Development of a Post-Offer Screening Tool for Patient Support Services

Krystal Laflin, Diane Aja, Nanci Banasiak

Key Words: back injuries • human engineering (technology) • preventive health services (facilities)

Objectives. The purpose of our project was to develop a post-offer screening tool that demonstrates interrater reliability, predictive validity, and face validity and that accurately represents the physical demands of the patient support services (lifting team) job at our health care facility.

Methods. The screening tool, which consists of 11 static and dynamic tasks, was developed, using the 13 incumbent staff members of the patient support services department, to determine whether the criteria established for each task matched the physical abilities of at least 80% of the total group tested. Test-retest design was used for this study. Intraclass correlation coefficients and the Kappa statistic were used to calculate interrater reliability. Face validity was determined through the Job Similarity Questionnaire completed by all subjects.

Results. Subjects did not meet criteria established for the static knee pull and the knuckle-to-elbow lift tasks, resulting in modification of these two criteria. Interrater reliability ranged from .22 for the maximum static pull wall task to .94 for the left-hand grip strength task. Face validity ranged from 53.9% to 92.4%.

Conclusion. Although face validity of the Job Similarity Questionnaire represented a wide range, we believe that the results were homogeneous enough to continue with the screening tool unchanged, except for lowering the expected outcome on two tasks. Interrater reliability was established for 75% of the tasks. The lack of variation of data for the other 25% prevented statistical analysis of these tasks but confirmed that all staff members met the physical criteria.

At the Medical Center Hospital of Vermont (MCHV) campus of Fletcher Allen Health Care, the patient support services staff (the primary lifting team) performs the majority of dependent patient lifts, averaging up to 30 a day. From 1990 to 1992, an average of 5 orderlies per year had injuries related to patient lifting tasks, which resulted in lost work time. With only 11.5 workers in the department, this number represented 43% of the total staff.

On average, injuries related to patient handling accounted for 36% of the total workers’ compensation costs at MCHV from 1989 to 1993. Most striking was a rise in hospital workers’ compensation insurance premiums from $656,786 to $1,200,000 during a 1-year period. To decrease the incidence of injuries related to patient handling tasks by better matching staff member strength with job demands, we developed a screening tool for the lifters in the patient support services department.

There is evidence that better physical matches be-
tween job demands and employee capabilities can prevent at least a portion of injury occurrences (Chaffin, Herrin, & Keyserling, 1978; Jackson, 1990; Waters, Putz-Anderson, Garg, & Fine, 1993). One must first define the physical demands of an employment situation and then develop a methodology to match the physical capabilities of a person to the job that is not in violation of the Americans With Disabilities Act of 1990 (ADA, Public Law 101–336) so that the job can be completed safely. Our intent was to study this premise for patient support services staff members who perform patient lifting and to support the use of post-offer testing in jobs with high risk for injury within our facility.

Literature Review

Lifting Teams

Many hospitals use orderly staff members or lifting teams to carry the bulk of patient lifting responsibilities. Hospital Employee Health examined the practice of using specially trained orderly staff members (or a lifting team) who had met the physical demands of the job, according to outcomes on strength tests, to perform patient lifting tasks on one nursing floor for 1 year ("Lifting Teams Can Help Hospitals," 1994). A 63-bed neurology floor was selected because of its frequent and heavy lifting activities. In addition to special instruction in lifting techniques, the lifting team members were screened for their ability to match the lifting requirements of the orderly job. The study found that during the time and shift that the lifting team was present, there were no injuries to nurses or the lifting team, whereas during the shift when no lifting team was present, injuries occurred at a rate similar to that before the study was initiated.

Biomechanical Job Stress

Spinal compressive forces during lifting are impressive. Owen and Garg's (1993) biomechanical analysis of the forces occurring during five different manual methods for transferring patients indicated that the two-person manual lifting method exceeded the limits of 770 lb of compressive disk force set by the Occupational Safety and Health Administration (Waters et al., 1993). In contrast, pulling transfer methods with a gait belt, walking belt, patient handling sling, and mechanical aids used forces under these limits.

