ORIGINAL RESEARCH

RELIABILITY OF STRENGTH AND PERFORMANCE TESTING MEASURES AND THEIR ABILITY TO DIFFERENTIATE PERSONS WITH AND WITHOUT SHOULDER SYMPTOMS

Aaron Sciascia, MS, ATC, PES[1]
Tim Uhl, PhD, ATC, PT, FNATA[1,2]

ABSTRACT

Background: Upper extremity physical performance measures exist but none have been universally accepted as the primary means of gauging readiness to return to activity following rehabilitation. Few reports have described reliability and/or differences in outcome with physical performance measures between individuals with and without shoulder symptoms.

Hypotheses/Purpose: The purpose of this study was to establish the reliability of traditional upper extremity strength testing and the CKCUEST in persons with and without shoulder symptoms as well as to determine if the testing maneuvers could discriminate between individuals with and without shoulder symptoms. The authors hypothesized that strength and physical performance testing would have excellent test/re-test reliability for individuals with and without shoulder symptoms and that the physical performance maneuver would be able to discriminate between individuals with and without shoulder symptoms.

Methods: Male and female subjects 18-50 years of age were recruited for testing. Subjects were screened and placed into groups based on the presence (Symptomatic Group) or absence of shoulder symptoms (Asymptomatic Group). Each subject performed an isometric strength task, a task designed to estimate 1-repetition maximum (RM) lifting in the plane of the scapula, and the closed kinetic chain upper extremity stability test (CKCUEST) during two sessions 7-10 days apart. Test/re-test reliability was calculated for all three tasks. Independent t-tests were utilized for between group comparisons to determine if a performance task could discriminate between persons with and without shoulder symptoms.

Results: Thirty-six subjects (18/group) completed both sessions. Test/re-test reliability for each task was excellent for both groups (intraclass correlations ≥.85 for all tasks). Neither strength task could discriminate between subjects in either group. Subjects with shoulder symptoms had 3% less touches per kilogram of body weight on the CKCUEST compared to subjects without shoulder symptoms but this was not statistically significantly different (p = .064).

Conclusions: The excellent test/re-test reliability has now been expanded to include individuals with various reasons for shoulder symptoms. Traditional strength testing does not appear to be the ideal assessment method for making discharge and/or return to activity decisions due to the inability to discriminate between the groups. The CKCUEST could be utilized to determine readiness for activity as it was trending towards being discriminatory between known groups.

Level of Evidence: Basic Science Reliability Study, Level 3

Keywords: CKCUEST, Physical Performance Testing, Reliability, Strength Testing

[1] Lexington Clinic, Shoulder Center of Kentucky, Lexington, KY, USA
[2] University of Kentucky, Division of Athletic Training, Lexington, KY, USA

CORRESPONDING AUTHOR

Aaron Sciascia
1221 South Broadway
Lexington, KY 40504
(859) 258-8506 office
(859) 258-8562 fax
E-mail: ascia@lexclin.com

INTRODUCTION

Functional testing is a mechanism that incorporates task or sports specific maneuvers into the traditional rehabilitation environment allowing the clinician to qualitatively and/or quantitatively assess a person's performance of a specific task. The testing provides the clinician with an observable depiction of dynamic physical function and/or a quantifiable result (time, strength, endurance, etc.), allowing judgments to be made regarding the successful resolution of impairments and/or the safe return to the sport of interest based on the performance of the task(s).[1] However, a recent report suggested the label "physical performance measure" is a more proper descriptor of such testing maneuvers because most maneuvers only assess one aspect of function (the physical aspect); therefore, broadly labeling a test as a measure "function" may not be accurate.[2]

Physical performance measures specific to the upper extremity exist but none have been universally accepted as the primary means of gauging readiness to return to activity following the completion of musculoskeletal rehabilitation. Unlike maneuvers described for the lower extremity which have reported injury prediction and performance value (in particular, the single leg hop and step-down maneuvers),[3-5] the upper extremity does not have a single best test to utilize for performance assessment likely due to the variation in the demands of different sports on the upper extremity. For example, the demands of an American football lineman require both closed and open chain arm movements which differ from the demands on a quarterback who is required to perform primarily open chain movements with the overhead throwing motion. Due to the absence of a gold standard of assessment for upper extremity physical performance, clinicians will often utilize some variation of strength testing as the post-intervention metric because strength is a basic physiological component of physical task performance permitting fundamental tasks to be executed (such is the rationale for routinely conducting manual muscle testing procedures during clinical examinations and throughout rehabilitation). Strength measures for the upper extremity are employed in the clinical setting to determine side to side differences between involved and non-involved limbs. The strength measures can be reliably implemented,[6-10] possibly adding justification for their routine use. However, they have not been examined in the literature for value regarding return to activity. Furthermore, as important as strength testing is for identifying potential impairments and assessing progress in the secure rehabilitation setting, it has been recognized that single component physiological measurements of strength, mobility, endurance, or pain do not necessarily translate to a patient's ability to perform a highly skilled dynamic task.[1,11]

