Research Report

Relationships Among Lifting Ability, Grip Force, and Return to Work

Background and Purpose. The relationship between functional capacity evaluation (FCE) data and work disability has not been studied. The validity of FCE testing results in terms of subsequent return to work (RTW) was the focus of this exploratory study. Subjects and Methods. Six hundred fifty adults of working age were evaluated as part of a standardized FCE. Clients were contacted by telephone 6 months after the FCE to determine their work status. Predictor variables were gender, age, time off work, maximum safe loads during 3 dynamic lifts, and isometric grip force. Other variables measured were whether or not the client returned to work (RTW-Y/N) and level of return to work (RTW level). Results. A multivariate logistic regression analysis demonstrated that the more time a worker was away from work, the less likely was RTW. Male subjects were less likely to return to work than female subjects. The more weight lifted from floor to waist, the more likely was RTW. The logistic regression equation correctly classified 80.3% of the subjects who returned to work and 56.6% of the subjects who did not return to work. Each of the 3 lift tests was related to RTW level, whereas the grip force tests were not related to either RTW-Y/N or RTW level. Discussion and Conclusion. Time off work and gender were the strongest predictors of RTW, but certain FCE subtests of lifting were related to RTW and RTW level for people with work-related chronic symptoms. Grip force was not related to RTW. [Matheson LN, Isernhagen SJ, Hart DL. Relationships among lifting ability, grip force, and return to work. *Phys Ther.* 2002;82:249–256.]

Key Words: *Functional capacity evaluation, Grip force, Kinesiophysical, Lift ability, Return to work, Work disability.*

Leonard N Matheson, Susan J Isernhagen, Dennis L Hart

For many years, hypothesized relationships among impairments, functional limitations, and disabilities have been a focus in rehabilitation.[1,2] Rehabilitation professionals have referenced the relationship between impairment measured at the organ level, functional limitation in terms of task performance, and disability in terms of participation in major life activities in the development of methods to treat[3,4] and evaluate[5] work disability. The disability benefits system that was established in 1956 by the US Social Security Administration (SSA) has stimulated interest in the validity of using impairment and functional limitation data to determine disability (inability to work). The SSA, in certain circumstances, uses impairment and functional limitation data to make decisions concerning disability and, thus, eligibility to receive disability benefits.

Recently, disability determination based on information about the functional capacity of the claimant has been the focus of the SSA's Re-Design Project.[6] In 1998, more than 2 million people applied for disability benefits with the SSA, and 11 million Americans received benefits based on disability, worth approximately $77 billion.[7] In addition to the federal system, each state in the United States and each province in Canada have a separate disability benefits system for people with work-related symptoms, as do all countries with modern economies. In addition, private insurance companies that insure against disability have their own disability determination systems. Many of these systems rely on health care professionals' opinions regarding the insured person's functional abilities and limitations to predict disability. This reliance on data about function occurs even though the validity of such data to predict disability has not been well studied.[8,9]

Information from what are commonly called "functional capacity evaluations" (FCEs) has been used in making decisions about disability for many years.[10,11] In this process, a person with medical impairment is provided benefits if his or her abilities are not adequate for the demands of work. Information from FCEs is used to translate the effect of the impairment in terms of ability to perform work tasks. In this article, *functional capacity evaluation* is defined as a "detailed examination and evaluation that objectively measures the client's current level of function in terms of the demands of competitive employment."[12(p48)] In this context, the primary purpose of the FCE is to compare a client's functional abilities with the demands of his or her work[12(p47)] in order to allow the safe return to work (RTW).[3,13]

As FCEs have evolved, standards for test development and service delivery have been promulgated.[3,13–16] In these standards, reliability and validity are of paramount importance. Although the reliability of some scores from FCE subtests[17,18] and some FCE batteries[19] has been studied, limited reliability data are available for most FCEs.[20] In general, FCE scores have been shown to be reliable when strict operational definitions are developed and implemented.[16,17] A few researchers[8,21] have investigated the validity of FCE scores in terms of defining the relationship between performance on functional tests and subsequent employment. The authors of a review of the validity of data from 28 work-related functional assessments reported that "there was . . . no instrument that demonstrated moderate to good validity in all areas. Very few work-related assessments were able to demonstrate adequate validity in more than one area, or with more than one study."[8(p145)] The authors emphasized that there is a need for further validity research.