In addition to the physical demands in the task of lifting patients, other work factors in nursing were found to contribute to the risk of injury (Garg & Owen, 1991; Hollenbeck, Ilgen, & Crampton, 1992; Jensen, 1990; Laflin & Aja, 1995; Snook & Ciriello, 1991). These include, but are not limited to, the following:

- **Staffing factors**: lifts done alone; failure to request assistance; cumulative stress from lifting and moving patients all day; repetitive bending at the waist; overestimation of the patient's ability to help, including "it just happened" (unpredicted incident); and lifting weights exceeding the 50th percentile of maximum acceptable weight at knuckle height, which is 63.8 lb for men (Snook & Ciriello, 1991)
- **Patient factors**: unexpected collapse, uncooperativeness or combativeness, poor endurance, and unexpected occurrences (e.g., patient moving left instead of right, patient's gown getting caught on the wheelchair)
- **Equipment factors**: brakes not working on wheelchairs, stretchers, and beds; tearing of transfer sheets; and too few or lack of lifting machines
- **Environmental factors**: rooms too small and cluttered with equipment
- **Policies and procedures**: unsafe lifting and transfer practices, lack of awareness of lifting aids, and poor compliance with preventive measures (Paterson et al., 1985)
- **Personnel**: employee's attitudes and behavior toward work; social influence of friends and family members; and organizational and managerial factors, such as lack of support for team lifting (Hollenbeck et al., 1992)

Preemployment or Post-Offer Screening

Preemployment (now termed post-offer with the passage of ADA) screening tests have been used to control occupational injuries in manual materials-handling jobs, particularly when these jobs cannot be redesigned or when their physical demands cannot be reduced. The purpose of these tests is to match a person's physical capacity to the job demands (Chaffin et al., 1978; Perry, Palmer, Weyrich, & Guo, 1994). Toledo Hospital in Ohio used a post-offer screening test at the time of hire for its potential lifting staff members. The test measured the ability to lift and carry a load of 100 lb for at least four steps and the ability to generate a push–pull force of 80 lb to 90 lb ("Lifting Teams Can Help Hospitals," 1994). These criteria were set to meet the physical demands for orderlies as outlined in the Dictionary of Occupational Titles (DOT) (U.S. Department of Labor, 1991). This job title was coded in the DOT as "heavy" work because the work required lifting loads in excess of 100 lb occasionally, in excess of 50 lb frequently, or in excess of 20 lb constantly when moving objects or materials (Peacock, 1992). Research by Pheasant and Stubbs (1992) reiterated that patient handling tasks are heavy in nature, and risk of injury is greater when a worker's physical capacities do not match the demands of the job.
The use of strength testing for post-offer screening is based on the hypothesis that the risk of injury is a function of the incompatibility of a person's strength with the job demands (Perry et al., 1994). In a study of employees of a food distribution company, Perry et al. (1994) found that those persons not tested with a post-offer screening tool had almost two and one-half times the risk of injury than those receiving post-offer screening. The post-offer screening tests included the following:

- Musculoskeletal evaluation (assessment of gross posture; upper-extremity, lower-extremity, and trunk range of motion [ROM]; strength)
- Psychophysical tests (floor-to-knuckle lift, 12 in.-to-knuckle lift, knuckle-to-shoulder lift, shoulder-to-overhead lift, carry, push, pull)
- Body mechanics and physical demand capacity rating
- Isometric strength tests for arm, torso, and leg

Keyserling, Chaffin, and Herrin (1980) found similar results in their study of post-offer screening of employees of a tire and rubber plant. On the basis of an analysis of medical visits, the incidence of injury for the control group (those not matched to the job on the basis of strength) was found to be more than three times that of the experimental group (those provided with post-offer screening). Although these results were marginally significant (p < .10) when only musculoskeletal visits were considered, none of these incidents were experienced by the experimental group. Post-offer screening tests used by the authors were limited to four isometric strength tests: (a) arm test, (b) torso test, (c) arm push-out test at 46 in. high, and (d) arm pull-in test at 60 in. high.

The intent of post-offer screening is to correctly identify persons who possess the capabilities of meeting or exceeding the job requirements. Post-offer screening also provides an accurate record of a person's preemployment capacity, which provides valuable information if a worker is being rehabilitated from a workplace injury.