Strength measures are possibly utilized as a rehabilitation progression or discharge metric because there is a lack of a gold standard for assessing upper extremity performance. Numerous physical performance measures for the upper extremity have been described in the literature. However, most maneuvers are either time consuming to implement, complex to perform, or are applicable to specific sports and do not translate across a variety of activities.[12,13] One test which could potentially overcome the implementation obstacles and may be applicable to a variety of sports would be the closed kinetic chain upper extremity stability test (CKCUEST).[14] The maneuver is performed in a weight-bearing position requiring the individual to alternately lift and horizontally adduct one hand, touching the opposite hand in a repetitive sequence while maintaining a weight-bearing position similar to the extended position of a push-up. The CKCUEST has been found to be reliable in asymptomatic subjects and subjects with subacromial impingement syndrome with test/re-test reliability being reported as excellent (ICC≥0.91).[14,15] Recently, Pontillo et al identified an association between decreased pre-season performance on the CKCUEST and the occurrence of shoulder injury during the season.[16] Athletes who sustained an in-season injury had a significantly lower number of touches at the beginning of the season during the CKCUEST compared to the athletes who did not sustain injury. These findings provide evidence that there may be a testing maneuver which can identify a reduction in physiological function which places individuals at risk for future injury. However, while the ability of the CKCUEST to predict injury is being studied, there is limited information reporting the discriminatory ability of the CKCUEST for persons currently with or without shoulder symptoms.[15]

Due to the limited reports describing reliability and differences in outcome with physical performance measures between individuals with and without shoulder symptoms, current clinical decision making regarding readiness to return to activity following rehabilitation has a marked shortcoming where return to activity decisions may be based on results from tests which cannot discern differences between patients with and without shoulder symptoms. Therefore, the purpose of this study was to establish the reliability of traditional upper extremity strength testing and the CKCUEST in persons with and without shoulder symptoms as well as to determine if the testing maneuvers could discriminate between individuals with and without shoulder symptoms. The hypotheses were: 1) strength testing and the CKCUEST would have excellent test/re-test reliability for both testing groups, and 2) asymptomatic individuals will demonstrate better performance on the CKCUEST than symptomatic individuals.

METHODS

Subjects
Male and female subjects between 18-50 years of age were recruited for testing. After reading and signing an IRB approved consent form, subjects were screened for placement into one of two groups based on the presence (Symptomatic Group [SG]) or absence (Asymptomatic Group [AG]) of shoulder symptoms. Presence of pain was determined via the completion of a numeric pain rating scale (NPRS), measured 0-10 with 0 = "no pain at all" and 10 = "worst pain ever felt". In addition to the NPRS, current physical functional status was assessed with the American Shoulder and Elbow Surgeons Score (ASES) where the patient reported level of perceived function from 0-100, with 0 = "not able to function" and 100 = "best function possible".[17] Inclusion criteria for the AG required a subject to score 90 or above on the ASES, report no pain or pain no greater than 2/10 on the NPRS, have no limitations in range of motion, no point tenderness to palpation of the shoulder complex, and no positive examination findings for tissue derangement or other conditions on the screening clinical examination. Subjects with pain ratings ≤2/10 were included in the AG if the ASES function component was unaffected by the presence of pain (a score of 50 on the function com-

ponent had to be reported) and the screening would suggest no injury was present. Inclusion criteria for the SG included the presence of pain greater than or equal to 3/10 on the NPRS and an ASES score below 89. Subjects could have limited range of motion but were required to demonstrate active elevation to at least 90°. Subjects may or may not have had point tenderness over their shoulder region and at least one positive clinical examination finding indicative of tissue derangement and/or other conditions (i.e. tendonitis, subacromial impingement, etc.). Subjects were excluded from this study if they had pain ≥8/10 on the NPRS and an ASES score ≤20. Subjects with pain ratings ≥8/10 were excluded out of concern for possibly advancing any possible underlying tissue lesion or exacerbating their symptoms to the point where the subjects would withdraw from the study. Subjects were also excluded if they had a current disease, illness, or condition medically disqualifying the individual from participating in vigorous activity, if he or she was currently participating in a post-surgical rehabilitation program, demonstrated signs of cervical radiculopathy,[18] or had shoulder and/or neck surgery in the past 24 months. Using a previously published sample size estimation method for reliability studies,[19] the target enrollment for a test/re-test design was 36 total subjects, which is based on an α of 0.05 and β of 0.20. This includes an assumption of a minimum acceptance of 0.70 intraclass correlation for reliability and upper limit acceptance of ≥0.90 reliability.

Procedure
Demographic information including age, sex, race, height, weight, and history of injury was recorded (Table 1). Following obtainment of the demographic information, a standard shoulder examination was conducted on both shoulders by a single certified athletic trainer with 15 years of clinical experience and expertise in shoulder evaluation and management to verify group assignment. The examination included palpation of anatomical structures of the glenohumeral joint and scapula, visual inspection of range of motion, manual muscle testing (break testing without a hand-held dynamometer), and special testing for the confirmation of presence or absence of tissue injury/irritation. The special tests included maneuvers with established acceptable

