Reliability, Minimal Detectable Change, and Normative Values for Tests of Upper Extremity Function and Power

Rodney J. Negrete,1 William J. Hanney,2 Morey J. Kolber,3 George J. Davies,4 Megan K. Ansley,5 Amanda B. McBride,5 and Amber L. Overstreet5

1Florida Hospital Celebration Health, Sports Medicine and Rehabilitation, Celebration, Florida; 2Department of Health Professions, University of Central Florida, Orlando, Florida; 3Department of Physical Therapy, Nova Southeastern University, Ft. Lauderdale, Florida; 4Department of Physical Therapy, Armstrong Atlantic State University, Savannah, Georgia; and 5Department of Physical Therapy, Armstrong Atlantic State University, Savannah, Georgia

ABSTRACT

Negrete, RJ, Hanney, WJ, Kolber, MJ, Davies, GJ, Ansley, MK, McBride, AB, and Overstreet, AL. Reliability, minimal detectable change, and normative values for tests of upper extremity function and power. J Strength Cond Res 24(12): 3318–3325, 2010—The purpose of this study was to examine the test–retest reliability, minimal detectable change (MDC), and determine normative values of 3 upper extremity (UE) tests of function and power. One hundred eighty participants, men (n = 69) and women (n = 111), were tested on 3 UE strength and power maneuvers in a multicenter study to determine baseline normative values. Forty-six subjects returned for a second day of testing within 5 days of the initial assessment for the reliability component of the investigation. Explosive power was assessed via a seated shot-put test for the dominant and nondominant arms. Relationships between the dominant and nondominant arms were also analyzed. A push-up and modified pull-up were performed to measure the amount of work performed in short (15-second) bursts of activity. The relationship between the push-up and modified pull-up was also determined. Analysis showed test–retest reliability for the modified pull-up, timed push-up, dominant single-arm seated shot-put tests, and nondominant single-arm seated shot-put tests to be intraclass correlation coefficient [13,1] 0.958, 0.989, 0.988, and 0.971, respectively. The MDC for both the push-up and modified pull-up was 2 repetitions. The MDCs for the shot put with the dominant arm and the nondominant arm were 17 and 18 in., respectively. The result of this study indicates that these field tests possess excellent reliability. Normative values have been identified, which require further validation. These tests demonstrate a practical and effective method to measure upper extremity functional power.

KEY WORDS: upper limb, functional tests, upper body strength and power

INTRODUCTION

Tests and measures play a vital role in the assessment of human performance in physical fitness, sport, and rehabilitation. Dynamic functional testing can assist the clinician or strength and conditioning professional to objectively evaluate progress, assess the effectiveness of the rehabilitation and training program, and help determine readiness to resume sport activity after injury. For functional testing to meet the diverse needs of the athletic and patient population, tests should include both weight-bearing (i.e., push-up, pull-up) and non–weight-bearing activities (i.e., throwing, shot putting).

Weight-bearing functional tests have been studied. Falsone et al. (6) proposed a 1-arm hop test to determine the effects of upper extremity (UE) dominance on test performance in American football and wrestling athletes with reported reliability values of 0.78 and 0.81, respectively. Rousch et al. (16) reported reference values for a closed kinetic chain (CKC) UE stability test to help develop goals and objectives for male collegiate baseball players recovering from arm injuries. Several researchers have published good reliability values for push-ups and modified pull-ups for determining upper body muscular function (3,11,13,15,19). Although many functional tests have been studied, they have not been described using short bursts of activity, which is generally more reflective of normal daily activity and sport participation. Also, the minimal detectable change (MDC) and relative relationships between pushing and pulling activities have not been reported.

Several articles describe throwing and putting as measures to determine performance capabilities of the UE
Jung and Dietz evaluated the muscle action potentials and efficiency during shot putting and ball throwing in right- and left-handed trained and untrained individuals (9). With shot putting, untrained individuals used arm extension as the primary force producing action and lacked the whole body coordination used by the trained subjects. However, the muscle actions of the arm and trunk were more coordinated with ball throwing. Collins and Hodges (2) reported that the Underkofler softball throw for distance produced a validity coefficient of 0.63 and a reliability coefficient of 0.95. Some functional tests including the softball throw and a 2 hand putting movements have been described in the literature (2,17). These tests are valuable; however, the softball throw test has only been studied in the dominant arm. Also, a 2-handed putting movement does not allow the examiner to evaluate differences between right and left UEs. Testing should allow the documentation of deficits or imbalances in strength and power characteristics. The relationship between the dominant and nondominant performances is necessary if future studies are to evaluate relative return to activity criterion. Unfortunately, many of these studies also do not report the MDC.

