeks. Patients commonly progressed through the WC program by attending 4 hours a day 5 days a week and progressing to 6 and then 8 hours a day over the course of 4 to 6 weeks total episode duration. Although the FCE and WC procedures were standardized, no attempt was made to determine how closely all clinicians adhered to the procedures.

Patients were retested functionally every week in the WC program to determine their progress towards their WC goals. When the therapist determined that the patient had met his or her WC program goals by assessing the weekly functional testing results and the referral source agreed, the patient was discharged from the WC program. The patient's functional abilities described at discharge were based on the results of the functional testing performed at the end of their last week in the program. The weekly and discharge functional testing during the WC program consisted of an abbreviated FCE, which followed the same guidelines as described above for the entry FCE. The abbreviated functional testing emphasized measures and activities related to the WC goals previously established following the FCE that were directly related to the targeted job skills and requirements and any other task or physical impairment that was identified as appropriate throughout the initial FCE, during WC, or during the weekly functional evaluations while in the WC. As stated above, for the purpose of this study, we report only changes in the patient's perception of their disability and pain, and the clinician's assessment of the patient's performance-based functional abilities.

2.4. Outcomes instruments and measures

There were two primary means of data collection: patient self-report and clinician observation during

functional testing. To assess change in the patient's perception of their disability and pain, patients completed three self-report questionnaires: first during the intake FCE, which was prior to FCE functional testing and initiation of the WC program, and second following the discharge functional testing. The questionnaires included the Oswestry Low Back Pain Questionnaire [8], the short-form McGill Pain Questionnaire [38], and an 11-point visual analog pain scale [26]. The Oswestry contains 10 items pertaining to the patient's perception of his or her disability and pain, such as their ability to stand. The patient responses are transformed to a perceived disability score that ranges from 0–100 with higher scores representing greater disability. The Oswestry scores have been shown to be reliable and valid over numerous studies of patients with lumbar spine impairments [8]. The McGill [38] contains 15 pain descriptor terms, like throbbing, pertaining to the patient's perception of their pain complaints. The patient categorizes their pain for each descriptor. The patient responses can be transformed to three pain scores, a sensory, affective, and a total score. For this study we used the total pain score to quantify the patient's perception of the quality of their pain. Scores range between 0–45 with higher scores representing greater concern for their pain complaints. The McGill pain scores have been shown to have good reliability and validity [38]. The visual analog scale allowed the patients to rate their maximal pain intensity over the previous 24 hours. Scores range from 0 (no pain) to 10 (worst possible pain) [26]. Pain scores from visual analog scales have been shown to have adequate reliability and validity [26,27].

Three clinician observations were made during functional testing, which was performed during the intake FCE and again weekly during the WC program using the abbreviated FCE. The performance-based functional tests allowed assessment of physical demand characteristic level (PDL) and workplace tolerance (WT). The patient's PDL was assessed using a lifting task from floor to waist level during the FCE following standardized operational definitions described by Isernhagen [24,25]. The method used to assess the lifting ability of the patient has been described as the kinesiophysical method of lifting where a specific testing protocol is utilized to standardize the procedures for repeatability and to establish observational criteria to quantify safe levels of work-related performance. The patient interacts with the therapist during the lifting task where weights are progressively added during the test. The therapist or the patient can terminate the test, but

Table 2
Physical Demand Level Scoring

Score	Label	Lifting Loads
1	Sedentary	10 lbs or less
2	Sedentary – Light	> 10 and ≤ 15 lbs
3	Light	> 15 and ≤ 20 lbs
4	Light – Medium	> 20 and ≤ 35 lbs
5	Medium	> 35 and ≤ 50 lbs
6	Medium – Heavy	> 50 and ≤ 75 lbs
7	Heavy	> 75 and ≤ 100 lbs
8	Very Heavy	Over 100 lbs

Adopted from Matheson [33].

it is the therapist who determines the test endpoint by utilizing safety criteria. To determine the test endpoint, the therapist observes the patient during the lifting task and determines if the patient's body movements in response to the load lifted are within what is operationally defined as safe. If the therapist terminates the testing by utilizing safety criteria, the lifting is considered a maximal effort by the patient and a true endpoint of the patient's abilities. The endpoint of the floor to waist lift, i.e., the maximum safe lifting amount in pounds of the kinesiophysical floor to waist lift, has been shown to have good inter-tester reliability for comparison of lifts that are "safe" and "unsafe" [47] and to have good intra- and inter-tester reliability [25,41]. The validity of the maximum safe lifting amount of the kinesiophysical floor to waist lift has been supported [34,42].