Table 1. Subject Demographics for Asymptomatic and Symptomatic Groups (N = 36)		
	Asymptomatic Group (Mean ± SD)	Symptomatic Group (Mean ± SD)
Total Subjects	18	18
Age	29 ± 7 years	30 ± 8 years
Height	171 ± 7 cm	172 ± 12 cm
Weight	71 ± 14 kg	76 ± 15 kg
Sex		
Male	9 (44%)	9 (50%)
Female	10 (56%)	9 (50%)
ASES Pain Score		
Dominant/Involved Arm*	49 ± 2	29 ± 9†
Non-Dominant/Non-Involved Arm*	50 ± 0	49 ± 4
ASES Function Score		
Dominant/Involved Arm*	50 ± 1	38 ± 10†
Non-Dominant/Non-Involved Arm*	49 ± 1	48 ± 3
ASES Total Score		
Dominant/Involved Arm*	99 ± 2	67 ± 15†
Non-Dominant/Non-Involved Arm*	99 ± 1	98 ± 6

ASES=American Shoulder and Elbow Surgeons; SD=standard deviation; cm=centimeters; kg=kilograms

*Involved and non-involved arm for the Symptomatic Group paralleled dominant and non-dominant arm for the Asymptomatic Group

†Symptomatic Group scores significantly less than Asymptomatic Group scores P<.001

clinical utility and/or those the research team has utilized in clinical practice and have become proficient at employing.[20-23] The maneuvers included: Spurling's test, Distraction, and Median Nerve Upper Limb Tension Test for cervical involvement; Painful Arc, Drop Arm Test, External Rotation and Internal Rotation Lag Signs, and Lift-Off Test for rotator cuff involvement; Hawkins-Kennedy and Neer Impingement Signs; Cross Body Adduction Test for AC Joint involvement; Modified Dynamic Labral Shear and Active Compression Tests for Labral involvement; Speed's and Upper Cut Tests for Biceps involvement; and the Scapula Dyskinesis Test for observational detection of altered scapular motion. No specific diagnosis was attempted to be made, rather these tests were used only to classify patients into either the AG or SG. Following the screening and group allocation, strength testing and the CKCUEST were administered in a randomized sequence.

Isometric Strength Testing of Shoulder Muscles[6]

In order to include a maneuver designed to assess strength that is commonly utilized in clinical practice, isometric shoulder elevation in the plane of the scapula was selected. Each subject was positioned standing with elevation of a single arm to 90° and 30° of horizontal abduction to place the arm in the scapular plane. A hand-held dynamometer (Lafayette Instrument Company, Lafayette, IN) was placed centered on the dorsal aspect of the forearm, half the distance between the distal radius and ulna and the elbow, parallel to the ground. The examiner resisted elevation in the scapular plane with the forearm in neutral and slight supination. In order to standardize the arm position for all subjects, a strap was placed through the handle of the dynamometer and secured to the bottom of a door via a bracket. The strap was adjusted for each subject to account for subject height and arm position

as described above (Figure 1). The limb to begin with was randomized. Each trial lasted five seconds with each subject instructed to give maximal effort. A minimum of 20 seconds rest was provided between each trial. Each limb was tested three times in alternating sequence (i.e. right, left, right, left, etc.) to facilitate strength recovery. The force output was recorded for each trial, with the average of three trials for each arm recorded in kilograms for data processing.

1-Repetition Maximum (RM) Estimate of Scapular Plane Elevation (scaption) Strength Test for the Upper Extremity[24]

The 1-RM scaption maneuver began with the subject standing and arms resting at the side of the body. Each subject was asked to self-select a free weight that he or she perceived as the maximal amount of weight which could be lifted no more than 10 times to shoulder level. The subjects were permitted to sample various weights in order to assist in selecting the most appropriate load with no more than three practice repetitions permitted per each weight sampled. Each subject was asked to elevate the arm up to 90° of elevation with the forearm in neutral (thumb up) which was controlled by a barrier placed at the appropriate height (Figure 2). The arm was required to maintain elbow extension during movement throughout the trial. A digital metronome was utilized and set at 47 beats per minute to control the pace of the arm. The pace of 47 beats per minute was established during pilot testing as it was the pace that subjects could accurately and comfortably maintain fluid arm motion. The arm was placed in the plane of the scap-