The National Institute of Occupational Safety and Health (NIOSH) outlined several criteria for developing an appropriate post-offer screening tool to test physical performance:

- The test should be safe to administer. The potential risks and exposure to the person should not outweigh the value of the data to be gained.
- The test should be reliable. Interrater and intrarater reliability should be tested for each measure.
- Only those tasks required or needed to perform the job should be tested.
- The test should be practical.
- The test should be able to predict future illness or injury and performance (Isernhagen, 1988).

For simulated, physically demanding strength and endurance tasks, isometric strength tests have been found to be valid and safe, whereas maximum psychophysical strength tests have inherent risks, especially when testing women for physically demanding jobs (Jackson, 1990). The average woman has only about 50% of the strength of her male counterpart (Jackson, 1990). Jackson demonstrated that although weight was positively correlated to strength, gender and age were negatively correlated.

Passage of ADA in 1990 has changed the way employers have been able to use preemployment or post-offer screening. Previously, employers could require physical screening of all applicants before a job offer, making it more difficult for persons with disabilities to have an opportunity to prove their ability to perform the essential functions of a job with reasonable accommodations. The ADA provides the following guidelines to assure fair treatment in the hiring practice:

- A functional job description must be designed before the design of the preemployment screen. Essential functions must be defined accurately for each job.
- Functional screening tests should be designed that meet the essential functions of the job. Characteristics of an essential function include the following: (a) the job exists to perform the functions, (b) the number of employees available to perform the function is limited, and (c) the function is highly specialized so that the incumbent is hired for his or her expertise or ability to perform the function.
- The test administrator will not indicate or make a hiring decision but will only score the employee as having "met criteria" or "did not meet criteria."
- Because of the short duration of the screening, endurance capacity and potential for cumulative trauma cannot be identified. The most that can be accomplished in post-offer screening is determining whether the potential employee has the physical capability to do the essential demands of the job on a short-term basis (U.S. Equal Employment Opportunity Commission & U.S. Department of Justice, 1991).

Development of the Screening Tool

We analyzed the lifting portion of the patient support services job at our facility, using NIOSH guidelines (Laffin & Aja, 1995; Waters et al., 1993). These guidelines use measurements of vertical height, horizontal reach, weight lifted, hand grasp, symmetry, and frequency of the lift. The information obtained in our analysis of the
patient support services job was combined with a review of the DOT (U.S. Department of Labor, 1991) job description for orderlies in order to establish physical criteria for the screening tool. Feedback from the department supervisor, who has performed all aspects of the job regularly, was used to help determine the tasks and tests to be used in the screening tool. Furthermore, the tasks selected were designed to assess the most difficult aspects of the lifting job.

The tests selected for the screening tool included a combination of six static and five dynamic muscle strength tests. Although the job of the orderly is dynamic by nature, dynamic tests are more variable in their results and performance than static tests and, therefore, potentially less reliable (Chaffin et al., 1978; Keyserling, Herrin, & Chaffin, 1978). Dynamic tests also require more time to administer. Dynamic tests of maximal exertion were avoided in our screening process because they increase potential for injury. Therefore, we set the lifting criteria at the stated pass-fail level to meet the job requirements on the basis of task analysis. We decided that a brief screening tool that used a combination of static and weight-limited dynamic tests would be more cost-effective and time efficient than lengthy dynamic tests and would allow us to assess a greater number of employees. This decision was supported by the knowledge that isometric strength of the back and arm muscles during lift tests were found to be related to musculoskeletal injuries (Keyserling et al., 1980).

Static pull at 37 in. and static push strength at 35 in., the most common height location of transfers done by the lifting team, is measured with the Chatillon1 dynamometer. A criterion of 90 lb was established because it equals the amount of force required to initiate the pulling of a 180-lb patient, which is the average weight of the patients transferred by the orderlies from bed to stretcher in our facility. The 90-lb criterion was determined in clinical trials on real persons before the start of the study. The participant is instructed to adopt the stance most often used to move an object and to gradually pull or push the Chatillon dynamometer against the wall, avoiding jerky motions. The average of three trials, using the peak score, is used to compensate for "learning" during the task.