The patient had the option to stop the test at any point or not accept the next increment of weight to be lifted. When the patient stops the test, he or she describes his or her reason for terminating the test. If the patient stops the test before the therapist determines the test endpoint by utilizing the safety criteria, the therapist documents that the patient limited their lifting. We operationally defined the maximum floor to waist level lift to be the maximum amount of weight lifted determined either by the therapist or the patient.

The maximum weight lifted from floor to waist was characterized by using the physical demand levels (PDL) originally described in the Dictionary of Occupation Titles (DOT) [52], and later modified by Matheson [33] (Table 2). Once the patient's PDL was characterized using the functional testing, a targeted job was identified for each patient, along with the expected PDL for the targeted job. The targeted job was the job to which the patient was expected to return following the WC program. We did not have targeted jobs for all patients. For those without targeted jobs, appropriate work-related targets were established as WC goals. The targeted job's or the work-related PDL became one of the goals of the WC program for each patient. The

Table 3
Patient Self-report Scores

Measure	Intake*	Discharge*	Change**
McGill	<i>n</i> = 44 14.6 ± 9.3, 1, 40	<i>n</i> = 32 11.8 ± 10.6, 0, 45	<i>n</i> = 30 2.8 ± 9.3, -25, 22, <i>P</i> = 0.334
Oswestry	<i>n</i> = 38 41.0 ± 18.1, 4, 82	<i>n</i> = 26 25.0 ± 17.2, 0, 63	<i>n</i> = 26 14.1 ± 14.1, -8, 56, <i>P</i> < 0.001
Pain Scale	<i>n</i> = 43 5.8 ± 2.6, 1, 10	<i>n</i> = 34 3.9 ± 3.1, 0, 9	<i>n</i> = 29 1.6 ± 2.5, -3, 8, <i>P</i> = 0.007

McGill = total McGill short form score (higher scores represent more concern for the patient's perception of their pain complaints); Oswestry = Oswestry Low Back Pain Disability score (higher scores represent more disability related to their lumbar complaints); Pain Scale = 11 point visual analog pain scale (higher scores represent more pain).

*Values are sample size, mean ± standard deviation, minimum, maximum.

**Values are sample size, mean ± standard deviation, minimum, maximum, probability of one-sample t-test ($H_0 = 0$).

final measure of improvement of lifting ability of the patient was a dichotomous variable of PDL goals met, which was coded as goals met (i.e., the patient lifted the amount of weight required on the job) or goals not met (i.e., the patient did not lift the amount of weight required on the job).

For the quantification of the patient's workplace tolerance (WT), the therapist observed the patient during the FCE performing up to 20 standardized tests for non-material handling tasks commonly performed in work settings. The non-material handling tasks included sitting, standing, bending, walking, climbing stairs or ladders, overhead and horizontal reaching, squatting, kneeling, crawling, balance, fine finger manipulation, simple hand grasp, forceful hand grasp, object assembly, operating foot controls, pushing, and pulling. On occasion, the patient identified non-material handling tasks not commonly tested that he/she had to perform on the job, and when this occurred, these tasks were simulated and tested as well, such as a waitress simulating carrying a platter of dishes, a laborer shoveling, or a housekeeper vacuuming or washing windows or mirrors. The therapist observed the patient performing each non-material handling activity to quantify the patient's level of functional capacity and workplace tolerance. The non-material handling tasks were performed using standardized testing procedures, and the therapists used standardized methods to observe and record their perceptions of how well the patient accomplished the task.

