Journal of Occupational Rehabilitation, Vol. 13, No. 2, June 2003 (© 2003)

The Effect of Order of Testing in Functional Performance in Persons With and Without Chronic Back Pain[1]

Andrew J. Haig,[2,3] Michael E. Geisser,[2] Carrie Nicholson,[2] Ebony Parker,[2] Karen Yamakawa,[2] Derrick Montomery,[2] and Ethan Booker[2]

Batteries of individually standardized physical and functional tests are commonly used to assess persons with chronic back pain disability. The order of testing may affect performance on later tests. One hundred and fifty patients with>3 months of back pain disability underwent a multidisciplinary Spine Team Assessment involving Physical Therapy, Occupational Therapy, Pain Psychology, and Vocational Rehabilitation Counselor assessments at a university spine clinic. Seventeen back healthy volunteers performed the physical component of the assessment. For the volunteers the order of testing was randomized to OT tests first or PT test first, with 0.5 h rest between the tests. For patients the order of testing was arbitrarily set by an alternating schedule, with 1 h psychological testing between the two components. For both the patients and volunteers, among the 14 test components, there was no significant difference ($p > 0.05$) in performance with order of testing. This held true for the subgroup of patients who put out good cardiac effort. Volunteers performed better than patients on all individual tests ($p < 0.001$). Results suggest that the order of physical testing during a Spine Team Assessment does not affect test performance either in chronic low back disabled patients or in volunteers.

KEY WORDS: functional capacity evaluation; back pain; exercise; chronic pain.

INTRODUCTION

The functional abilities of persons with back pain are of substantial importance. Functional assessments are used for many reasons, ranging from determination of appropriate treatment course to assessment of treatment progress to measurement of treatment effectiveness to determining work disability (1–5).

Functional test batteries for persons with chronic back pain disability are commonly scored as if each component were performed in isolation. But individual test norms are usually established in isolation from other fatiguing events. Given fear, fatigue, pain, and

[1]This paper was presented at the International Society for the Study of the Lumbar Spine, June 15–17, Edenburgh, Scotland.

[2]Department of Physical Medicine and Rehabilitation, The University of Michigan, Ann Arbor, Michigan.
[3]Correspondence should be directed to Andrew J. Haig, MD, The Spine Program, University of Michigan, 325 E. Eisenhower, Ann Arbor, Michigan 48108; e-mail: andyhaig@umich.edu.

the effect of warm-up exercises, it is possible that performance on sequential tests will be different from performance on individual tests.

The current study looks at a typical multidisciplinary team assessment for chronic back pain. Its null hypothesis is that there is no effect of test order on physical and functional test performance in asymptomatic volunteers and in persons with chronic back pain disability.

METHODS

Subjects

Volunteer subjects were recruited from the community via posters and personal contact, and provided consent for an institutional review board approved study. Volunteers denied any previous back surgery, back pain disability, current back pain. Other disabling conditions were *not* excluded. Volunteers completed an initial questionnaire and were screened for any medical contraindications to aggressive exercise testing under supervision prior to randomization.

Patients included persons who underwent a multidisciplinary Spine Team Assessment (STA), at a university based spine program (6,7). This assessment is intended for adults age 18–55 with substantial chronic back pain disability of more than 3 month duration. Patients are prescreened for spinal, cardiac, and other contrandications to aggressive exercise testing.

Testing

The clinical spine team assessment is a half-day multidisciplinary assessment that includes assessments by a physical therapist, occupational therapist, psychologist, and rehabilitation (or vocational) counselor. A team meeting is held with a physiatrist, and a report with suggestions for assessment is sent to the referral source.

The testing by the physical therapist (PT) and by the occupational therapist (OT) are fairly physical, with the OT testing assessing more whole body function and the PT testing assessing more isolated performance measures. For patients, the PT and OT components of testing are always separated by an hour of physical rest, typically an interview with a counselor. The volunteers rested for at least 0.5 h between test components, during which they filled out psychological tests. Among the multiple written tests taken by the groups was the SF-36, a widely established test of health-related quality of life. A combined physical component score, described by Fanuele and *et al.* was used in this study (8).