Dynamic pull or push force at a height of 35 in., the amount of pull or push force needed to move a sled filled with boxes of various weights and shapes 3 ft, is measured with the Chatillon dynamometer. To develop a pull or push force of 90 lb (the criterion established), the sled is weighted with 336 lb to 341 lb. Weighting the sled with various versus fixed weight boxes allows more finite evaluation of actual pull or push force in persons unable to meet the 90-lb criterion. This method also allows more accurate replication of the 90 lb of pull-push force, while accommodating changes in the coefficient of friction between the sled and floor because of changes in floor surfaces due to weather and cleaning.

Static knee-level pull-up is used to measure leg strength. Peak torque force is measured in pounds, using the protocol for position established by Keyserling et al. (1978) for the leg–arm–torso dynamometer. Participants are tested by pulling up on the dynamometer from a squat position. A criterion of 100 lb is set to reflect the estimated effort of a single lifter participating in a two-man “dead lift” of a 180-lb to 200-lb patient and in consideration of the DOT job description for an orderly.

Static waist-level pull (bilateral biceps and arm curl test) is used to measure arm strength. Peak torque force is measured in pounds, using the static leg–arm–torso dynamometer and the position protocol established by Keyserling et al. (1978). Participants are tested by pulling up on the dynamometer with the elbows bent at a 90° angle. A criterion of 100 lb is set to reflect the half weight of a 180-lb to 200-lb patient and in consideration of the DOT job description for an orderly.

Static hand grip strength has been determined by NIOSH (Waters et al., 1993) to be an important component in lifting ability. This test has participants squeeze the Jamar dynamometer2 positioned in the second setting, arm at their side, and elbow flexed and averages three trials for both hands. The criterion established for both hands is the mean for the normative population on the basis of gender and age (Mathiowetz et al., 1985).

Maximum lift knuckle-elbow level is used to measure dynamic arm strength. It is designed to simulate the extended arm reach position used for transfers from recliner to bed, the most difficult of all transfers. The criterion of 100 lb set for this task is half the 180-lb to 200-lb weight of the average patient in our hospital (Laflin & Aja, 1995). This test measures the amount of weight one can hold on a 6-ft bar for a 3-sec count with complete extension of the arms at 90° shoulder flexion (see Figure 1). Three seconds is the average lifting interval used for NIOSH calculations (Waters et al., 1993) and the time recommended to develop maximum torque and to avoid fatigue (Keyserling et al., 1978).

Dynamic floor-to-waist lifting strength testing requires participants to lift a 45-lb, 6-ft bar with 55-lb disc weights added to simulate the length of patients typically lifted (see Figure 2). The criterion established is 100 lb to reflect

1 Manufactured by Charillon Scales, PO Box 35668, Greensboro, North Carolina 27425-5668.
2 Manufactured by Charillon Scales, PO Box 35668, Greensboro, North Carolina 27425-5668.
at least half the weight of the average patient lifted and to match the DOT job description for an orderly.

**Application of body mechanics** (low back position and stance) during the floor-to-waist lift task uses a scale similar to Matheson's (see Cole, 1995). Low back position is graded 0 if held lordotic during the entire lift, -1 if it shifted from lordotic to neutral during the lift, -2 if it became kyphotic during the lift, and -3 if excessive kyphosis occurred. Stance is graded 0 if legs were positioned shoulder width apart with one leg slightly ahead of the other, -1 if legs were positioned shoulder width apart and parallel, -2 if feet were less than shoulder width apart, and -3 if feet were touching. The passing criterion established is -2 for the combined scores.

**Cardiovascular fitness/tolerance for repetitive lifting** uses heart rate measured after the two lifting tasks by taking the carotid pulse for a 10-sec count and multiplying by 6. The criterion for target heart rate \((220 - \text{age} \times 80\%)\) is set at a level to allow new hires of general physical fitness to pass the screen.

**Dynamic strength assessing ability to perform transfers** from bed to gurney, from gurney to bed, from recliner to bed, and from bed to recliner is measured with a 165-lb weighted dummy (see Figure 3). Four transfers are done in each scenario for a total of eight to assess capacity to accomplish the transfer from both sides. Although 165 lb is less than the average patient weight of 180 lb to 200 lb, the lifting staff members reported that because of the dummy's stiff characteristic and dead weight, it subjectively felt heavier.