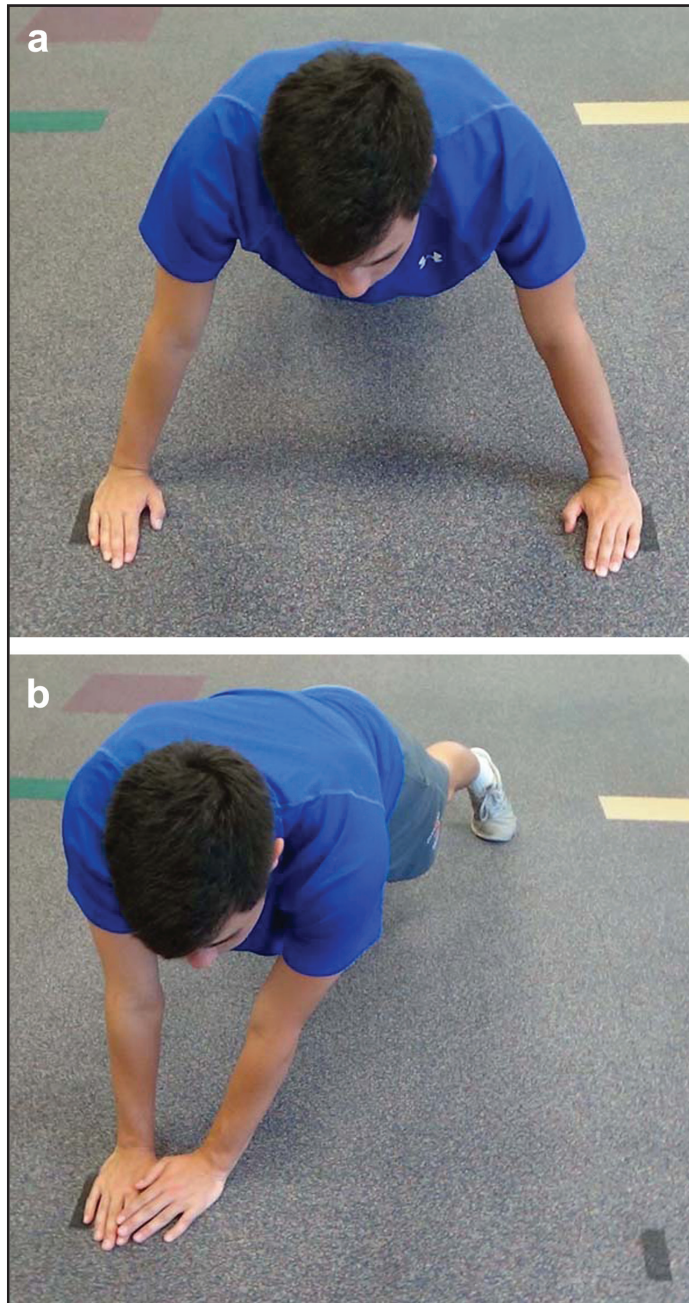**Figure 2.** *One repetition maximum elevation task.*

ula with the subject performing 10 repetitions of scapular plane elevation. Each arm was tested separately for one trial. The test was discontinued if the subject could not perform elevation to the required target or if the subject reported pain and/or self-limited him or herself. The subject was stopped by the investigator if observable compensations of the trunk and body were used to lift the weight. The number of repetitions completed and weight lifted were used to estimate 1-RM via the calculation described by Brzycki[24]: Estimated 1-RM = weight lifted/1.0278-0.0278x (where x = the number of repetitions performed).[24] This task was selected because it was considered to be more functional and more challenging than traditional manual muscle testing due to its dynamic design and it allows for the incorporation of an individual's perception of task performance.

Closed Kinetic Chain Upper Extremity Stability Test (CKCUEST)[14]

The CKCUEST was selected for inclusion in this study because it can be implemented in any clinical setting

Figure 1. *Isometric strength testing.*

and is an upper extremity-specific physical performance measure that is not designed exclusively for overhead athletes. Two pieces of tape were placed on the floor parallel to each other 36 inches apart. The subject began in the elevated position similar to a standard push-up with one hand on each piece of tape, the body straight and parallel to the floor, and feet no greater than shoulder width apart (Figure 3a). When the test began, the subject removed one hand from the floor, touched the opposing hand on the opposite line and then replaced the hand on the original line (Figure 3b). The subject then removed the other hand from the floor, touching the opposite line and returning it to the original line. A single test consisted of alternating touches for 15 seconds. Subjects were instructed to attempt as many touches as possible during the 15 seconds while maintaining proper push-up form. Each subject was permitted to perform a submaximal trial prior to performing the maximal effort attempts in order to become familiar with the test demands. Subjects performed two maximal effort trials each lasting 15 seconds with 45 seconds of rest in between the trials. Verbal cues were provided by a member of the research team if a subject was not maintaining proper body position during the testing. In the event a subject did not return the hand to the tape or did not touch the opposing hand during a repetition, the repetition was not recorded. The average number of touches between the two trials was calculated and recorded.

Figure 3. *Closed kinetic chain upper extremity test beginning position (figure a), and active position (figure b).*

Data Analysis

Descriptive statistics for all subjects were calculated with means and standard deviations reported for continuous variables and frequencies and percentages reported for categorical variables. The results from both the isometric strength task and the 1-RM estimate task were recorded in pounds then converted to kilograms. The results from all three tests were normalized to each subject by dividing each individual's test result by the body weight in kilograms prior to performing any comparative analyses in order to account for anthropometric differences between subjects. Since this study included subjects with shoulder symptoms, normalization to body weight was preferred over subject height in order to account for joint loading which could be a potential confounding variable within individuals who may or may not had compromised shoulder anatomy. The intraclass correlation coefficient (ICC), standard error of measurement (SEM), and minimal detectable change at the 90% confidence level (MDC_{90}) were calculated for all three tasks. In order to examine the inter-session reliability of the maneuvers, subjects were retested following the identical protocol no less than seven days and no more than 10 days after the initial testing session. ICC values were calculated using the two-way random effects model with absolute agreement [ICC

(2,1)].[25,26] An ICC greater than 0.75 was interpreted as excellent while values between 0.40–0.75 were considered fair to good and < 0.40 was considered poor.[27] Prior to determining if any test could discriminate between subjects with and without shoulder symptoms, a formal test of normality was initially utilized for each dependent variable. The Shapiro-Wilk test for normality was employed revealing the variables were normally distributed which allowed independent t-tests to be utilized for between group comparisons. Statistical significance was set at $\alpha = p < 0.05$. All statistical calculations were performed using STATA/IC (version 13.1 for Windows, StataCorp, LP, College Station, TX).