"PT tests" were performed by a physical therapist and an exercise physiologist. They included the following: the Sorensen test of trunk extension endurance involves prone lying with the trunk extended off the end of a table for as long as possible, with the test truncated at 2 min (9). The upper abdominal test involved holding the back off the floor with arms crossed in front of the chest, and with knees bent, for as long as possible, up to 2 min. The lower abdominal test involved lying in a prone position and raising the straight legs a few inches off the ground, holding this position for up to 2 min.

A submaximal cardiac fitness test using the YMCA protocol is performed on an exercise bicycle (10). Projected maximum oxygen consumption (MVO_2) is calculated. If a person does not maintain pedal cadence long enough for measurement of at least two heart rates between 110 and 85% of the person's age predicted maximum heart rate, with two corresponding workloads, projected MV_{O_2} is not calculated. In addition to projected cardiovascular fitness, heart rate is used as a measure of physiologic effort, with persons who achieve more than 70% of their predicted maximum heart rate (220−age) considered to have made a good physiologic effort. In addition, the volunteers, but not the patients, underwent an upper body endurance test of repeated seated bench press at 0.5 Hz at 60% of ideal body weight. They also underwent a lower body endurance test, involving repeated seated leg extension of 60% ideal body weight at 0.5 Hz.

"OT tests" were performed by an Occupational Therapist. They included The Progressive Isoinertial Lifting Evaluation (PILE) (11). This commonly used test involves lifting progressively heavier weights from floor to waist, and from waist to shoulder at a rate of 12/min. Maximum weight lifted and heart rate are outcome measures. In our center the subjects are not told how much weight they are lifting. The Slow PILE, an addition to the PILE intended to reflect strength, rather than endurance, involves lifting at one lift per minute of progressively increased weights, beginning at the highest weight lifted on the PILE (12). The Functional Activities Screening Test (FAST) involves 5 min of repeated stooping and bending, 5 min of repeated reaching and twisting, 2 min of static kneeling, static squatting, and static crouching (13).

Randomization

The volunteers were randomized by drawing straws to determine whether the "PT tests" or the "OT tests" were performed first. Patients were assigned to "PT tests" first or "OT tests" first based on a predetermined clinic schedule. Because the clinic schedule had tree "OT first" appointments for each "PT first" appointments, the distribution of patients was approximately three "OT tests" first to one "PT tests" first. No patient related factors entered into the determination of whether PT or OT testing occurred first.

Data Analysis

Data were entered into a Microsoft access database and checked for errors. The Statistical Package for the Social Studies (SPSS), version 9, was used for statistical analysis. t-test and chi-square test were performed to compare the differences for volunteers and patients, and within each group, by the testing order, "PT first" or "OT first." Statistical test results with $p < 0.05$ level are accepted as significant in this study.

RESULTS

Subject demographics are described in Table I. The patients and the volunteers differed significantly in age (35.3 years vs. 40.4 years), weight (154.0 1b vs. 193.3 1b), % ideal body weight (101.6 vs. 127.7), and the calculated SF-36—Physical Component Summary (PCS) score (55.5 vs. 25.6). The volunteers averaged 5.5 points higher in the PCS score than the U.S. norm of 50.0 (8). No statistically significant difference in demographics was found between volunteers randomized to "PT first" vs. "OT first," or between patients randomized to "PT first" and "OT first."