The purpose of our study was to determine the validity of functional capacity tests of lifting ability and grip force in terms of whether or not clients seeking rehabilitation subsequently returned to work and the level of work to which they returned. These functional capacity tests are components of a standardized FCE that is used at multiple clinical sites in North America. Previous research has shown these tests to yield reliable measurements,[17,22] but the validity of data from these tests in terms of guiding RTW has not been studied. The lift test

LN Matheson, PhD, is Assistant Professor, Program in Occupational Therapy, Washington University School of Medicine, 4444 Forest Park Ave, St Louis, MO, 63108 (USA) (mathesonl@msnotes.wustl.edu). Address all correspondence to Dr Matheson.

SJ Isernhagen, PT, is President, Isernhagen Work Systems, Duluth, Minn.

DL Hart, PT, PhD, is Director of Consulting and Research, Focus On Therapeutic Outcomes Inc, Knoxville, Tenn.

Dr Matheson assisted with the research concept/design, performed the data analyses, provided institutional liaison, and performed the majority of the writing. Ms Isernhagen assisted with the research concept/idea, was responsible for data collection, provided consultation, and assisted with the writing and editing. Dr Hart assisted with the research concept/design, provided statistical consultation, and assisted with the writing and editing. The authors thank Jessica Brown, OTR, MA, for her assistance with data entry and auditing of data records and Joanna Penny, OTR, MSOT, for her assistance with review of the scientific literature.

This project was reviewed and approved by the Human Subjects Committee of the Washington University School of Medicine.

This article was submitted January 30, 2001, and was accepted October 3, 2001.

Table 1.
Demographic Comparison of Groups

	Total Sample Size	Returned to Work			Did Not Return to Work		
		n	$\bar{X}$	SD	n	$\bar{X}$	SD
Age (y)[a]	650	349	40.1	9.8	301	43.1	10.2
Time off work (mo)[a]	602	315	6.8	8.6	287	17.1	16.9
Percentage of males/females	650	349	59.3/40.7		301	61.2/38.1	

[a] Independent one-way analysis of variance ($P<.05$).

intratester reliability was examined in a study in which 12 blinded judges viewed random videotape segments of a lifting evaluation of 3 subjects.[17] Depending on the lift test segment, agreement among raters was found to range from $\kappa=.68$ to $\kappa=1.00$ using the Cohen kappa statistic adjusted for chance agreement. In a study of 27 subjects with an average interval of 1 week, Pearson product moment correlations of grip force test-retest reliability were found to range from $r=.79$ to $r=.93$ for either the dominant or nondominant hand, depending on whether single or multiple measurements were taken.[22]

Our study examined whether there is a relationship between subjects' RTW (eg, "I returned to a new job in the same company") or level of return to work (RTW level) and performance of lifting or hand-grasp subtests in the FCE. We studied FCE performance data collected approximately 6 months prior to a telephone follow-up interview, at which time information was collected concerning RTW status.

Method

Subjects

In this retrospective study, we used a sample of convenience of clients from 25 clinics in 16 states in the United States and one province in Canada affiliated with the Isernhagen Work System (IWS).* All clients were working-age adults who were not working due to reported functional limitation. All clients had received an FCE following the IWS format (IWS-FCE) between January 1991 and April 1998 as part of the RTW decision-making process. In the normal course of practice, attempts to interview by telephone all clients tested at participating clinics were made 6 months after the IWS-FCE. In total, 803 clients (60.5% males) were located and interviewed. The telephone follow-up interview was accomplished a mean of 6.8 months (SD=2.5, range=4–9) after completion of the IWS-FCE. At the time of the interview, all clients were questioned concerning whether they had received work hardening, work conditioning, or any other medical or rehabilitative intervention in the interval between the IWS-FCE and the follow-up. Because our research concerned the validity of FCE performance for subjects who had completed active treatment, the group of 148 clients who reported that they had received any of these clinical services between IWS-FCE and follow-up were excluded from subsequent analyses. This was the only exclusionary criterion.