RESULTS

Subjects

A total of 36 subjects completed both testing sessions with 18 subjects in each group thus satisfying the sample size estimate (AG: females 10, males 8; SG: females 9, males 9). A summary of the descriptive statistics for all subjects is reported in Table 1. Per the ASES self-reported questionnaire, the SG had an average ASES score of 67 ± 15 points out of a possible 100 points. The ASES pain score, function score, and total ASES score were all significantly less for the SG compared to the AG ($p < .001$). The screening revealed the following possible diagnoses, for descriptive purposes: labral injury (7 subjects),

rotator cuff tendonitis/impingement (7), biceps tendonitis (1), rotator cuff injury (1), multidirectional instability (1), and concurrent rotator cuff and labral injury (1).

Reliability

The test/re-test reliability for all three tasks was considered excellent for both groups with the AG (CKCUEST = 0.85, isometric task = 0.98 for each arm, 1-RM estimate = 0.94 for the dominant arm and 0.96 for the non-dominant) and SG (CKCUEST = 0.86, isometric task = 0.97 involved arm and 0.95 for non-involved arm, 1-RM estimate = 0.93 for each arm) having similar ICC values. The SEM and MDC$_{90}$ values for each test and group are presented in Table 2.

Discriminatory Analysis

Across all tests, prior to normalizing the test results to body weight, there were no statistically significant differences in the performance of any task between the AG and SG. After applying the body weight correction, neither the isometric task for the dominant/involved arm ($p = .89$) or for the non-dominant/non-involved arm ($p = .99$), nor the 1-RM estimate for the dominant/involved arm ($p = .36$) or for the non-dominant/non-involved arm ($p = .17$) could discriminate between subjects with or without shoulder symptoms (Table 3). Subjects with shoulder symptoms had 3% less touches per kilogram of body

Table 2. *Reliability Results for Asymptomatic and Symptomatic Groups*

Asymptomatic Group n=18	Isometric Task Dominant	Scaption Non-Dominant	CKCUEST	1-RM Estimate Dominant	1-RM Estimate Non-Dominant
ICC	0.98	0.98	0.85	0.96	0.94
95% CI Lower Bound	0.95	0.95	0.42	0.89	0.82
95% CI Upper Bound	0.99	0.99	0.95	0.99	0.98
Mean	12kg	12kg	22 touches	8kg	8kg
SD	4kg	4kg	5 touches	4kg	4kg
SEM	1kg	1kg	2 touches	1kg	1kg
MDC$_{90}$	1kg	1kg	4 touches	2kg	2kg
Symptomatic Group n=18	Isometric Task Involved	Scaption Non-Involved	CKCUEST	1-RM Estimate Involved	1-RM Estimate Non-Involved
ICC	0.97	0.95	0.86	0.93	0.93
95% CI Lower Bound	0.91	0.86	0.11	0.81	0.83
95% CI Upper Bound	0.99	0.98	0.96	0.98	0.98
Mean	12kg	13kg	22 touches	10kg	10kg
SD	4kg	4kg	5 touches	4kg	5kg
SEM	1kg	1kg	2 touches	1kg	1kg
MDC$_{90}$	2kg	2kg	4 touches	3kg	3kg

CKCUEST=closed kinetic chain upper extremity stability test; RM=repetition max; ICC=intraclass correlation coefficient; 95% CI=95% confidence interval; SD=standard deviation; SEM=standard error of measurement; MDC$_{90}$=minimal detectable change at 90% confidence level; kg=kilogram

Table 3. *Task Results Normalized to Body Weight (in kilograms) for Asymptomatic and Symptomatic Groups*

	Asymptomatic Group (n=18)	Symptomatic Group (n=18)	P-Value (95% CI)
Isometric Task Dominant/Involved	16% ± 4%	16% ± 3%	P=0.89 (15%, 17%)
Isometric Task Non-Dominant/Non-Involved	16% ± 4%	16% ± 4%	P=0.99 (15%, 17%)
CKCUEST	32% ± 7%	29% ± 6%	P=0.064 (29%, 32%)
1-RM Estimate Dominant/Involved	12% ± 4%	12% ± 4%	P=0.36 (11%, 13%)
1-RM Estimate Non-Dominant/Non-Involved	12% ± 4%	13% ± 4%	P=0.17 (11%, 13%)
CKCUEST=closed kinetic chain upper extremity stability test; RM=repetition max; 95% CI=95% Confidence Interval			

weight on the CKCUEST compared to subjects without shoulder symptoms which was trending towards statistical significance (p = .064).

Discussion

Clinical decision making for determining the successful completion of a rehabilitation program and thus safe return to activity can be challenging. Clinicians have many tools at their disposal to assist them in making discharge and return to activity decisions, with most clinicians opting to use some variation of a strength measure as a means of determining cessation of treatment or activity readiness. With the understanding that strength measures may not serve as an exclusive surrogate for making discharge and/or return to activity decisions, physical performance measures were developed and have been advocated as more challenging options to determine readiness for activity.[1,2,12,14,28] Examining both traditional strength measures and an upper extremity-specific physical performance measure in this study led to one of the two study hypotheses being supported with all tasks having excellent test/re-test reliability in both subjects with and without shoulder symptoms. The hypothesis that the CKCUEST could distinguish between individuals with and without shoulder symptoms was partially rejected as the evidence was trending towards supporting the hypothesis (p = .064) but was by definition (p < .05) not statistically different between the performances of the two subject groups.