Table I. Subject Demographics Examined for the Group Difference Between Patients and Volunteers, and Within Each Group Between Persons Randomized to Physical Therapy First or Occupational Therapy First

					Volunteer				Patient			
			t-test/chi-square test				*t*-test/chi-square test				*t*-test/chi-square test	
	Volunteers all	Patients all	*t*	Significance	OT first	PT first	*t*	Significance	OT first	PT first	*t*	Significance
Number	17	150			8	9			112	38		
Age	35.29 (13.39)	40.41 (8.25)	−2.253	0.026*	30.25 (10.14)	39.78 (14.86)	−1.523	0.148	40.95 (8.20)	38.84 (8.31)	1.363	0.175
Sex (% male)	47.1%	52.7%	0.192	0.661[a]	50.0%	44.4%	0.052	0.819[a]	52.7%	52.6%	0.000	0.996[a]
Weight (1b)	154.03 (32.79)	193.29 (46.80)	−3.357	0.001**	157.31 (33.42)	151.11 (33.96)	0.379	0.710	192.75 (44.09)	194.93 (54.21)	−0.238	0.812
% ideal body weight	101.55 (16.21)	127.65 (26.98)	−3.900	0.000**	100.66 (12.37)	102.33 (19.75)	−0.206	0.840	126.93 (25.33)	129.85 (31.84)	−0.555	0.580
Exercise/week	2.47 (1.07)	na			2.75 (1.04)	2.22 (1.09)	1.019	0.325	na	na		
SF 36—Physical Component Summary	55.52 (3.56)	25.61 (6.94)	16.877	0.000**	56.92 (3.15)	54.43 (6.65)	1.431	0.174	25.66 (6.73)	25.42 (7.78)	0.148	0.883

Note. Values given in the parentheses are standard deviation (SD).
[a] Chi-square test.
* $p = 0.05$ (two-tailed); ** $p = 0.01$ (two-tailed).

Table II presents test performance of the patients and volunteers. The volunteers performed significantly better than the patients in every single test—cardiovascular fitness, PILE, Slow PILE, and FAST. The data show that the order of testing did not seem to affect the performance within the volunteers or within the patients. Figures 1 and 2 illustrate the findings in two representative tests. In both the the Sorenson test (a "PT test") and the PILE (an "OT test"), the patients performed worse than controls, but within these groups performance was not affected by the order of testing.

One might wonder if the patients simply did not put out much effort, and thus were not fatigued. Cardiac response to exercise is one measure of effort. As shown in Fig. 3, when the subjects who put out good cardiac effort (>70% of maximum age predicted heart rate) were isolated, there remains no trend.

DISCUSSION

The results of this study show that the order of testing resulted in no difference in test performance for either able bodied persons or persons with chronic back pain. Results also showed that persons with back pain disability are deficient in both functional performance and physical conditioning compared to back healthy volunteers. The results may be useful to clinicians and researchers who wish to evaluate the abilities of persons with back pain, but generalization should be tempered by the limitations of this methodology.

It was remarkable that no test was affected by the performance of other tests beforehand. For example, a "PT test" such as the Sorenson test is intended to fatigue back muscles that are subsequently used in "OT test" such as the PILE. Yet performance in the PILE was no worse when the Sorenson test was performed beforehand. One might conclude that this is because the back pain patients did not fatigue muscles on one test sufficiently to affect their performance on the next test. The percent of maximum heart rate achieved on the bicycle ergometer suggests less than full physiologic effort by many of the back pain patients. But the back healthy subjects appeared to put out good effort on the bicycle ergometer, and they also did not have a significant deterioration of performance after a previous test.

Two possibilities remain that may explain the lack of change in performance: The rest interval may have been sufficient to allow full recovery, and the tests do not fatigue the same muscle fibers, even though the same general muscle groups may be used. The first possibility leads us to believe that standardization of functional assessments is important. If in fact a different protocol provides less rest, it may result in contamination of test results based on previous tests performed. Designers of functional capacity evaluations and multidisciplinary assessment protocols for rehabilitation planning should consider this. The second possibility points out a difficulty with functional capacity evaluations as measures of true functional capacity. Different tests may be so specific as to not be generalizable to life functions that are somewhat different from the test circumstances. For this reason we advocate for the use of functional assessments primarily for developing rehabilitation planning, and less for declaration of present or future abilities.