Following WT testing, the therapist identified the non-material handling tasks that were performed in a less than optimal manner, such as a task performed with poor body mechanics or only tolerated for a short period of time. Although there is no operational definition of performing a non-material handling task in a "less than optimal manner", and therefore there are no

reliability or validity statistics on these observations, the clinicians attempted to be consistent with their observations and recordings. The successful performance of the tasks performed in a less than optimal manner was identified as a goal of each patient in the WC program. As described for the PDL measure above, the final measure of improvement of workplace tolerance for the patient was a dichotomous variable of WT goals met, which was coded as goals met or goals not met.

2.5. Data analysis

Descriptive statistics were used to characterize the raw data. Changes in patient self-reported scores over the WC program were assessed using one-sample t-tests ($H_0 = 0$). The associations between clinician reported performance-based improvement of lifting ability and workplace tolerance and patient self-reported pain and perceived disability were assessed using two-sample t-tests. Retrospective power (power = $1 - \beta$) analyses [5] were performed on all non-significant results ($\alpha = 0.05$).

3. Results

Change in patient self-reported total McGill, Oswestry, and visual analog pain scores are displayed in Table 3. For all scores, the higher value represents more pain or disability. The change scores were calculated as intake minus discharge, so positive change scores represent improvement. Average scores for all scales improved. The improvement in lumbar disability and visual analog pain scores was significant ($P < 0.05$), but the McGill total scores were not different than 0, which represents no change in the patient's concern

Table 4
Physical Demand Level Classification from FCE

PDL Level	Intake FCE (<i>n</i> = 50)	Discharge FCE (<i>n</i> = 47)
Sedentary	6	2
Sedentary – Light	4	0
Light	48	19
Light – Medium	20	23
Medium	16	32
Medium – Heavy	4	17
Heavy	2	2
Very Heavy	0	4

Values are percents.

about their pain complaints ($P = 0.33$, Power = 0.36) over the course of the WC program.

Raw PDL levels related to performance-based floor to waist lifting ability from initial FCE to WC program termination are displayed in Table 4. The scores represent an improvement overall in the patients' ability to lift more weight at WC discharge. Twenty-eight percent of patients did not change their PDL over the course of the WC program, but 38% of the patients increased the PDL one level, 30% increased their PDL two levels, and 4% increased their PDL three levels. Forty-one percent of the patients met their PDL goals, and 64% of the patients met their workplace tolerance goals.

The associations in improvement in performance-based lifting ability and workplace tolerance and improvement in patient self-reported pain and disability are displayed in Table 5. None of the associations was significant ($P > 0.05$, Power values range 0.07 to 0.47). There was a tendency for patients who met their PDL or WT goals to have better improvement in disability and visual analog pain scores, but the differences were not significant.

4. Discussion

The results of this study demonstrate improvement in 1) patient self-reported disability and pain intensity and 2) clinician reported performance-based lifting ability and workplace tolerance over the course of a work-conditioning program for patients with lumbar spine impairments. A majority of patients improved their PDL (72%) or met their WT goals (64%), but only 41% of the patients met their PDL goals. The results also show that we could not confirm ($P > 0.05$) associations between patient self-reported improvement in disability, pain concerns, and pain intensity and clinician reported assessment of patients improvement in lifting

ability and workplace tolerance, although small sample sizes reduced the power of these results. The primary finding of this study is that in spite of continuing concern about pain complaints, patients can decrease their perceived disability and pain intensity, increase their lifting ability, and improve their workplace tolerance over the course of a short (i.e., average 19 days) work conditioning program, but our data did not support an association between patient self-report of pain and disability and clinician assessed change in lifting ability and workplace tolerance.

Although there have been noteworthy recent studies concerning functional capacity evaluations [10–15,19–21,25,28,29,32,34–36,40–45,47], only a limited number of studies describing the success of work conditioning intervention were found [31]. Similar to our study, most descriptions of WC studies use samples of convenience because well controlled, i.e., randomized controlled trials, are extremely difficult to implement in the field of WC, which is predominated by patients receiving workers' compensation benefits many of whom may be represented by legal counsel. Given this significant limitation, for the moment, convenience data provide a place to start the study of effectiveness of WC programs. However, samples of convenience suffer from potential selection bias, offer no control of extraneous variables, and more rigorous designs are recommended.