Therefore, the purpose of this study was to determine the test–retest reliability, minimum detectable change, and normative values for 3 new UE tests: a seated 1 arm shot put for distance, a timed modified pull-up test, and a timed push-up test.

**METHODS**

**Experimental Approach To The Problem**

To achieve our purpose, 3 tests of upper body strength and power were performed in a group of normal healthy adults in a random order: (a) one arm seated shot put with a 6-lb medicine ball, testing both the right and left UEs; (b) push-ups for 3 sets of 15 seconds with a 45-second rest period; and (c) modified pull-ups for 3 sets of 15 seconds with a 45-second rest period on a Smith Press machine.

A subgroup of 46 subjects were asked to repeat the push-up, pull-up, and shot-put tests between 3 and 5 days after initial testing to determine reliability coefficients because they have not been previously reported.

**Subjects**

One hundred and eighty healthy recreationally active adults (women, n = 111 and men, n = 69) between the ages of 18
and 45 years volunteered to participate in the multicenter study. Subjects were tested in 3 locations: Florida Hospital Celebration in Celebration, FL; the University of Central Florida (UCF) in Orlando, FL; and Armstrong Atlantic University in Savannah, GA. A subgroup of 46 subjects (women $n = 34$ and men $n = 12$) from the Florida locations returned between 3 and 5 days after initial testing. These data were used to investigate the intrareliability of these 3 testing techniques. The Institutional Review Boards of Florida Hospital, UCF, and Armstrong Atlantic University approved the methods and procedures used in this study. Informed consent was obtained from the subjects before testing. Each subject completed a health history form, and anyone with a history of UE injury, fracture, or surgery was excluded from participation. To be included in the study, subjects must be between the ages of 18 and 45. Subjects were excluded if they had a recent history of UE orthopedic disorders, unable to complete the tests as prescribed, or were unable to read, write, and communicate in English.

**Instruments**

Instruments used in this study included a measuring tape (Stanley Inc. Measuring, New Britain, CT, USA), digital stop watch (NuLine Products, Carlsbad, CA, USA), a smith machine (Prostar, Kansas City, MO, USA), and upper Body ergometer (Biodex Inc. Shirley, NY, USA).

**Procedures**

On obtaining consent, subjects were brought to a climate controlled indoor testing area. Subjects wore loose comfortable clothing that would not encumber physical movements. All subjects watched a short video demonstrating the technique for each of the 3 tests to be evaluated. After watching the video, subjects began a 5-minute warm-up of self-selected moderate intensity, on a seated Upper Body Ergometer (Biodex Inc.). This was followed by 3 minutes of general upper body stretching. Each subject was given a random ordered data sheet to take to each testing station for recording the results of each test.

**The Single Arm Seated Shot-Put Test**

The 1 arm seated shot put was performed with the subject...
position for women in which the subjects assumed the hands and knees position. Subjects were positioned prone with hands shoulder width apart with the trunk held in a rigid straight position. Push-ups were performed through the full range of motion and as quickly as possible. Subjects began in the “up” position with their elbows fully extended. On descending the body toward the ground, subjects flexed their elbows until the upper arm was parallel to the testing surface (Figure 2). The subject was instructed to limit head and trunk motion. A warm-up trial was completed before 3 maximal trials. The maximal effort trials were performed for 15 seconds each followed by 45 seconds of rest. The number of push-ups completed in each of the 3 15-second bouts was recorded. Verbal encouragement was consistently given to all subjects to ensure a high-intensity effort.

The Modified Pull-Up Test

The modified pull-up was performed on a Smith Press (Cybex, Medway, MA, USA) or similar apparatus, and a bench was used to position the subject’s feet or lower legs. Subjects assumed a supine position with their heels on a bench and using an overhand grip to clench the Smith Press bar. The bar was positioned just above the subject's reach when the subject was supine on the floor. Men performed the pull-up with their legs supported at their heels. Women had their entire lower legs supported just below the knees. A modified pull-up was performed through a full range of motion; the subjects started by hanging from the bar with arms fully extended and pulled up high enough so that the upper arms were parallel to the floor. The subjects then lowereded themselves back to the elbows fully extended position (Figure 3). The subjects were instructed to continuously keep a rigid straight trunk posture limiting trunk and head motion. Each subject completed as many pull-ups as possible in 15 seconds during the 3 maximal trials. A 45-second rest period was given between each maximal test bout. Verbal encouragement was given to ensure maximal intensity.