Data were analyzed on an anonymous basis after confidentiality safeguards for the study had been reviewed and accepted by the senior author's institutional review board. There were no group-wise differences, as determined by independent one-way analyses of variance (ANOVAs), between the excluded and included samples in terms of age, time off work, time to follow-up interview, or gender proportion (all $P>.05$). The study sample of 650 subjects had a mean age of 41.5 years (SD=10.1, range=19–71). Subjects reported being off work a mean of 14.1 months (SD=11.7, range=0–84) prior to the IWS-FCE. The sample consisted of 349 subjects (53.6%) who had returned to work and 301 subjects (46.4%) who had not returned to work. Demographic comparison of the groups is presented in Table 1. As a group, subjects who returned to work were younger (40.1 years versus 43.2 years) and had been off of work for a shorter period of time (6.9 months versus 17.0 months) ($P<.05$) than those who did not return to work.

Procedure

Subjects were tested with the IWS-FCE protocol.[23] The IWS-FCE protocol was finalized in 1988 and has not been revised. There were no changes in the protocol over the course of the data collection period. Across data collection sites, standardized training was required of all evaluators. The IWS-FCE is a standardized test battery designed to quantify safe physical abilities of the client[24] for the purpose of assisting in the RTW decision-making process. The IWS-FCE is administered by a physical therapist or occupational therapist who has been formally trained. The IWS-FCE involves 29 functional assessment subtests (10 strength tasks that include or are

* Isernhagen Work Systems, 1015 E Superior St, Duluth, MN 55802.

components of manual materials handling and 19 movement tasks that either do not involve materials handling or in which the materials handling does not focus on strength). Each subtest in the IWS-FCE has an operational definition that has been described in an examiner's manual.[23]

For our study, subtests of the IWS-FCE were selected for study because each had previously been shown to have what we considered acceptable reliability.[17,25,26] The subtests were 3 measures of lifting ability and 2 measures of grip force, each of which is described below.

Lifting ability. The standard protocol for evaluating lifting ability uses a method that has been described.[17,23] This protocol consists of *progressive performance testing*, which we operationally define as sequential lifting events in which the amount of weight lifted is increased until a "safe maximum lift" is achieved.[17] The starting position is the same for all subjects: standing erect with elbows flexed at a right angle and with hands positioned on the handles of a 30.48- × 30.48- × 30.48-cm (12- × 12- × 12-in) box placed on a shelf that creates the proper degree of elbow flexion. The starting position is the same for all subjects. A lifting event consists of 5 lift repetitions performed at sequential weights, beginning at a weight that can be lifted easily by the subject and gradually progressing to safe maximum endpoint. The *safe maximum endpoint* is operationally defined as the highest load (measured in pounds) lifted 5 times where a stable spine is maintained and predetermined physiological limits are not exceeded (eg, heart rate), as described in the reference manual[24] and training program. The ability of trained evaluators to identify a safe maximum lift endpoint has been shown to be reliable.[17,25] These operational definitions are applied to 3 different lifting ability subtests:

1. Floor-to-waist lift. The person being evaluated begins in the starting position, lowers the box to the floor, and lifts it back to the starting position.

2. Waist-to-crown-level lift. The person being evaluated begins in the starting position and lifts the box to a shelf until the hands are at crown (top of the head) height. The box is then lowered to the starting position.

3. Horizontal lift. The person being evaluated begins in the starting position, lifts the box from the shelf, and carries the box 1.2 m (4 ft) horizontally to a shelf at the same height, where the box is placed on the shelf. The box is then lifted again and carried to the starting position.

Table 2.
Return-to-Work Levels (RTW Level)

RTW Level	Description of RTW Level
1	I returned to my prior job without modification in the same company.
2	I returned to my prior job with modification in the same company.
3	I returned to a new job in the same company.
4	I returned to a new job in a new company.

Grip force. Whole-hand isometric grip force is measured (in pounds) with the Jamar analog hydraulic dynamometer† using the procedure described by Mathiowetz et al[27] in all 5 grip span positions for each hand. We used the average of 3 maximum-effort squeezes for the grip span position that was greatest for the individual as the hand strength datum for each hand. In addition, because age and gender are important covariates for grip force,[27] a *z*-score transformation[28] of each average hand strength value was calculated based on published normative data.[27] This calculation provided an absolute value as well as a norm-referenced value.