There were substantial differences between the back healthy volunteers and the patients in all aspects of physical performance. While one time performance on functional tasks such as the PILE and the FAST might be attributed by some to pain, fear, or other psychological factors, the difference in performance on the cardiovascular fitness testing as well as the

Table II. Volunteer and Patient Performance on Physical and Occupational Therapy Tests in Relation to Order of Testing

					Volunteer				Patient			
			t-test/chi-square test				*t*-test/chi-square test				*t*-test/chi-square test	
	Volunteers all	Patients all	*t*	Significance	OT first	PT first	*t*	Significance	OT first	PT first	*t*	Significance
Number	17	150			8	9			112	38		
PT tests												
Upper abs (s)	98.65 (34.91)	na			110.63 (26.52)	88.00 (39.39)	1.370	0.191	na	na		
Lower abs (s)	103.06 (35.14)	na			100.75 (37.70)	105.11 (34.86)	−0.248	0.808	na	na		
Sorenson (s)	94.12 (33.78)	36.87 (41.21)	5.463	0.000**	89.63 (35.00)	98.11 (34.23)	−0.505	0.621	33.41 (38.93)	46.67 (46.36)	−1.524	0.130
Bench press (1b)	84.88 (21.75)	na			90.71 (17.18)	80.33 (24.74)	0.944	0.361	na	na		
Leg extension (1b)	83.63 (23.75)	na			90.71 (17.18)	78.11 (27.54)	1.057	0.308	na	na		
MET level	10.90 (3.05)	5.53 (3.54)	5.814	0.000**	11.72 (2.82)	10.25 (3.23)	0.954	0.356	5.60 (3.52)	5.31 (3.66)	0.403	0.687
Maximum heart rate on the bicycle	80.68 (6.18)	78.26 (10.97)	0.886	0.377	78.78 (7.01)	82.37 (5.15)	−1.213	0.244	78.18 (11.40)	78.54 (9.60)	−0.149	−0.882
% invalid bicycle test	5.9%	25.2%		[a]	12.5%	0.0%	[a]		24.8%	26.7%	0.045	0.832[b]
OT tests												
PILE low lift % expected weight lifted	114.69 (33.18)	36.89 (18.76)	14.690	0.000**	125.98 (32.22)	104.65 (32.45)	1.357	0.195	36.74 (19.30)	37.36 (17.19)	−0.170	0.865
% maximum heart rate	78.31 (7.09)	65.33 (11.40)	4.581	0.000**	78.99 (7.93)	77.70 (6.69)	0.363	0.721	65.73 (11.58)	64.06 (10.89)	1.095	0.275
PILE high lift % expected weight lifted	105.24 (45.66)	43.64 (24.26)	8.835	0.000**	120.50 (51.08)	91.69 (38.04)	1.329	0.204	45.28 (25.08)	38.54 (21.04)	1.434	0.154
% maximum heart rate	70.94 (9.05)	63.29 (11.72)	2.593	0.010**	70.36 (10.62)	71.44 (8.03)	−0.239	0.815	63.97 (12.00)	61.06 (10.60)	1.231	0.220
Good cardiac effort (heart rate > 70% projected maximum)—Patients												
PILE low lift (% expected)									45.22 (19.05)	46.41 (13.71)	−0.191	0.849
PILE high lift (% expected)									51.01 (20.94)	44.49 (20.56)	0.747	0.460
Slow PILE low lift (1b)	77.94 (21.65)	30.88 (19.68)	8.477	0.000**	86.88 (13.87)	70.00 (24.87)	1.695	0.111	30.35 (19.74)	32.40 (20.11)	−0.345	0.731
Slow PILE high lift (1b)	54.06 (19.34)	25.60 (17.92)	5.530	0.000**	64.29 (14.84)	46.11 (19.33)	2.055	0.059	26.16 (18.73)	24.00 (15.83)	0.400	0.691
FAST total time (min)	15.89 (0.38)	10.45 (4.14)	5.229	0.000**	16.00 (0.00)	15.77 (0.54)	1.207	0.248	10.36 (4.08)	10.75 (4.37)	−0.494	0.622
FAST (% total completer)	87.5	15.0		[a]	100.0	75.0	[a]		14.4	16.7	0.108	0.742[b]
Stoop5 (% completers)	93.8	24.5		[a]	100.0	87.5	[a]		25.2	22.2	0.133	0.716[b]
Reach5 (% completers)	100.0	51.0		[a]	100.0	100.0	[a]		46.8	63.9	3.159	0.075[b]
Knee12 (% completers)	100.0	74.1		[a]	100.0	100.0	[a]		73.9	75.0	0.018	0.893[b]
Stoop2 (% completers)	100.0	33.3		[a]	100.0	100.0	[a]		30.6	41.7	1.490	0.222[b]
Squat2 (% completers)	87.5	49.0		[a]	100.0	75.0	[a]		47.7	52.8	0.275	0.600[b]