Statistical Analyses

All data were entered into SPSS version 12.0 program for statistical analysis. Descriptive statistics for all variables were tabulated and presented in Table 1. Test–retest reliability of
the timed push-up, timed modified pull-up, and the seated shot-put functional tests was evaluated using intraclass correlation coefficient (ICC) model 3, K. The mean value from each testing session was used for the analysis. Intraclass correlation coefficient values may be influenced by intersubject variability of scores, because a large ICC may be reported despite poor trial-to-trial consistency if the intersubject variability is too high (14,21). The SEM is not affected by intersubject variability (21); therefore, it was reported in conjunction with the ICCs using the formula: $\text{SEM} = \frac{SD}{\sqrt{1 - r}}$ (14). The MDC was calculated for the interrater measurements using the following formula: $\text{MDC}_{90} = 1.65 \times \text{SEM} \times \sqrt{2}$ to determine the magnitude of change that would exceed the threshold of measurement error at the 90% confidence level (8,14). The MDC values were rounded to the nearest point measured for each test. Bland and Altman plots were used to provide a visual illustration of the relationship between trials 1 and 2 from the reliability component of this investigation for each of the 3 tests. A ratio was determined between the push-up test and the modified pull-up test to assess symmetry of the anterior and posterior musculature. Lastly, a bilateral comparison of the dominant to nondominant seated shot-put performance was made.

**RESULTS**

All 180 subjects successfully completed the testing procedures. The descriptive analysis, including means and SDs for the entire group and both genders, is presented in Tables 1–3. Also, values were stratified by age for both men and women and can be found in Tables 4 and 5. A subset of 46 subjects returned for retesting to determine reliability coefficients for the timed pull-up, push-up, and seated single arm shot-put tests. Intraclass correlation coefficient ranged from 0.958 to 0.989 for these 3 tests suggesting excellent reliability (Table 6). Scatterplots used to provide a visual relationship between the variables are presented in Figures 4–7.

The MDC was calculated at the 90% confidence interval. The MDC for the push-up and modified pull-up tests was 2 repetitions. This represents that a patient or client would need to have a change of ≥2 push-ups or modified pull-ups to represent true change that exceeds measurement error. The MDC for the shot put is 17 in. for the dominant arm and 18 in. for the nondominant arm.

The ratio between contrasting muscle actions, push-up to modified pull-up, was established for the entire group of

---

**TABLE 5.** Women stratified by age (total $n = 111$).

<table>
<thead>
<tr>
<th>Age (range) and Test</th>
<th>Mean (range)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women age 18–22 ($n = 46$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU score</td>
<td>12.04 (6.33–18.00)</td>
<td>2.79</td>
</tr>
<tr>
<td>MPU score</td>
<td>4.07 (0.00–9.00)</td>
<td>2.57</td>
</tr>
<tr>
<td>SSP dominant score (in.)</td>
<td>60.25 (35.43–130.00)</td>
<td>22.98</td>
</tr>
<tr>
<td>SSP nondominant score (in.)</td>
<td>54.83 (32.43–126.00)</td>
<td>21.71</td>
</tr>
<tr>
<td><strong>Women age 23–27 ($n = 44$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU score</td>
<td>13.33 (6.67–20.00)</td>
<td>3.27</td>
</tr>
<tr>
<td>MPU score</td>
<td>4.63 (0.00–13.33)</td>
<td>3.09</td>
</tr>
<tr>
<td>SSP dominant score (in.)</td>
<td>85.92 (35.03–169.00)</td>
<td>38.95</td>
</tr>
<tr>
<td>SSP nondominant score (in.)</td>
<td>78.10 (12.33–165.33)</td>
<td>37.16</td>
</tr>
<tr>
<td><strong>Women age 28–32 ($n = 11$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU score</td>
<td>13.81 (8.00–20.33)</td>
<td>3.37</td>
</tr>
<tr>
<td>MPU score</td>
<td>5.21 (0.00–11.67)</td>
<td>4.16</td>
</tr>
<tr>
<td>SSP dominant score (in.)</td>
<td>99.35 (51.57–146.67)</td>
<td>31.90</td>
</tr>
<tr>
<td>SSP nondominant score (in.)</td>
<td>91.74 (49.20–144.00)</td>
<td>30.86</td>
</tr>
<tr>
<td><strong>Women age 33–37 ($n = 3$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU score</td>
<td>14.89 (13.00–17.00)</td>
<td>2.01</td>
</tr>
<tr>
<td>MPU score</td>
<td>9.44 (7.67–10.67)</td>
<td>1.58</td>
</tr>
<tr>
<td>SSP dominant score (in.)</td>
<td>88.98 (44.60–117.00)</td>
<td>38.87</td>
</tr>
<tr>
<td>SSP nondominant score (in.)</td>
<td>77.10 (36.63–106.00)</td>
<td>36.10</td>
</tr>
<tr>
<td><strong>Women age 38–42 ($n = 7$)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PU score</td>
<td>10.90 (4.33–15.00)</td>
<td>3.81</td>
</tr>
<tr>
<td>MPU score</td>
<td>6.14 (1.67–12.00)</td>
<td>4.09</td>
</tr>
<tr>
<td>SSP dominant score (in.)</td>
<td>78.82 (39.10–126.33)</td>
<td>35.58</td>
</tr>
<tr>
<td>SSP nondominant score (in.)</td>
<td>75.08 (39.63–127.33)</td>
<td>35.53</td>
</tr>
</tbody>
</table>