In addition to the concern about the design of studies of WC program effectiveness, researchers and clinicians should use measures that are reliable, valid and sensitive to change. Fortunately, there is a growing body of literature supporting the reliability and elucidating the validity of many measures collected during an FCE, particularly the lifting tasks [25,28,32,34–36,41–43,47]. However, because the results of these studies are mixed, much psychometric investigation is still needed.

Of clinical importance, some involved in the care of patients with chronic pain complaints contend patients should not be referred to work conditioning programs and participate in physical activities if they have pain complaints. However, investigators describing guidelines for functional capacity evaluations contend that if the patient's pain complaints are "medically stable", i.e., operationally defined as the consistent presence of a set of signs and symptoms, the continued pain complaints are not a contraindication to participation in an FCE [18]. We believe those FCE guidelines can apply to patients entering work conditioning programs. Our data support the fact that patients with chronic pain

Table 5
Association Between Change in Clinician Assessment of Lifting Ability and Workplace Tolerance and Change in Patient Self-Report of Pain and Disability

Self-Report Scales	Lifting Ability		Workplace Tolerance	
	Met Goals	Some Goals Met	Met Goals	Some Goals Met
McGill Change				
n	14	16	20	9
Mean(SD)	-0.6(9.9)	5.8(8.0)	2.6(9.9)	4.3(8.2)
t (df)	-2.0(24.9)		-0.5(18.8)	
P	0.06		0.63	
Power (1- β)	0.47		0.07	
Oswestry Change				
n	14	12	18	7
Mean(SD)	17.2(14.4)	10.4(13.3)	15.3(15.8)	11.0(10.2)
t(df)	1.2(23.9)		0.8(17)	
P	0.22		0.44	
Power (1- β)	0.22		0.10	
Pain Change				
n	12	17	19	9
Mean(SD)	2.0(2.5)	1.3(2.6)	2.0(2.9)	0.7(1.6)
t(df)	0.8(24.1)		1.4(25)	
P	0.42		0.16	
Power (1- β)	0.11		0.23	

McGill = total McGill short form score (higher scores represent more concern for the patient's perception of their pain complaints); Oswestry = Oswestry Low Back Pain Disability score (higher scores represent more disability related to their lumbar complaints); Pain Scale = 11 point visual analog pain scale (higher scores represent more pain).

n = sample size; SD = standard deviation; t = two-sample t-test; df = degrees of freedom; P = probability.

complaints can participate in properly designed, safe, physical activities such as those in work conditioning programs, and when they do, self-report of their disability and pain intensity tend to improve concurrently with improvement in performance-based measures of lifting ability and workplace tolerance.

Importantly, recent studies [11,12] support FCEs as behavioral tests influenced by multiple factors, including psychological. Data support the fact that psychological concerns affect performance-based tests during patient examinations, i.e., patients with psychological concerns report their perceptions of results of clinical movement tests differently compared to patients without psychological concerns, particularly for patients with low back pain [51], and data support that patients commonly under-report their actual performance-based abilities as assessed during an FCE [44]. Although not measured, we expect patient psychological concerns affected our data performance-based observations. Therefore, caution should be exercised when interpreting results of studies on FCEs and WC programs, particularly where there are limited number of papers describing the reliability or validity of the clinician's observation of performance-based assessment of non-material handling tasks and workplace tolerance [4]. In our study, we used standardized observations and

recording processes in an attempt to reduce variability of the WT scores. Future studies should examine precise reliability, validity and responsiveness of the WT scores, including any psychosocial influence on the measures.

Return-to-work is a common goal of WC programs. We did not assess return-to-work because whether a person returns to work is frequently not associated with the patient's medical condition or the treatment rendered [3,50]. Furthermore, exactly what constitutes a successful return-to-work outcome is difficult to ascertain. Although an operational definition for return-to-work in good health has been proposed and used successfully for clinical decision-making [3,7], we did not collect return-to-work information. Future research should initiate the study of factors associated with successful return-to-work following WC program attendance.