Note. Values given in the parentheses are standard deviation (SD).
[a] Chi-square test not performed because of low (< 5) count in one or more cells.
[b] Chi-square test.
** $p = 0.01$ (two-tailed).

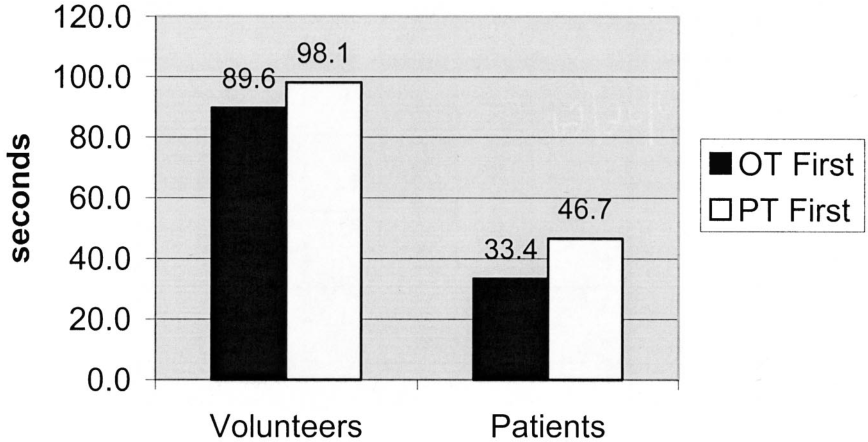

Fig. 1. An example of a "PT test." Sorensen test trunk endurance in patients and volunteers.

increase in BMI are evidence of long-standing physical deconditioning. As obvious as these findings may seem to rehabilitation clinicians, the literature does not contain many direct comparison between controls and chronic back pain patients on a battery of physical performance tests such as the Spine Team Assessment. This data adds support to the idea that deconditioning is a factor in back pain disability.