**Method**

**Procedure**

The intent of this project was to develop a screening tool that accurately represents the physical demands of the patient support services job and to evaluate interrater reliability and face validity for the individual tasks comprising this tool. The screening tool was administered to all current orderly staff members who comprise the lifting team (13 men) in order to ascertain their ability to meet the physical criteria established for the tool. All screening was done at the beginning of the subject's shift to help control for variables related to fatigue or job stress. Subjects were evaluated in pairs to simulate the transfer portion of their job.

The subjects signed informed consent after the intent of the screening and use of the data in research were reviewed with them. They were assured that the data were confidential and were made aware that compliance was not mandatory and that refusal would in no way jeopardize their jobs.

The subjects were assigned a number on arrival at the testing site and randomly assigned to one of the two evaluators. Demographic data were obtained, including age, weight, height, years on the job, previous injury his-
Figure 3. Simulating recliner-to-bed transfer with a dummy.

...tory, and physical activity level (see Table 1). Physical activity level was categorized as very active (exercising or playing sports three times a week), moderately active (exercising or playing sports one to two times a week) or inactive (rarely exercising or playing a sport). After the evaluator reviewed 10 basic body mechanics concepts with the subject, the subject completed a true or false test that covered these concepts. The test results were reviewed until the subject achieved 100% accuracy. The subject was then led through a warm-up that consisted of simple stretches as demonstrated by the evaluators. These included stretching arms over the head and behind the neck, neck ROM, trunk ROM, squats, and toe raises.

Both evaluators independently tested each subject, administering the 11 tasks of the screening tool sequentially. The second evaluator was blind to the results of the first in order to prevent bias. Additionally, subjects were not told their results of the first screening to prevent attempts to duplicate results for the second evaluator.

The screening tool took 30 min to administer per evaluator. After completing the first nine strength tasks, the subject was paired with another to perform the transfer portion of the screen. During this task, evaluators assessed the body mechanics of the first subject they tested. If the subject pair reported that they generally relied on a nurse to assist with a patient’s legs during a transfer, one of the evaluators or the occupational therapy aide provided that degree of assist.

After the transfer portion of the screen, the subject was given a 15-min rest break during which he was asked to complete the Job Similarity Questionnaire, which was developed to indicate how well the subjects believed each task on the screening tool matched the strength requirements of the job. Each task is rated from 1 to 9. Scores falling in the 1 to 3 or 7 to 9 range reflect an underestimation or overestimation of strength required for the actual job performance, whereas scores falling in the 4 to 6 range reflect a close match of task with job requirements. After the subjects’ rest break and completion of the questionnaire, the screening tool was administered a second time by the second evaluator.

Data Analysis

For continuous measures (the six static and five dynamic strength tests) the intraclass correlation coefficient (ICC) was used to estimate intrarater reliability (Winer, 1962). The Kappa statistic (Fleiss, 1981) was used as a measure of agreement for categorical variables (pass–fail work status). In some cases, measures of agreement were not calculated when there was little or no data variability (i.e., everyone passed or everyone failed). For the ICC measures, this included four tasks where all subjects met the criteria set with no variation in their scores: dynamic push or pull cart, floor-to-waist lift, heart rate, and ability to perform transfers. For the Kappa statistics assessing pass–fail work status, data with no variability included eight tasks: static pull wall (all passed), dynamic push or pull cart (all passed), static knee pull (all passed), floor-to-waist lift (all passed), knuckle-to-elbow lift (all failed), heart rate (all passed), and ability to perform transfers (all passed). Face validity data from the Job Similarity Questionnaire were summarized with descriptive statistics.

Results

Face validity of the Job Similarity Questionnaire ranged from 53.9% to 92.4% agreement, indicating that the screening tool closely matches the strength demands for the job tasks with the criterion set at 80% (see Table 2). Four tasks in the screening tool were rated as having met or exceeded established criteria (i.e., floor-to-waist lift, static waist-level pull-up (arm pull), “patient” lift recliner to bed, “patient” lift bed to gurney) and two tasks (dynamic push cart, dynamic pull cart) were at 77%, slightly below the 80% criteria. Although subjects reported that the dynamic push and pull cart tests may exceed requirements for patient handling, they noted that they were very similar to another job task, that of pushing the aphereser unit.