**TABLE 6.** Test–retest reliability intraclass correlation coefficients ICC (3,K) ($n = 46$).*

<table>
<thead>
<tr>
<th>Test</th>
<th>ICC (3,K) (95% CI)</th>
<th>SEM</th>
<th>MDC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Push-up</td>
<td>0.958 (0.908–0.979)</td>
<td>1 Rep</td>
<td>2 Reps</td>
</tr>
<tr>
<td>Modified pull-up</td>
<td>0.989 (0.980–0.994)</td>
<td>1 Rep</td>
<td>2 Reps</td>
</tr>
<tr>
<td>Shot-put dom. UE</td>
<td>0.988 (0.978–0.993)</td>
<td>7 in.</td>
<td>17 in.</td>
</tr>
<tr>
<td>Shot-put nondom. UE</td>
<td>0.971 (0.947–0.984)</td>
<td>8 in.</td>
<td>18 in.</td>
</tr>
</tbody>
</table>

*UE = upper extremity; MDC = minimal detectable change; ICC = intraclass correlation coefficient.
subjects and the men and women separately. The results from the descriptive statistics (mean values) were used to generate this ratio for each group. The ratio found for the entire group of subjects was 2.01:1. The push-up to modified pull-up ratio for the male subjects was 1.57:1. The women demonstrated a ratio of 2.72:1 for the pushing to pulling musculature. These results suggest for our group of normal recreationally active subjects, the upper body musculature used for pushing are 1½–2½ times stronger than the musculature involved in pulling.

The bilateral comparison of the seated single arm shot-put maneuver demonstrated a 9% difference between the dominant and nondominant UEs for both the men and women. The mean scores for the dominant and nondominant UEs were used for this calculation.

**DISCUSSION**

As functional testing procedures are proposed, the use of normative data and reliability data are imperative for the evaluation and analysis of patient or athletic performance. The information obtained allows the clinician to assess the balance of the agonist and antagonist muscle performance and the symmetry between contralateral limbs. The simple field tests proposed in this project may be a valuable resource.
Upper Extremity Function and Power

for rehab and strength and conditioning professionals. The descriptive statistics can provide important guidelines during the rehab or training process. When using the normative data, the clinicians must consider the population being tested and assessed, to make the appropriate comparison of results.

Published bilateral comparisons with strength and power ratios using isokinetic data from isolated UE joint testing are routinely analyzed by clinicians (1). The current functional testing data can allow clinicians or strength and conditioning professionals, without access to isokinetic devices, to perform meaningful assessments of upper body function and power.

The results from this study indicate the timed push-up, modified pull-up, and seated single arm shot put, demonstrate significant reliability. Concurring with our results, reliability for similar field tests of UE functional performance have been reported (3,7,11,13,15,19).

In addition to the reproducibility of a test, the clinician conducting reassessments must consider the issue of change and whether observed differences actually reflect true change. Calculating the MDC for the 3 tests gives the clinician more information to assist with the analysis of retest results during a comprehensive rehabilitation or training program. With this information, the clinician can be 90% confident that changes in the MDC represent true improvement in the patient’s test score and exceed measurement error (14). Consideration of the MDC is important when monitoring the progress of athletes or patients because intertrial variation may incorrectly suggest a change that has not exceeded the threshold of error. This calculation can play an important role in goal setting for both the clinician and strength and conditioning professional.