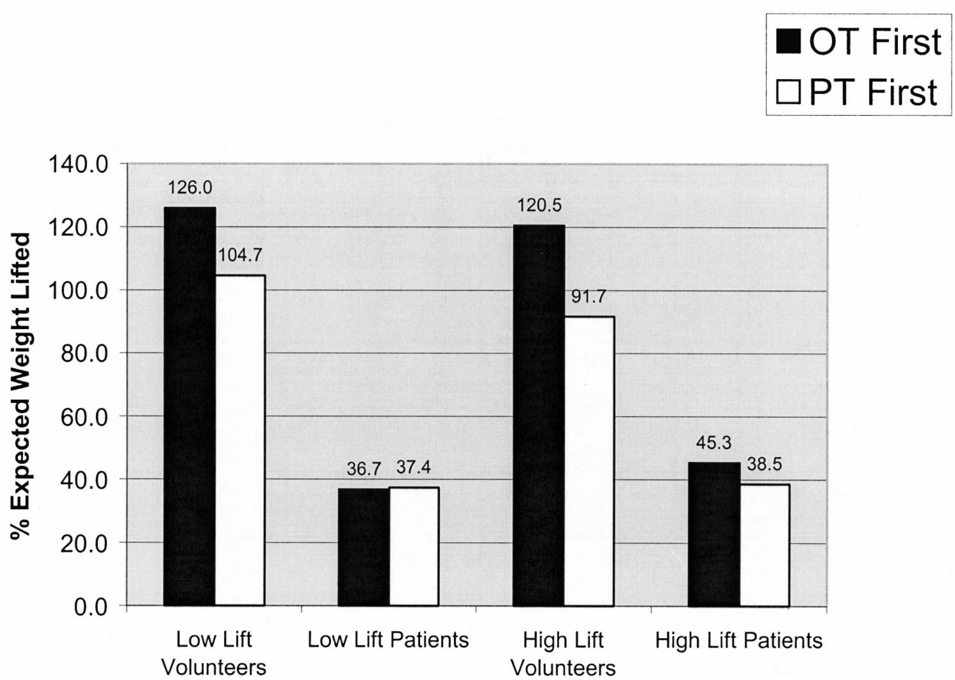

Fig. 2. An example of an "OT test." Pile results in patients and volunteers.

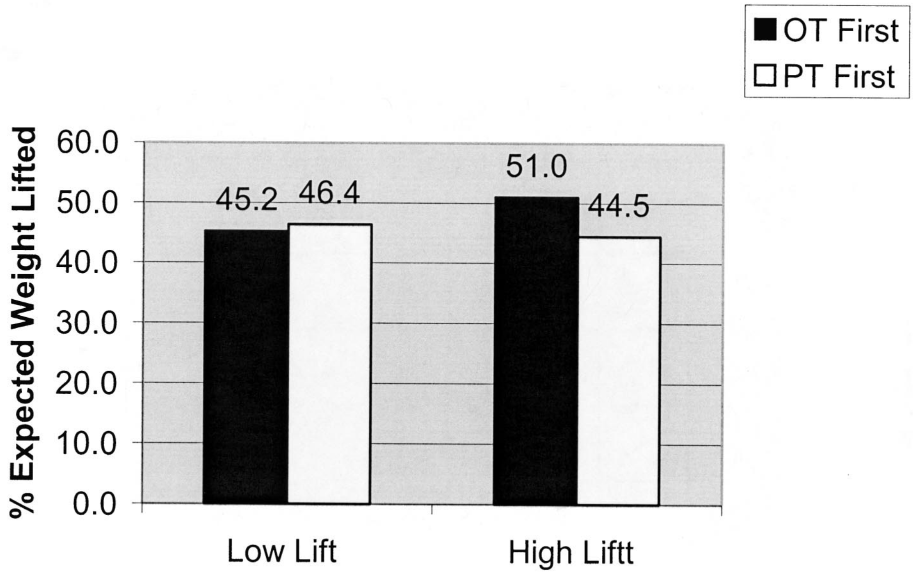

Fig. 3. PILE performance in patients who put out good physiologic effort during testing. Thirty-two percent of the patients who performed in low lift and 25.3% performed in high lift that put out good cardiac effort (heart rate >70% projected maximum) are included.

The subject population probably reflected the population of persons with back pain disability in general. At the University of Michigan, almost all referrals from primary care physicians and specialists for evaluation of back pain disability go through the Spine Program. The back pain population had similar scores on the SF-36 as another large cohort of persons with back pain disability (8). The back healthy subjects were slightly more functional than the general public on the SF-36, and performed somewhat better than the norms on the PILE test. One would expect from a population of volunteers, as the general population includes persons whose physical disability would make them less likely to volunteer. While there were statistically significant 5-year age differences between the volunteers and patients, this may not be clinically significant, especially in light of the drastic differences between these controls and the patients. This was a relatively small control population, so extrapolation should be done with caution.