All tests could be reliably performed over multiple days amongst individuals with and without shoulder symptoms. Both the isometric strength task and 1-RM estimate had excellent test/re-test reliability with ICC values being ≥0.93. These findings parallel previous studies which have also examined the test/re-test reliability of clinical strength testing of the shoulder.[6] The ICC values in this study for the CKCUEST were slightly lower (ICC = 0.85) but still similar to the values reported in the original reliability study (ICC = 0.93) and a study involving subjects with subacromial impingement syndrome (ICC≥0.91).[14,15] While the original report examining the reliability of the CKCUEST exclusively focused on the outcome of task performance in asymptomatic individuals, the current study chose to also include persons with current complaints of shoulder pain in order to provide a clearer picture of the upper extremity assessment measure's clinical value. Additionally, the original report did not provide SEM and MDC_{90} values. However, calculation of these metrics could be performed from the original results, offering an SEM of 0.5 touch and MDC_{90} of 1.2 touches.[14] The current study's SEM of 2 touches and MDC_{90} of 4 touches were larger than both the original report[14] and the report involving subjects with subacromial impingement syndrome.[15] The difference in SEM and MDC_{90} values was likely due to the performance of one less trial in the current study. The decision to utilize one less trial was based on the methodology

from a recent study[16] and also to lessen the effects of fatigue during testing since multiple tasks were employed.

An important finding from the current study is the lack of a side-to-side difference in the performance of the isometric strength task in the SG. Clinicians routinely utilize manual muscle testing during initial evaluation procedures or periodically throughout rehabilitation to determine if strength deficits exist or if strength imbalances are resolving. Manual muscle testing was originally employed to assess the strength ability of patients with paralytic conditions.[29] In conditions where neurological integrity is compromised, manual muscle testing may have clinical value. However, manual muscle testing may not have robust value as an individual evaluation tool for musculoskeletal injury with an absence of nerve injury or neurological dysfunction. The SG demonstrated no side-to-side difference which can be explained in part as no neurological involvement was reported by these participants. Furthermore, although the subjects in the SG reported a pain level resulting in a significantly lower pain score on the ASES pain score compared to the subjects in the AG, the subjects with painful shoulders were not actively being treated for their shoulder pain suggesting that pain level is not always equitable to perceived or demonstrated dysfunction. Therefore, it is important to not assume weakness will routinely coincide with the presence of pain.

The dynamic 1-RM estimate was employed to serve as a more challenging variation to the static, isometric strength assessment. Furthermore, acknowledging the paradigm shift from the traditional medical model of healthcare (expert opinion) to the biopsychosocial model (patient as a consumer and active participant in treatment), the utilization of a performance task where the patient was permitted to self-select a weight based on perceived ability to perform was considered to be complementary to the biopsychosocial framework.[30] Although the task was deemed appropriate because of the subject-perception aspect, no statistical differences in side-to-side strength were noted in either group (dominant to non-dominant arm in the AG and involved to non-involved arm in the SG). To assist in the selection of the appropriate weight, the subjects were permit-

ted to sample various weights and to perform no more than three practice repetitions prior to finalizing their decision on the weight to use for the full 10 repetition trial. However, although the weights could be sampled by the subject prior to final load selection, the lack of difference between the arms during the 1-RM estimate task creates the possibility that some individuals may have underestimated the amount of weight that could be lifted a maximum of 10 repetitions.

Although the three tasks could be reliably reproduced by the two groups over multiple days, the tests could not distinguish performance outcome between individuals with and without shoulder symptoms. The CKCUEST was trending towards being able to distinguish between the two groups (where p = .064) suggesting the more involved physical performance measure may provide clinicians with different information than the traditional strength measures regarding the ability to perform. While Tucci et al found a distinct difference in the number of CKCUEST touches performed between subjects with (10-12 touches) and without (23-28 touches) subacromial impingement syndrome, the subacromial impingement syndrome subjects were 24 years older on average compared to the healthy group. Therefore, the difference between the groups could have been due to age rather than injury presence which limits the interpretability and comparability of the findings to the current study.[15]

Unlike the lower extremity which is sensitive to the effects of injury because of the impact injury can have on stability and mobility, the upper extremity has the advantage of having a separate and independent non-involved extremity which can be utilized for task performance. This phenomenon was demonstrated in the current study where the non-involved arm of the subjects in the SG outperformed the non-dominant arm of the subjects in the AG by 1.5kg (which equates to an approximate difference of three pounds). Although not statistically different, the 1.5kg difference may suggest that the individuals with shoulder symptoms have learned to adapt and modify task performance by utilizing the non-involved arm in a more efficient manner. The decreased effect of injury on the upper extremity is further highlighted in the medical impairment rating literature where the rat-

ings for an injured arm have higher thresholds than the similar impairment ratings for an injured knee.[31] For example, an 8% upper extremity impairment equates to a 5% whole body impairment rating while an 8% lower extremity impairment equates to a 20% whole body impairment rating.[31] It is therefore possible that a general measure of physical performance such as the CKCUEST may help overcome the shortcomings of traditional strength testing as a metric for determining return to activity because of its more challenging requirements thus giving it the ability to potentially better distinguish between persons with and without shoulder symptoms. It is not suggested that traditional strength testing be eliminated from physical assessments because they can have value with detecting certain pathological conditions i.e. rotator cuff injury[32,33] but should be reconsidered as clinical measures for determining cessation of treatment and/or activity readiness.