Good reliability (ICC > .75) was calculated for five of the seven data sets of raw scores analyzed (see Table 3). The highest of these, .94 for left grip and .81 for right grip, were not surprising given previous research that the Jamar dynamometer has one of the highest accuracy rates of ±3° (“Testing Grip and Pinch Strength,” 1995). ICCs for static push and pull were below .75, indicating only fair to poor reliability. Table 3 also includes the mean results in pounds and range of raw scores.

In the work status category, which noted pass or fail of the task, only four tasks (grip, arm pull, body mechanics,
Table 1
Subject Demographics

<table>
<thead>
<tr>
<th>Subject</th>
<th>Age (Years)</th>
<th>Height</th>
<th>Weight (lb)</th>
<th>Years of Experience</th>
<th>Activity Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27</td>
<td>6 ft 5 in.</td>
<td>315</td>
<td>1</td>
<td>Very active</td>
</tr>
<tr>
<td>2</td>
<td>26</td>
<td>5 ft 6 in.</td>
<td>176</td>
<td>1.25</td>
<td>Very active</td>
</tr>
<tr>
<td>3</td>
<td>40</td>
<td>5 ft 11 in.</td>
<td>210</td>
<td>19</td>
<td>Moderately active</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>5 ft 0 in.</td>
<td>155</td>
<td>1</td>
<td>Moderately active</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>6 ft 2 in.</td>
<td>187</td>
<td>1</td>
<td>Inactive</td>
</tr>
<tr>
<td>6</td>
<td>23</td>
<td>6 ft 4 in.</td>
<td>215</td>
<td>1.5</td>
<td>Very active</td>
</tr>
<tr>
<td>7</td>
<td>23</td>
<td>5 ft 11 in.</td>
<td>190</td>
<td>3</td>
<td>Moderately active</td>
</tr>
<tr>
<td>8</td>
<td>52</td>
<td>5 ft 1 1/2 in.</td>
<td>205</td>
<td>8</td>
<td>Very active</td>
</tr>
<tr>
<td>9</td>
<td>22</td>
<td>5 ft 11 in.</td>
<td>205</td>
<td>0.75</td>
<td>Very active</td>
</tr>
<tr>
<td>10</td>
<td>26</td>
<td>5 ft 10 in.</td>
<td>180</td>
<td>0.5</td>
<td>Very active</td>
</tr>
<tr>
<td>11</td>
<td>21</td>
<td>6 ft 2 in.</td>
<td>250</td>
<td>2</td>
<td>Moderately active</td>
</tr>
<tr>
<td>12</td>
<td>43</td>
<td>5 ft 3.5 in.</td>
<td>165</td>
<td>11</td>
<td>Inactive</td>
</tr>
<tr>
<td>13</td>
<td>34</td>
<td>5 ft 10 in.</td>
<td>190</td>
<td>5.5</td>
<td>Moderately active</td>
</tr>
</tbody>
</table>

Note: N = 13. M = 31 years for age, 5 ft 11.27 in. for height, 203.3 lb for weight, and 4.17 years for experience. Forty-six percent of subjects were very active, 38% were moderately active, and 15% were inactive.

push wall) had variability between the first and second test of subjects, lending the results to analysis of interrater reliability. A Kappa of 80% was calculated for grip work status, representing good agreement between evaluators. Analysis of arm pull work status (at waist level), body mechanics work status, and push wall work status yielded values less than 50%.

Discussion
The intent of this study was to develop a post-offer screening tool for the lifting team of a patient support services department and to evaluate this tool for interrater reliability and face validity. After the testing, only two tests on the screening tool were adjusted to the mean of the 13 subjects. The criterion for the static waist-level pull-up was adjusted from 100 lb to 60 lb, and the criterion for maximum knuckle-to-elbow lift was adjusted from 100 lb to 45 lb. Although individual subjects could functionally lift 100 lb of patient weight, as a group, their mean static arm lift strength was 60 lb. One theory for this discrepancy is that the static strength tests did not allow for the contribution of momentum and accessory muscles in upper-extremity strength that functional or dynamic tests require. By modifying the criteria on the screening tool on the basis of the performance of the incumbent staff, we are requiring new hires to perform job tasks at a level equal to or stronger than that of the current staff members versus merely measuring physical performance on the basis of the mechanical analysis of job demands. The latter does not take into account the physics of multidimensional human movement.