Data for 2 self-report variables—RTW and RTW level—were collected during a telephone interview approximately 6 months after the IWS-FCE. The first variable (RTW), concerning whether or not the subject had returned to work (RTW-Y/N), was recorded as the subject's response to the question, "Have you returned to work?" This response was recorded as "Yes" or "No." If the subject answered "Yes" to the RTW-Y/N question, a question addressing the second variable was posed: "What level of work are you doing?" The subject was asked to select one response from the 4-response variable (RTW level) described in Table 2.

Data Analysis

To guard against bias in sample selection, and confirm that the selected sample was not different from the sample that was excluded, univariate analyses of demographic variables were conducted using chi-square analysis for nominal data and one-way ANOVA for ratio data. To study the differences between the group that returned to work and the group that did not return to work, we performed a chi-square analysis for nominal data and a one-way ANOVA for each of the ratio-level performance variables. To examine the relationship between IWS-FCE variables and whether or not a person returned to work, a separate one-way ANOVA was conducted on each performance variable across RTW-Y/N groups. Multiple logistic regression analysis[28] was per-

† Asimow Engineering, Santa Fe Springs, CA 90670.

Table 3.
Performance Comparison of Groups

	Total Sample Size	Returned to Work				Did Not Return To Work			
		n	$\bar{X}$	SD	Range	n	$\bar{X}$	SD	Range
Floor-to-waist lift (pounds of load)[a]	627	345	37.7	23.8	0–110	282	22.9	15.9	0–115
Waist-to-overhead lift (pounds of load)[a]	637	344	27.9	15.8	0–70	293	21.1	15.0	0–65
Horizontal lift (pounds of load)[a]	640	345	45.0	23.8	0–140	295	32.9	22.3	0–110
Right grip force (pounds of force)[a]	624	334	86.4	34.3	0–193	290	79.1	40.2	0–172
Left grip force (pounds of force)[a]	622	332	83.4	32.1	0–196	290	75.6	35.2	0–154

[a] Independent one-way analysis of variance ($P<.05$).

Table 4.
Multivariate Logistic Regression Analysis of Return-to-Work Likelihood (n=539)[a]

Variable	Regression Coefficient (β)	Standard Error	χ^2	*P*	Odds Ratio
Gender: male[b]	−.815	.276	8.71	.003	0.443
Age	−.012	.010	1.64	.200	0.988
Time off work[b]	−.066	.011	39.66	<.001	0.936
Floor-to-waist lift[b]	.018	.008	4.81	.028	1.018
Waist-to-overhead lift	.003	.013	0.07	.793	1.003
Horizontal lift	−.001	.010	0.01	.926	1.001
Right grip force	−.002	.005	0.23	.632	0.998
Left grip force	.010	.005	3.25	.072	1.010
Intercept[b]	.595	.548	1.18	.278	1.812

[a] Multivariate logistic model: log-likelihood=−314.46, χ^2=115.18, $P<.0001$. Observations correctly classified=69.4%.
[b] $P<.05$.

formed on the predictor variables of gender, age, time off work, and the 5 performance variables against RTW-Y/N as the outcome variable. This analysis was performed on the set of 539 subjects with complete data sets for these variables. Of the original sample of 650 subjects, 48 subjects did not have data for time off work, 28 subjects were missing grip force data, and 35 subjects were missing lifting ability data. To address the relationship between IWS-FCE performance and RTW level, each performance variable and RTW level was compared using a series of one-way ANOVAs with Scheffé *post hoc* analyses when a one-way ANOVA was significant. All tests of significance were conducted with $\alpha=.05$.

Results

Relationships Among FCE Performance and Return to Work

The comparisons between the group that returned to work and the group that did not return to work are presented in Table 3. The separate one-way ANOVAs demonstrated that, for each performance variable, those who returned to work performed better than those who did not return to work (all $P<.05$).

The relative contribution of the performance data in the context of the demographic variables was assessed with multivariate logistic regression analysis using RTW-Y/N. These results are summarized in Table 4 for the sample of 539 subjects with complete data sets. The combination of variables in the logistic regression equation correctly classified 80.3% of the subjects who returned to work and 56.6% of the subjects who did not return to work, with an overall correct classification of 69.4%. The overall percentage was higher ($P<.05$) than the chance rate of 53.8%. Time off work ($\chi^2_{539}=39.66$, $P<.001$) and subjects' gender ($\chi^2_{539}=8.71$, $P=.003$) were related to RTW-Y/N, with more time off work and male gender predicting a lower likelihood of returning to work. Age ($\chi^2_{539}=1.64$, $P=.200$) was not related. Of the performance variables, only floor-to-waist lift ($\chi^2_{539}=4.81$, $P=.028$) was related to RTW-Y/N, with greater lift ability related to improved likelihood of RTW.