The Spine Team Assessment was chosen by the Michigan Rehabilitation Engineering Research Center for Ergonomic Solutions for Employment to provide a codified, scientifically measurable, yet individualized, and effective multidisciplinary assessment for persons with chronic back pain disability. The methodology has been described in detail, and software has been devised to support duplication (6,7,14). It is probably similar in scope and intensity to other team assessments. The current findings may not apply to more intensive assessments, or those with less rest between tests.

CONCLUSION

The current study suggests that, for assessments similar to the Spine Team Assessment, the order of testing does not affect test performance among patients with back pain and back healthy volunteers. Back pain patients appear to be substantially deconditioned compared to back healthy volunteers.

ACKNOWLEDGMENTS

This is a publication of the University of Michigan Rehabilitation Engineering Research Center which is funded by the National Institute on Disability and Rehabilitation Research of the United States Department of Education under Grant #H133E980007. The opinions contained in this publication are those of the grantee and do not necessarily reflect those of the United States Department of Education.

REFERENCES

1. Tramposh AK. The functional capacity evaluation: Measuring maximal work abilities. *Occup Med* 1992; 7(1): 113–124.
2. Kraus J. The independent medical examination and the functional capacity evaluation. *Occup Med* 1997; 12(3): 525–556.
3. Mayer TG, Gatchel RJ, Kishino N, Keeley J, Capra P, Mayer H, Barnett J, Mooney V. Objective assessment of spine function following industrial injury. A prospective study with comparison group and one-year follow-up. *Spine* 1985; 10(6): 482–493.
4. Harten JA. Functional capacity evaluation. *Occup Med* 1998; 13(1): 209–212.
5. Hart DL, Isernhagen SJ, Matheson LN. Guidelines for functional capacity evaluation of people with medical conditions. *J Orthop Sports Phys Ther* 1993; 18(6): 682–686.
6. Haig AJ, Geisser ME, Theissen M, Michel B, Yamakawa K. The Spine Team Assessment—Physical and psychosocial performance of 500 adults with chronic low back pain disability. *J Occup Rehabil* (manuscript submitted for pulication).
7. Fanuele JC, Birkmeyer NJO, Abdu WA, Tosteson TD, Weinstein JN. The impact of spinal problems on the health status of patients. Have we understimated the effect? *Spine* 2000; 25(12): 1509–1514.
8. Haig AJ, Theissen M, Geisser M, Michel B, Yamakawa K. Team decision making for Spine Team Assessment—Standardizing the multidisciplinary assessment for chronic back pain. *J Occup Rehabil* (manuscript submitted for pulication).
9. Novy DM, Simmonds MJ, Olson SL, Lee CE, Jones SC. Physical performance: Differences in men and women with and without low back pain. *Arch Phys Med Rehabil* 1999; 80: 195–198.
10. Golding LA, Myers CR, Sinning WE. (ed.). *The Y's way to physical fitness.* Champaign, IL: Human Kinetic Publishers, 1989, pp. 88–98.
11. Mayer TG, Barnes D, Kishino ND, Nichols G, Gatchel RJ, Mayer H, Mooney V. Progressive isoinertial lift evaluation 1. A standardized protocol and normative database. *Spine* 1988; 13: 993–1002. (Published erratum appears in *Spine* 1990; 15:5.
12. Chaudhuri S, Haig AJ, Buccholz R, Yamakawa K, Lamphiear R. The slow progressive isoinertial lifting evaluation: What does it mean? *Arch Phys Med Rehabil* 2000; 81: 1281.
13. Ruan CM, Haig AJ, Geisser ME, Yamakawa K, Buchholz RL. Five functional screening tests predict poor outcomes in physical and psychosocial testing in individuals with chronic spinal pain. *J Occup Rehabil* 2001; 11: 119–132.
14. Rehabilitation Team Assessments, LLC. 3316 The Glade, Ann Arbor, MI. (www.rehab-team.com).