Given the results on the Job Similarity Questionnaire, we believe that the post-offer screening tool has face validity and accurately represents the job demands of the orderly staff (lifting team) at our facility. The wide range found in percentage agreement probably resulted from the small number of subjects tested. For the “patient” lift recliner to bed, “patient” lift bed to gurney, and hand grip tests, only 1 subject rated these tasks as underestimating the strength demands of the job, whereas the other 12 rated them as meeting or exceeding the strength demands. Any task scores in the 4 to 9 range will likely identify persons who meet the physical demands of the job and who are at low risk for injury. Therefore, although hand grip

Table 2
Results of Job Similarity Questionnaire

<table>
<thead>
<tr>
<th>Screening Task</th>
<th>Frequency Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Underestimates Strength Demands for Job Tasks</td>
</tr>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Static pull wall</td>
<td>0</td>
</tr>
<tr>
<td>Static push wall</td>
<td>0</td>
</tr>
<tr>
<td>Dynamic pull cart</td>
<td>0</td>
</tr>
<tr>
<td>Dynamic push cart</td>
<td>0</td>
</tr>
<tr>
<td>Static knee pull</td>
<td>0</td>
</tr>
<tr>
<td>Static waist-level pull-up</td>
<td>0</td>
</tr>
<tr>
<td>Knuckle-to-elbow lift</td>
<td>0</td>
</tr>
<tr>
<td>Floor-to-waist lift</td>
<td>0</td>
</tr>
<tr>
<td>Hand grip</td>
<td>0</td>
</tr>
<tr>
<td>“Patient” lift recliner to bed</td>
<td>0</td>
</tr>
<tr>
<td>“Patient” lift bed to gurney</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: N = 13.
had only 61.6% of scores falling in the 4 to 6 range (closely matches strength demands for job tasks), this task is still likely to be predictive of ability to meet the job requirements.

Static push and pull wall also had only 61.6% of subjects rate this task in the 4 to 6 range, with two rating this task as underestimating the strength demands of the job. With the majority of subjects (65%) rating the task as closely matching or exceeding strength demands of the job, this task is still likely to be predictive of ability to safely perform the job.

The task with the lowest percentage of rating in the 4 to 6 range was static knee pull. Seven subjects rated this task as closely matching the strength demands of the job, whereas the other six rated this task as exceeding the strength demands. With no ratings in the 1 to 3 range, this task is also highly likely to correctly identify persons capable of performing the physical demands of the job. The static knee pull is an isometric task requiring maximal exertion. With the average score of 300 lb exceeding the requirement of 100 lb, it is understandable that almost half the subjects rated this task as exceeding the strength demands of the job. Although only four tasks met or exceeded the 80% criteria, we believe that overall the tasks selected will correctly identify persons who meet the physical demands of the job.

With regard to interrater reliability, all but two tests on the post-offer screening tool (static push wall, static pull wall) had good reliability (ICC ≥ .75). The less than satisfactory ICC results on the push wall test (.59) and static pull wall test (.22) may have been due to error in test design. Keyserling et al. (1978) showed static testing to be highly reliable if specific testing criteria are met. These criteria include consistent instruction about body posture, avoidance of jerky motions, and maintenance of the force for 4 to 5 sec. Although we instructed subjects to gradually pull or push and hold for 3 to 4 sec, we did not require them to adhere to the same position or assume the same stance. This may have resulted in variation between and within subjects in these areas. Although we reviewed general body mechanics techniques before screening, we purposely did not instruct subjects in a specific stance during actual testing of static push or pull wall so that we could assess their current strategy for performing these activities. This was in contrast to good reliability at the .76 and .77 level for the static waist-level pull-up and static knee pull tests. The adherence to specific postures for these tests likely account for the good interrater reliability. Because of the artificial nature of pushing and pulling statically against the wall and wide variation in scores, we would advocate standardizing this task in future testing situations and eliminating it if changes in instruction do not result in increased interrater reliability.