Return-to-Work Level

The relationships between IWS-FCE subtests and RTW level are summarized in Table 5. For comparison, the values of subjects who did not return to work are also included, although their data were not part of the statistical analysis. The ANOVA for each performance variable demonstrated that floor-to-waist lift ($F_{3,341}=6.91$, $P=.0002$), waist-to-overhead lift ($F_{3,340}=3.64$, $P = .0131$), and horizontal lift ($F_{3,341} = 5.03$, $P = .002$) were related to RTW level, whereas neither right grip force

Table 5.
Comparison of Performance Characteristics (in Pounds) Across Level of Return to Work (RTW Level)

	Floor-to-Waist Lift[a,b]		Waist-to-Overhead Lift[b]		Horizontal Lift[a,b]		Right Grip Force		Left Grip Force	
RTW Level	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
1	44.3	25.6	30.8	15.9	50.5	24.3	85.4	32.9	80.2	31.1
2	35.1	19.3	26.1	14.8	40.8	21.0	88.0	32.2	84.7	31.2
3	28.6	19.2	23.2	12.8	37.9	20.8	77.0	34.1	76.7	30.6
4	36.0	25.3	28.6	17.7	44.9	26.0	92.1	38.1	91.0	35.0
Did not return to work[c]	24.4	22.8	20.9	14.9	32.6	22.2	78.7	39.9	75.5	35.1

[a] Difference between level 1 and level 2, $P<.05$.
[b] Difference between level 1 and level 3, $P<.05$.
[c] Data from clients who did not return to work are provided for comparison; their data were not part of the statistical analyses.

($F_{3,330}=2.17$, $P=.091$) nor left grip force ($F_{3,328}=2.59$, $P=.053$) were related.

The Scheffé *post hoc* analyses demonstrated that floor-to-waist lift and horizontal lift both could be used to differentiate subjects who returned to their prior job without modification from those who returned to either a modified job or to a new job in the same company. Waist-to-overhead lift differentiated subjects who returned to their prior job without modification from those who returned to a new job in the same company.

Discussion

Our results are in agreement with previous research with a wide variety of rehabilitation populations, that the amount of time off work is strongly related[9,29–32] to whether or not a person returns to work (RTW Y/N), with gender making modest contributions.[9,31,33] In addition to time off work and gender, we found that small amounts of variance in RTW were related to differences in performance of IWS-FCE lifting subtests. Although greater trunk force and leg force have each been associated with improved RTW,[34] our study is the first to show that the greater the lifting ability the greater the likelihood of RTW. We found that certain measures of performance are more strongly related to RTW-Y/N than are others. Results of multivariate logistic regression suggest colinearity of the performance tests. After the floor-to-waist lift variable was considered, the remaining performance variables did not provide additional explanatory power. Thus, when it is not possible to administer more than one test, measuring floor-to-waist lift is likely to be most useful.

We also found that IWS-FCE performance was related to RTW level. Performance on the floor-to-waist lift and horizontal lift differentiated between subjects who returned to work in the same company at their prior job without modification and those who returned to work in a modified job. This same combination of tests differentiated between subjects who returned to work in the same company at their prior job without modification and those who returned to work in a new job. Thus, if employers use FCE to screen employees who return to work after a period of disability, these tests may be useful. The waist-to-overhead lift is less often useful, differentiating subjects who returned to work in the same company at their prior job without modification and those who returned to work in a new job. Grip force was not related to RTW level. At least for this population, measurement of grip force is not likely to be useful for RTW decisions.

The positive relationship between lift ability and both measures of RTW is a reasonable finding and suggests that lift ability is a valuable characteristic to measure when FCE is used to guide RTW decisions. In contrast, although hand force is a widely measured performance characteristic, we found that it was not related to either RTW-Y/N or RTW level. The widespread use of grip force to predict RTW, in our opinion, should be reconsidered. Its use may be justified for other purposes,[35] but not for prediction of RTW in this population.