Finally, the upper extremity physical performance measure literature has suggested that a testing battery comprised of several measures may better assist clinicians in making well-informed clinical decisions about the complex upper extremity and return to activity.[2,28] While this observation has merit, the composition of the testing battery has yet to be established. Recently, Pontillo et al employed an upper extremity pre-season testing battery comprised of isometric strength measures, fatigue tasks, and the CKCUEST in an attempt to predict the occurrence of shoulder injury sustained during a competitive football season.[16] They found that although isometric forward elevation strength and prone-Y to fatigue performance in pre-season were predictive of future injury to the right arm, the CKCUEST was the only maneuver predictive of injury to either arm with a clinical utility of 0.79 sensitivity, 0.83 specificity, and 18.75 positive likelihood ratio.[16] These findings are in contrast to the findings in the current study where the CKCUEST could not clearly discriminate between individuals with and without shoulder symptoms. This contrast however is likely due to differences in the timing of testing (the subjects with shoulder pain in the current study had been experiencing pain from months to years rather than acutely) and the variation in diagnoses identified in each study. Specifically, the current study included diagnoses strictly based on clinical examination without imaging where only half

of the population had suspected internal derangement, while the diagnoses reported by Pontillo et al were primarily cases of instability with verified labral lesions and acromioclavicular separations.[16]

LIMITATIONS

There are several limitations to note in this study. First, the Symptomatic Group was comprised of individuals with various diagnoses. Although the various conditions could allow the results to be generalized, focusing on a specific pathology or condition may have yielded different results. Additionally, none of the subjects were evaluated by a physician and thus no advanced imaging or diagnostic testing (i.e. nerve conduction, diagnostic arthroscopy, etc.) was performed to verify the extent of tissue derangement (assuming any existed). Second, the 1-RM estimate procedure allowed for each subject to self-select the weight he or she perceived as the maximum weight which could be lifted for 10 repetitions. It is possible that some subjects underestimated the weight that could have been lifted and thus limited the chance of finding differences within or between subjects. Third, the closed chain design of the CKCUEST may not provide specific information regarding the ability to perform open chain tasks such as overhead throwing with success. However, the CKCUEST appears to provide different information compared to traditional strength testing highlighting the idea that physical performance measures may allow for the simultaneous assessment of multiple physiological systems better than strength testing. The higher demands of the CKCUEST are likely producing the difference in information but may be one of multiple metrics to utilize for upper extremity performance. Finally, strength was the primary physiological component of physical function that was examined in this study. It is understood that multiple areas of physical function or performance should be considered since human task execution rarely, if ever, utilizes just one component of function during performance. However, strength was the main area of focus since it is commonly considered during the evaluation and rehabilitation of musculoskeletal injury.

CONCLUSIONS

Similar to previous literature, the strength tasks and physical performance measure examined in this

study were found to have excellent test/re-test reliability. The excellent test/re-test reliability has now been expanded to include individuals with various reasons for shoulder symptoms. Traditional strength testing does not appear to be the ideal assessment method to utilize for making discharge and/or return to activity decisions due to the lack of performance differences between the testing groups. Although the tests could be reliably performed, no test could clearly distinguish between individuals with and without shoulder symptoms however; the CKCUEST could have a role as a task to determine readiness to return to activity as it was trending towards being able to discriminate between known groups. Further research needs to exclusively examine subjects with specific pathological conditions such as labral injury, rotator cuff injury, and instability to confirm the clinical utility of the CKCUEST in patients with distinct diagnoses as well as in overhead athletes.

REFERENCES

1. Keskula DR, Duncan JB, Davis VL, Finley PW. Functional outcome measures for knee dysfunction assessment. *J Athl Train.* 1996;31(2):105-110

2. Reiman MP, Manske RC. The assessment of function: How is it measured? A clinical perspective. *J Man Manip Ther.* 2011;19(2):91-99

3. Augustsson J, Thomee R, Linden C, Folkesson M, Tranberg R, Karlsson J. Single-leg hop testing following fatiguing exercise: reliability and biomechanical analysis. *Scand J Med Sci Sports.* 2006;16:111-120

4. Brumitt J, Heiderscheit BC, Manske RC, Niemuth PE, Rauh MJ. Lower extremity functional tests and risk of injury in division III collegiate athletes. *Int J Sport Phys Ther.* 2013;8(3):216-227

5. Earl JE, Monteiro SK, Snyder KR. Differences in lower extremity kinematics between a bilateral drop-vertical jump and a single-leg step-down. *J Orthop Sports Phys Ther.* 2007;37(5):245-252

6. Hayes K, Walton JR, Szomor ZL, Murrell GAC. Reliability of 3 methods of assessing shoulder strength. *J Shoulder Elbow Surg.* 2002;11:33-39