We chose the Chatillon dynamometer strain gauge to test waist-level pull-up, knee pull, and push and pull wall on the basis of cost, portability, and availability in most clinics. However, these features of the Chatillon violated the recommendations by Keyserling et al. (1978) for static testing. They recommended electronic force monitors that provide a timed average of force produced versus the peak torque that cable tensiometers or strain gauges, such as the Chatillon, measure. The latter method in their research was reliable at the .76 level for the arm pull strength test and .77 level for the knee pull strength test, which was replicated in this study. Further study of inter-reliability with the Chatillon strain gauge for the static push and pull wall tests should be done to determine whether standardizing posture and instructions will result in good reliability as indicated by Keyserling et al. and in this study for waist-level pull-up and knee pull. If not, the static push and pull wall tests may be eliminated from the screening tool, or a different measuring device may need to be considered.

In the work status section of our screening tool, few of the tasks could be analyzed because all 13 subjects passed all but two tests—knuckle-to-elbow lift and static waist-level pull-up. Because the subjects were already performing the job, this high pass rate prevented analysis but perhaps confirmed the validity of the criteria selected for each task. To increase the number of data sets for evaluation, we used the Kappa statistic. However, the Kappas calculated may be misleading because our sample size was so small and, therefore, all analyses were extremely sensitive to even small differences in values. In many ways, the Kappa analysis is a reflection of interrater reliability. Task score determines work status. Variation in task results, as observed with the static push wall (ICC = .59) results, in variation in pass or fail status, with a resultant low Kappa score of .24. Tasks with good interrater reliability, such as hand grip (ICC = .81), have a good Kappa score of .80.

Although we tested or screened all current employees, the small number in the group makes analysis and generalization of results to other lifting populations difficult.
We intend to collect additional data to continue to study both interrater reliability and predictive validity of this screening tool. The predictive validity will be determined by the ability of the tool to determine which employees passing the tests are less likely to be injured. The Appendix reflects the functional job tests portion of the final version of the post-offer screening tool.

Conclusion

The post-offer screening tool was well received by the patient support services staff and by the hospital’s self-insured workers’ compensation team. The evaluators and orderlies agreed that the lifting tasks closely replicated the expectations of an orderly’s job. Reliability of the raw scores on tests, except for the static push and pull wall, were considered satisfactory, and these tests will continue to be used unmodified. The static push and pull wall parts of the screen have been modified to include specific postural instruction. Interrater reliability will be reassessed in future studies to determine whether acceptable levels can be achieved or a different testing device or a different test needs to be selected.

Acknowledgments

We thank Brad Muir, supervisor of patient support services, and all his staff members for supporting our efforts to develop this screening tool. We also thank Fletcher Allen Health Care’s workers’ compensation team for its consistent respect and encouragement of our research interests.

Appendix

Functional Job Tests of the Revised Post-Offer
Screening Tool

The worker is evaluated on the following functional job tests on a pass–fail basis:

- Average static pull achieved (requirement = 90 lb)
- Average static push achieved (requirement = 90 lb)
- Dynamic pull achieved at 35 in. high (requirement = 90 lb)
- Dynamic push achieved at 35 in. high (requirement = 90 lb)
- Static knee-level pull-up achieved, using 22% of height from lordotic to neutral during lift, -2 = low back became kyphotic during the lift, -3 = excessive kyphosis during the lift (requirement = 0 to -2).
- Body mechanics score—stance: 0 = legs positioned shoulder width apart, one leg slightly ahead of the other; -1 = legs positioned shoulder width apart and parallel; -2 = feet positioned less than shoulder width apart; -3 = feet touching (requirement = 0 to -2).
- Heart rate achieved during lift-and-carry tests (requirement = < 80% mphr
- Transfer 165-lb "patient" with help of one other person from bed to gurney (bedside), gurney to bed (bedside), bed to gurney (gurney side), and gurney to bed (gurney side) (requirement = smooth transfer)
- Transfer 165-lb "patient" with help of one other person from bed to recliner (bedside), recliner to bed (bedside), bed to recliner (recliner side), and recliner to bed (recliner side) (requirement = smooth transfer)

References


Coming in January:

**Special Issue on Community Health**

- What Do We Need to Know?
- How Do We Pay for It?
- Documenting Future Planning With Clients
- Pediatric Lead Exposure Prevention

Turn to *AJOT* for the latest information on occupational therapy treatment modalities, aids and equipment, legal and social issues, education, and research.