There was no difference in lift ability or grip force between subjects who returned to work at a new job at the same employer and those who did not return to work at all. This finding suggests that the employer's ability to provide a different job to people who otherwise would not return to work may determine outcome rather than some of the variables we measured. This requires more study.

Return to work is an outcome that is so broadly determined[36,37] that some have challenged its utility.[38] A review of several studies that used RTW as an outcome variable following rehabilitation for low back pain[9] showed that RTW is a multi-determinant outcome that often was not well defined and included several types of vocational outcomes. Because RTW is so important for individuals with disabilities and for the institutions that provide disability benefits, some authors[9,36,38] recom-

mended that standardized and consistent definitions of RTW be used, but that may mean that successful results are not individualized to each client's needs and work setting. One of the potentially useful results of this study occurred as we considered RTW as both a dichotomous variable (RTW-Y/N) and as a multi-level variable (RTW level). Because RTW level has important financial consequences, the use of a multi-level RTW variable is recommended. This study demonstrated the utility of RTW as an outcome variable that can be studied effectively.

There were limitations in our study. This exploratory retrospective study needs to be confirmed by a prospective study with a new data set. In addition, information about job demands and job availability was not assessed. Success in RTW is dependent on both of these factors,[39] in addition to the demographic and performance variables studied. Another limitation has to do with the absence of a uniform method for comparing information about an individual's impairment with his or her functional limitations, as other authors[9] have recommended. Without this information, we were unable to study the effect on the individual of the impact of impairment on functional limitation and work disability. Finally, the time over which the data were collected is a threat to the utility of these findings. Although efforts were taken to minimize differences between sites and across time in the administration of the IWS-FCE, patterns of change in the environment such as rates of unemployment and changes in the availability of certain jobs in a local economy were more likely to be pertinent.[9]

Conclusion

Validity of data from lifting and hand-grasp subtests of a standardized FCE in terms of RTW-Y/N and the level of work to which the client returned was assessed. Although the 2 factors that had the strongest relationships to RTW Y/N were gender and the time a worker was away from work, the amount of weight lifted from floor to waist was also related to RTW-Y/N. Logistic regression models correctly classified 80.3% of the subjects who returned to work and 56.6% of the subjects who did not return to work. Lift tests were positively related to RTW level, whereas the grip force tests were not related to either RTW Y/N or RTW level.

References

1 Nagi S. Disability concepts and prevalence. In: *Mary Switzer Memorial Seminar*. Cleveland, Ohio: National Rehabilitation Association; 1975.

2 Jette A. Physical disablement concepts for physical therapy research and practice. *Phys Ther.* 1994;74:380–386.

3 Hart DL, Berlin S, Braeger P, et al. Development of clinical standards in industrial rehabilitation. *J Orthop Sports Phys Ther.* 1994;19:232–241.

4 Matheson L, Ogden L, Violette K, Schultz K. Work hardening: occupational therapy in industrial rehabilitation. *Am J Occup Ther.* 1985;39:314–321.

5 Matheson L, Gaudino E, Mael F, Hesse B. Improving the validity of the impairment evaluation process: a proposed theoretic framework. *J Occup Rehabil.* 2000;10:311–320.

6 Chater S. *Plan for a New Disability Claim Process.* Washington, DC: Social Security Administration; 1994:70.

7 Apfel K. *Social Security and Supplemental Security Income Disability Programs: Managing for Today Planning for Tomorrow.* Washington, DC: Social Security Administration; 1999:24.

8 Innes E, Straker L. Validity of work-related assessments. *Work.* 1999; 13:125–152.

9 Feuerstein M, Menz L, Zastowny T, Barron B. Chronic back pain and work disability: vocational outcomes following multidisciplinary rehabilitation. *J Occup Rehabil.* 1994;4:229–251.

10 McBride E. *Disability Evaluation and Principles of Treatment of Compensable Injuries.* Philadelphia, Pa: JB Lippincott Co; 1963.

11 Smith W. *Principles of Disability Evaluation.* Philadelphia, Pa: JB Lippincott Co; 1959.

12 *Occupational Health Guidelines: Evaluating Functional Capacity.* Alexandria, Va: American Physical Therapy Association; 1997.

13 Hart DL, Isernhagen SJ, Matheson LN. Guidelines for functional capacity evaluation of people with medical conditions. *J Orthop Sports Phys Ther.* 1993;18:682–686.