7. Reimann BL, Davies GJ, Ludwig L, Gardenhour H. Hand-held dynamometer testing of the internal and external rotator musculature based on selected positions to establish normative data and unilateral ratios. *J Shoulder Elbow Surg.* 2010;19:1175-1183

8. Michener LA, Boardman Iii ND, Pidcoe PE, Frith AM. Scapular muscle tests in subjects with shoulder pain and functional loss: Reliability and construct validity. *Phys Ther.* 2005;85:1128-1138

9. Turner N, Ferguson K, Mobley BW, Riemann B, Davies G. Establishing Normative Data on Scapulothoracic Musculature Using Handheld Dynamometry. *J Sport Rehabil.* 2009;18:502-520

10. Westrick RB, Duffey ML, Cameron KL, Gerber JP, Owens BD. Isometric shoulder strength refrence values for physically active collegiate males and females. *Sports Health.* 2013;5(1):17-21

11. Michener LA. Patient- and clinician-rated outcome measures for clinical decision making in rehabilitation. *J Sport Rehabil.* 2011;20(1):37-45

12. Negrete RJ, Hanney WJ, Kolber MJ, Davies GJ, Riemann B. Can upper extremity functional tests predict the softball throw for distance: A predictive validity investigation. *Int J Sports Phys Ther.* 2011;6(2):104-111

13. Kumta P, MacDermid JC, Mehta SP, Stratford PW. The FIT-HaNSA demonstrates reliability and convergent validity of functional performance in patients with shoulder disorders. *J Orthop Sports Phys Ther.* 2012;42(5):455-464

14. Goldbeck TG, Davies G. Test-retest relability of the closed chain upper extremity stability test: A clinical field test. *J Sport Rehabil.* 2000;9:35-45

15. Tucci HT, Martins J, de Carvalho Sposito G, Ferreira Camarini PM, de Oliveira AS. Closed Kinetic Chain Upper Extremity Stability test (CKCUES test): a reliability study in persons with and without shoulder impingement syndrome. *BMC Musculoskeletal Disorders.* 2014;15:1-9

16. Pontillo M, Spinelli BA, Sennett BJ. Prediction of in-season shoulder injury from preseason testing in division I collegiate football players. *Sports Health.* 2014

17. Richards RR, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG, et al. A standardization method for the assessment of shoulder function. *J Shoulder Elbow Surg.* 1994;3(6):347-352 10.1016/S1058-2746(09)80019-0.

18. Wainner RS, Fritz JM, Irrgang JJ, Boninger ML, Delitto A, Allison S. Reliability and diagnostic accuracy of the clinical examination and patient self-report measures for cervical radiculopathy. *Spine.* 2003;28(1):52-62

19. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* Jan 15 1998;17(1):101-110

20. Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT, et al. Physical examination tests of the shoulder: A systematic review with meta-analysis of individual tests. *Br J Sports Med.* 2008;42:80-92

21. Hegedus EJ, Goode AP, Cook CE, Michener L, Myer CA, Myer DM, et al. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests *Br J Sports Med.* 2012;46:964-978

22. Myer CA, Hegedus EJ, Tarara DT, Myer DM. A user's guide to performance of the best shoulder physical examination tests *Br J Sports Med.* 2013;47:903-907

23. Sciascia AD, Spigelman T, Kibler WB, Uhl TL. Frequency of use of clinical shoulder examination tests by experienced shoulder surgeons. *J Athl Train.* 2012;47(4):457-466

24. Brzycki M. Strength testing - Predicting a one-rep max from reps-to-fatigue. *J Phys Ed Rec Dance.* 1993;64(1):88-90

25. Denegar CR, Ball DW. Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement. *J Sport Rehabil.* 1993;2:35-42

26. Shrout PE, Fleiss JL. Intraclass correlation: uses in assessing rater reliability. *Psychol Bull.* 1979;86(2):420-428

27. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess.* 1994;6(4):284-290

28. Negrete RJ, Hanney WJ, Kolber MJ, Davies GJ, Ansley MK, McBride AB, et al. Reliability, minimal detectable change, and normative values for tests of upper extremity function and power. *J Strength Cond Res.* 2010;24(12):3318-3325

29. Lovett RW, Martin EG. Certain aspects of infantile paralysis: With description of a method of muscle testing. *J Am Med Assoc.* 1916;66:729-733

30. Adams JR, Drake RE. Shared decision-making and evidence-based practice. *Comm Mental Health.* 2006;42(1):87-105 10.1007/s10597-005-9005-8.

31. Cocchiarella L, Andersson GBJ. *Guides to the Evaluation of Permanent Impairment*. 5th ed. Chicago: AMA Press; 2001.

32. Itoi E, Minagawa H, Yamamoto N, Seki N, Abe H. Are Pain Location and Physical Examinations Useful in Locating a Tear Site of the Rotator Cuff? *Am J Sports Med.* 2006;34(2):256-264

33. McCabe RA, Nicholas SJ, Montgomery KD, Finneran JJ, McHugh MP. The Effect of Rotator Cuff Tear Size on Shoulder Strength and Range of Motion. *J Orthop Sport Phys Ther.* 2005;35:130-135