The final measures of lifting ability used two different stopping rules for the performance-based measure: clinician determination of maximum safe lifting ability and patient determination of when to stop lifting. The maximum safe lifting ability determined by the clinician has been shown to be reliable [25,41,47] and valid [34,42,47]. Studies have been published concerning psychophysical lifting where workers select the

maximum amount of weight they believe they can comfortably lift throughout a work shift [48], but to our knowledge, these studies have never been replicated by patients with acute or chronic lumbar impairments [25]. Therefore, we are not aware of any reliability or validity statistics for the maximum amount of weight lifted as determined by the patient. However, given the current litigious environment, particularly for patients in WC, we elected to err on the conservative side and accept the patient's wishes when they elected to stop lifting even though they had not reached their maximum safe limit as described by Isernhagen [25]. Unfortunately, the data concerning who stopped the tests was not collected consistently enough to allow statistical analyses. This could have affected our results.

This retrospective study suffered from missing data (Tables 3–5), poor design, and poor controls, which could have affected our results for several reasons, including selection bias. Studies have delineated the effect of patient selection bias [39,46] and this study is no exception. Therefore, the results should be interpreted with caution. Future studies should endeavor to have complete data sets and assess the impact of the missing data, if possible.

5. Conclusion

This study investigated the association between patient self-report of pain and disability and clinician observed performance-based measures of floor to waist lifting ability and workplace tolerance during a work-conditioning program for patients with lumbar spine impairments. Patient self-reported disability and pain intensity and clinician reported performance-based lifting ability and workplace tolerance improved over the course of the work-conditioning program. A majority of patients improved their PDL or met their WT goals, but no associations between patient self-reported improvement in disability, pain concerns, and pain intensity and clinician reported performance-based measures of lifting ability and workplace tolerance were identified. A small sample size reduced the power of the results. However, in spite of continuing concern about pain complaints, patients decreased their perceived disability and pain intensity, increased their lifting ability, and improved their workplace tolerance over the course of the work-conditioning program.

Acknowledgements

We would like to acknowledge Crystal Carter, PTA for data collection and Kerri Patterson for data entry.

References

- [1] American Physical Therapy Association, *Guidelines for Programs for Injured Workers*, 1995, American Physical Therapy Association.
- [2] American Physical Therapy Association, *Occupational Health Guidelines: Evaluating Functional Capacity*, 1997, American Physical Therapy Association Board Of Directors.
- [3] M.L. Baldwin, W.G. Johnson and R.J. Butler, The error of using returns-to-work to measure the outcomes of health care, *Am J Industr Med* **29** (1996), 632–641.
- [4] S. Brouwer, M.F. Reneman, P.U. Dijkstra, J.W. Groothoff, J.M. Schellekens and L.N. Goeken, Test-retest reliability of the Isernhagen Work Systems Functional Capacity Evaluation in patients with chronic low back pain, *J Occup Rehabil* **13**(4) (2003), 207–218.
- [5] J. Cohen, *Statistical Power Analysis for the Behavior Sciences*, (2nd Edition), New York, NY: Academic Press Inc; 1988.
- [6] L.E. Darphin, Work-hardening and work-conditioning perspectives, in: *The Comprehensive Guide to Work Injury Management*, S.J. Isernhagen, ed., Aspen Publishers, Gaithersburg, MD, 1995, pp. 443–462.
- [7] C.E. Dionne, R. Bourbonnais, P. Fremont, M. Rossignol, S.R. Stock and I. Larocque, A clinical return-to-work rule for patients with back pain, *CMAJ* **172**(12) (2005), 1559–1567.
- [8] J.C.T. Fairbank, J. Couper, J.B. Davies and J.P. O'Brian, The Oswestry Low Back Pain Disability Questionnaire, *Physiotherapy* **66** (1980), 271–273.
- [9] J.M. Fritz and J.J. Irrgang, A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec Back Pain Disability Scale, *Phys Ther* **81**(2) (2001), 776–788.
- [10] V. Gouttebauge, H. Wind, P.P.F.M. Kuijer and M.H.W. Frings-Dresen, Reliability and validity of functional capacity evaluation methods: a systematic review with reference to Blankenship system, Ergos work simulator, Ergo-kit and Isernhagen work system, *Int Arch Occup Environ Health* **77** (2004), 527–537.
- [11] D.P. Gross, Measurement properties of performance-based assessment of functional capacity, *J Occup Rehabil* **14**(3) (2004), 165–174.
- [12] D.P. Gross and M.C. Battie, Factors influencing results of functional capacity evaluations in workers' compensation claimants with low back pain, *Phys Ther* **85**(4) (2005), 315–322.
- [13] D.P. Gross and M.C. Battie, The prognostic value of functional capacity evaluation in patients with chronic low back pain: part 2, *Spine* **29**(8) (2004), 920–924.
- [14] D.P. Gross, M.C. Battie and J.D. Cassidy, The prognostic value of functional capacity evaluation in patients with chronic low back pain: part 1, *Spine* **29**(8) (2004), 914–919.
- [15] D.L. Hart, Relation between three measures of function in patients with chronic work-related pain syndromes, *J Rehabil Outcomes Meas* **2**(1) (1998), 1–14.
- [16] D.L. Hart, Tests and measurements in returning injured workers to work, in: *The Comprehensive Guide to Work Injury Management*, S.J. Isernhagen, ed., Aspen Publishers, Gaithersburg, MD, 1995, pp. 345–367.
- [17] D.L. Hart, S. Berlin, P.E. Brager et al., Development of clinical standards in industrial rehabilitation, *J Orthop Sports Phys Ther* **19**(5) (1994), 232–241.
- [18] D.L. Hart, S.J. Isernhagen and L.N. Matheson, Guidelines for functional capacity evaluation of people with medical conditions, *J Orthop Sports Phys Ther* **18**(6) (1993), 682–686.