14 *Standards for Educational and Psychological Testing.* Washington, DC: American Psychological Association, American Educational Research Association, National Council on Measurement in Education; 1999.

15 Johnston M, Keith R, Hinderer S. Measurement standards for interdisciplinary medical rehabilitation. *Arch Phys Med Rehabil.* 1992;73: S3–S23.

16 Task Force on Standards for Measurement in Physical Therapy. Standards for tests and measurements in physical therapy practice. *Phys Ther.* 1991;71:589–622.

17 Isernhagen SJ, Hart DL, Matheson LN. Reliability of independent observer judgments of level of lift effort in a kinesiophysical functional capacity evaluation. *Work.* 1999;12:145–150.

18 Matheson LN, Mooney V, Grant J, et al. A test to measure lift capacity of physically impaired adults, part 1: development and reliability testing. *Spine.* 1995;20:2119–2129.

19 Lechner D, Jackson J, Roth D, Straaton K. Reliability and validity of a newly developed test of physical work performance. *J Occup Med.* 1994;36:997–1004.

20 Innes E, Straker L. Reliability of work-related assessments. *Work.* 1999;13:107–124.

21 Lechner D, Roth D, Straaton K. Functional capacity evaluation in work disability. *Work.* 1991;1:37–47.

22 Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg [Am].* 1984;9:222–226.

23 Isernhagen SJ. *Functional Capacity Evaluation Procedure Manual.* Duluth, Minn: Isernhagen Work Systems; 1988.

24 Isernhagen SJ. Functional capacity evaluation. In: Isernhagen SJ, ed. *Work Injury: Management and Prevention.* Rockville, Md: Aspen; 1988:139–174.

25 Smith RL. Therapists' ability to identify safe maximum lifting in low back pain patients during functional capacity evaluation. *J Orthop Sports Phys Ther.* 1994;19:277–281.

26 Mathiowetz V. Reliability and validity of grip and pinch strength measurements. *Phys Rehab Med.* 1991;2:201–212.

27 Mathiowetz V, Kashman N, Volland G, et al. Grip and pinch strength: normative data for adults. *Arch Phys Med Rehabil.* 1985;6: 69–74.

28 Zar J. *Biostatistical Analysis.* 2nd ed. Englewood Cliffs, NJ: Prentice-Hall Inc; 1984.

29 Black-Schaffer R, Osberg J. Return to work after stroke: development of a predictive model. *Arch Phys Med Rehabil.* 1990;71:285–290.

30 Lancourt J, Kettelhut. Predicting return to work for lower back pain patients receiving workers' compensation. *Spine.* 1992;17:629–640.

31 Lusted M. Predicting return to work after rehabilitation for low back injury. *Australian Journal of Physiotherapy.* 1993;39:203–210.

32 Milhous RL, Haugh LD, Frymoyer JW, et al. Determinants of vocational disability in patients with low back pain. *Arch Phys Med Rehabil.* 1989;70:589–593.

33 Eklund M. Chronic pain and vocational rehabilitation: a multifactorial analysis of symptoms, signs, and psycho-socio-demographics. *J Occup Rehabil.* 1992;2:53–66.

34 Fredrickson B, Trief P, Van Beveren P, et al. Rehabilitation of the patient with chronic back pain: a search for outcome predictors. *Spine.* 1988;13:351–353.

35 *Guides to the Evaluation of Permanent Impairment.* 4th ed. Chicago, Ill: American Medical Association; 1993.

36 Gatchel R, Mayer T, Hazard R, et al. Functional restoration: pitfalls in evaluating efficacy [editorial]. *Spine.* 1992;17:988–994.

37 Rainville J, Sobel J, Hartigan C, Wright A. The effect of compensation involvement on the reporting of pain and disability by patients referred for rehabilitation of chronic low back pain. *Spine.* 1997;22: 2016–2024.

38 Baldwin M, Johnson W, Butler R. The error of using returns to work to measure the outcomes of health care. *Am J Ind Med.* 1996;29: 632–641.

39 Berkowitz M, Hill M. Disability and the labor market: an overview. In: Berkowitz M, Hill M, eds. *Disability and the Labor Market: Economic Problems, Policies, and Programs.* New York, NY: ILR Press; 1986:1–27.