Specificity principles are especially important when developing an UE exercise program in rehabilitation and physical conditioning. Selection should be based on the mechanics and loading that achieve the appropriate muscle activity (4). Non-weight-bearing exercises exhibit speed involving more free movement and less stabilization, whereas weight-bearing exercise requires more stabilization and less acceleration (4).

Functional rehabilitation should reintroduce and reproduce the positions and forces that will occur during the activity of daily living (ADL) or sport the patient is returning to.

Previous researchers advocate the use of both open kinetic chain (OKC) and CKC exercises throughout the UE rehabilitation process (12,18,22). Wilk et al. emphasized a continuum of functional movement patterns based on proper patient selection and the ultimate goal of the rehab program (22). Ellenbecker and Davies introduced the concept of a functional testing algorithm with emphasis on isokinetic testing. They outlined an integrated OKC isokinetic and functional testing method for patients with UE injury (1).

It has been theorized that a balance in strength should exist between opposing muscle or muscle actions to help avoid injury and enhance performance (1). Based on this theory, Baker and Newton (1) evaluated the ability of 2 field tests to assess the strength balances between contrasting muscle actions of the shoulder girdle. They stated the ratio between upper body pressing and upper body pulling should be approximately 100% if the 2 different muscle groups required for each action were addressed equally in training. Using 42 highly skilled athletes for test subjects, Baker and Newton had each subject perform 1 repetition maximum (1RM) bench press for assessing the strength of upper body pressing and also a 1RM body weight pull-up for assessing the strength of upper body pulling. The strength ratio was found to be 97.7% and the correlation, 0.81 for the 2 tests (1). In our study, 2 field tests including the timed push-up and modified pull-up were performed to find the ratio between these opposing muscle groups of the upper body. This study had markedly different results as in the Baker and Newton study. The ratios calculated between the push-up and modified pull-up ranged between 1.57 and 2.72:1. The muscles involved in the pulling movement were 64 and 37% of the pushing musculature for the men and women, respectively.

Differences between these 2 studies are the fitness level of the subjects and the different test procedures. Baker and Newton used highly skilled male semiprofessional and professional athletes in their study, and the current project included male and female subjects who were recreationally active. The study by Baker and Newton used the 1RM bench press and vertical body weight pull-up. Our study used a standard body weight push-up and the horizontal body weight modified pull-up test. These tests are somewhat similar in that they assess muscle performance of pushing and pulling actions, but it may be difficult to make a direct comparison.

Ellenbecker and Davies recommended increasing the isokinetic shoulder external rotation (ER) to internal rotation strength ratio of 66–76% to bias the ER strength for patients with shoulder pathology (5). This ER bias would help create a “posterior dominant” shoulder to increase dynamic stability and aid in preventing reinjury in throwing and racquet-sport athletes (5). The pushing to pulling ratio calculated in our study may indicate a similar scenario for the recommendation of biasing the ratio in favor of the pulling musculature to create a more symmetrical strength and power relationship.

Lastly, a bilateral comparison for the seated single arm shot-put test revealed a 9% difference between the dominant and nondominant scores for both men and women. These results are in line with Jung and Dietz who reported an 8% dominant to nondominant difference among untrained persons performing a standing shot-put maneuver (9). This finding may have significance while evaluating and rehabsilitating a nondominant or nonthrowing arm after injury. For the patient or athlete to progress to higher levels of activity, they would need to have a <10% difference in the scores between each arm.
In applying these results, there is a need to develop a series of functional performance tests that can be incorporated into a training or rehabilitation program to assess the efficacy of the training, aid in the progression through the rehabilitation process, and help determine readiness to return to ADL or sport.

**Practical Applications**
A series of functional tests assessing different aspects of neuromuscular performance, much like the tests in our study, would be superior to any one test alone for evaluating upper body strength and power. Full return to ADL or sport should be based on the appropriate performance of all tests in the sequence. The current study may assist in the development of a serial testing paradigm for use during the training and rehabilitation of patients with UE pathology. The findings of this study are limited to a healthy nonpatient population. Future research is needed to determine the relationships between the different tests and examine the ability to predict performance of a functional task such as throwing.

**Acknowledgments**
The authors would like to acknowledge Michelle Dolske, PhD, Katie Welsh, MPT, Tammy Stone, MPT, and Bryan Riemann, PhD, ATC, for their assistance with this project.

**References**