- [19] E. Innes and L. Straker, Reliability of work-related assessments, *Work* **13** (1999), 107–124.
- [20] E. Innes and L. Straker, Validity of work-related assessments, *Work* **13** (1999), 125–152.
- [21] E. Innes, N.L. Tuckwell, L. Straker and T.E. Barrett, Test-retest reliability on nine tasks of the physical work performance evaluation, *Work* **19** (2002), 243–253.
- [22] S.J. Isernhagen, Contemporary issues in functional capacity evaluation, in: *The Comprehensive Guide to Work Injury Management*, S.J. Isernhagen, ed., Aspen Publishers, Gaithersburg, MD, 1995, pp. 410–442.
- [23] S.J. Isernhagen, Functional Capacity Evaluation: rationale, procedure, utility of the kinesiophysical approach, *J Occup Rehabil* **2**(3) (1992), 157–168.
- [24] S.J. Isernhagen, Functional Capacity Evaluation and work hardening perspectives, in: *Contemporary Care For Painful Spinal Disorders*, T. Mayer, V. Mooney and R. Gatchel, eds, P.A. Malvern, Lea and Febiger, Philadelphia, PA, 1991, pp. 328–345.
- [25] S.J. Isernhagen, D.L. Hart and L.M. Matheson, Reliability of independent observer judgments of level of lift effort in a kinesiophysical Functional Capacity Evaluation, *Work* **12** (1999), 145–150.
- [26] M.P. Jensen, P. Karoly and S. Braver, The measurement of clinical pain intensity: a comparison of six methods, *Pain* **27** (1986), 117–126.
- [27] M.P. Jensen, J.A. Turner, J.M. Romano and L.D. Fisher, Comparative reliability and validity of chronic pain intensity measures, *Pain* **83** (1999), 157–162.
- [28] S. Kersnovske, L. Gibson and J. Strong, Item validity of the physical demands from the Dictionary of Occupational Titles for functional capacity evaluation of clients with chronic low back pain, *Work* **24** (2005), 157–169.
- [29] P.M. King, N. Tuckwell and T.E. Barrott, A critical review of functional capacity functions, *Phys Ther* **78** (1998), 852–865.
- [30] W. Kuijjer, E.H. Gerrits and M.F. Reneman, Measuring physical performance via self-report in healthy young adults, *J Occup Rehabil* **14**(1) (2004), 77–87.
- [31] D. Lechner, Work hardening and work conditioning interventions: do they affect disability? *Phys Ther* **74**(5) (1994), 471–493.
- [32] D. Lechner, J. Jackson, D. Roth and K. Straaton, Reliability and validity of a newly developed test of physical work performance, *J Occup Med* **36** (1994), 997–1004.
- [33] L.N. Matheson, *Work Capacity Evaluation: Interdisciplinary Approach to Industrial Rehabilitation*, Anaheim, CA, Employment and Rehabilitation Institute of California, 1984.
- [34] L.N. Matheson, S.J. Isernhagen and D.L. Hart, Relationships among lifting ability, grip force, and return to work, *Phys Ther* **82** (2002), 249–256.
- [35] L.N. Matheson, V. Mooney, J.E. Grant, M. Affleck, H. Hall, T. Melles, R.L. Lichter and G. McIntosh, A test to measure lift capacity of physically impaired adults. Part 1 – development and reliability testing, *Spine* **20**(19) (1995), 2119–2129.
- [36] L.N. Matheson, L.C. Rogers, V. Kaskutas and M. Dakos, Reliability and reactivity of three new functional assessment measures, *Work* **18**(1) (2001), 41–50.
- [37] R. Melzack, The McGill Pain Questionnaire: major properties and scoring methods, *Pain* **1** (1975), 275–299.
- [38] R. Melzack, The short-form McGill Pain Questionnaire, *Pain* **30** (1987), 191–197.
- [39] D.B. Pryor and K.L. Lee, Methods for the analysis and assessment of clinical databases: the clinician's perspective, *Stat Med* **10**(4) (1991), 617–628.
- [40] M.F. Reneman, *Functional Capacity Evaluation in Patients with Chronic Low Back Pain: Reliability and Validity*, Thesis University of Groningen, the Netherlands, 2004.
- [41] M.F. Reneman, P.U. Dijkstra, M. Westmaas and L.N. Goeken, Test-retest reliability of lifting and carrying in a 2-day functional capacity evaluation, *J Occup Rehabil* **12**(4) (2002), 269–275.
- [42] M.F. Reneman, A.S. Fokkens, P.U. Dijkstra, J.H. Geertzen and J.W. Groothoff, Testing lifting capacity: validity of determining effort level by means of observation, *Spine* **30**(2) (2005), E40–E46.
- [43] M.F. Reneman, S.M.H.J. Jaegers, M. Westmaas and L.N.H. Goeken, The reliability of determining effort level of lifting and carrying in a functional capacity evaluation, *Work* **18**(1) (2002), 23–27.
- [44] M.F. Reneman, W. Jorritsma, J.M. Schellekens and L.N. Goeken, Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain, *J Occup Rehabil* **12**(3) (2002), 119–129.
- [45] M.F. Reneman, W. Jorritsma, J.M. Schellekens and L.N. Goeken, Relationship between kinesiophobia and performance in a functional capacity evaluation, *J Occup Rehabil* **13**(4) (2003), 227–285.
- [46] L. Resnik and D.L. Hart, Using clinical outcomes to identify expert physical therapists, *Phys Ther* **83**(11) (2003), 990–1002.
- [47] R.L. Smith, Therapists' ability to identify safe maximum lifting in low back pain patients during functional capacity evaluation, *J Orthop Sports Phys Ther* **19**(5) (1994), 277–281.
- [48] S.H. Snook and V.M. Ciriello, The design of manual material handling tasks revised tables of maximum acceptable weights and forces, *Ergonomics* **34**(9) (1991), 1197–1213.
- [49] C.A. Velozo, P.J. Lustman, D.M. Cole, J.A. Montag and E. Eubanks, Prediction of return to work by rehabilitation professionals, *J Occup Rehabil* **1**(4) (1991), 271–279.
- [50] G. Waddell and A.K. Burton, Occupational health guidelines for the management of low back pain at work: evidence review, *Occup Med* **51**(2) (2001), 124–135.
- [51] M. Werneke and D.L. Hart, Centralization: association between repeated end-range pain responses and behavioral signs in patients with acute non-specific low back pain, *J Rehabil Med* **37**(5) (2005), 286–290.
- [52] United States Department of Labor, *Dictionary of Occupational Titles*, (4th ed.), US Government Printing Office, Washington, DC, 1991.