

# Convergent Validity of the Constant-Murley Outcome Measure in Patients with Rotator Cuff Disease

*Helen Razmjou, Andrea Bean, Joy C. MacDermid, Varda van Osnabrugge, Niki Travers, and Richard Holtby*

## ABSTRACT

**Purpose:** Despite the popularity of the Constant-Murley score (CMS), a combined patient-report and objective (health professional administered) shoulder outcome measure, minimal information exists on its reliability or validity. The purpose of this study was to examine (1) internal consistency and convergent validity of the CMS and (2) the degree to which other shoulder measures explain variation in range of motion and strength components of the CMS.

**Methods:** In this cross-sectional study of outcome measures, 45 patients undergoing conservative or surgical treatment for rotator cuff disease participated. Two disease-specific outcome measures, one shoulder-specific measure, and one upper extremity outcome measure were compared to the CMS.

**Results:** Correlations between competing measures and relative total CMS were moderate (0.56 to 0.75). The American Shoulder & Elbow Score, the Upper Extremity Functional Index, the Western Ontario Rotator Cuff Index, and the Rotator Cuff Quality of Life index explained 38%, 28%, 26%, and 10% of variance of the objective component of the CMS respectively.

**Conclusions:** The patient-report component of the CMS measures a multidimensional concept. The strength component had moderate correlations with isometric strength measures of the shoulder external rotators and abductors. The total CMS appears to measure a construct that is not totally captured by competing measures. The unexplained variance may be due, in part, to the lack of importance of the patient's physical impairment to symptoms or activity limitations as measured by other instruments.

**Key words:** Constant-Murley, outcome measure, reliability, rotator cuff disease, validity

**Razmjou H, Bean A, MacDermid JC, van Osnabrugge V, Travers N, Holtby R. Convergent Validity of the Constant-Murley Outcome Measure in Patients with Rotator Cuff Disease. *Physiother Can.* 2008;60:72-79.**

## RÉSUMÉ

**Objectif:** Bien que le Constant-Murley Score (CMS), qui inclut une autoévaluation de l'épaule par le patient et une mesure de résultat objective (administrée par le thérapeute), soit populaire, peu d'information existe sur sa fidélité et sa validité. Cette étude avait pour buts d'examiner 1) la constance interne et la validité convergente du CMS et 2) le degré auquel d'autres mesures de l'épaule expliquent la variation dans l'amplitude articulaire et dans les éléments de la force du CMS.

**Méthodes:** 45 patients qui subissaient un traitement conservateur ou chirurgical pour une maladie de la coiffe des rotateurs ont participé cette étude transversale des mesures d'impact. Deux mesures d'impact spécifiques la maladie ont été comparées au CMS: une mesure spécifique de l'épaule et une mesure des membres supérieurs.

**Résultats:** Les corrélations entre les mesures concurrentes et le CMS total relatif étaient modérées (0,56 0,75). L'American Shoulder & Elbow Score, l'Upper Extremity Functional Index, le Western Ontario Rotator Cuff Index et le Rotator Cuff Quality of Life Index expliquaient une variance de 38%, 28%, 26% et 10%, respectivement, dans la mesure de résultat objective du CMS.

**Conclusions:** L'élément d'autoévaluation par le patient du CMS mesure un concept multidimensionnel. L'élément de force avait des corrélations modérées avec les mesures isométriques de force des rotateurs et des abducteurs externes de l'épaule. Le CMS total semble mesurer un concept qui n'est pas

*Helen Razmjou, MSc (PT), PhD(C), Cred. MDT:* Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario; Department of Physical Therapy, University of Toronto, Toronto, Ontario.

*Andrea Bean, BSc PT, Cred. MDT:* Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario; Department of Physical Therapy, University of Toronto, Toronto, Ontario.

*Joy C. MacDermid, PT, PhD:* School of Rehabilitation Science, McMaster University, Hamilton, Ontario; Clinical Research Laboratory, Hand and Upper Limb Centre, St. Joseph's Health Centre, London, Ontario.

*Varda van Osnabrugge, BSc PT, Cred. MDT:* Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario; Department of Physical Therapy, University of Toronto, Toronto, Ontario.

*Niki Travers, BPE, BHSc PT, RMT:* Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario.

*Richard Holtby, MB, BS, FRCSC:* Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, Toronto, Ontario; Department of Surgery, University of Toronto, Toronto, Ontario.

This study received ethics approval from the Human Research Ethics Board of the Sunnybrook Health Science Center, Toronto, Canada.

This study was funded by the research funds of the Practice Based Research, Sunnybrook Health Sciences Centre. Helen Razmjou is supported by a Doctoral Fellowship Award from the Canadian Institutes of Health Research.

Address correspondence to: *Helen Razmjou*, Holland Orthopaedic and Arthritic Centre, Sunnybrook Health Sciences Centre, 43 Wellesley Street East, Toronto, ON M1Y 1H1; Tel: 416-967-8654; Fax: 416-967-8716; E-mail: Helen.Razmjou@sunnybrook.ca.

DOI:10.3138/physio/60/1/72

complètement compris par les mesures concurrentes. La variance inexpiquée peut être due, en partie, au manque d'importance de la déficience physique ou des limites de l'activité du patient sur les symptômes mesurés par d'autres instruments.

**Mots clés:** Constant-Murley, mesure d'impact, fiabilité, maladie de la coiffe des rotateurs, validité

## BACKGROUND

Clinical outcome measures are used to represent patient status and to evaluate important clinical changes that result from treatment. Based on the International Classification of Functioning, Disability and Health (ICF) defined by the World Health Organization,<sup>1</sup> the term "functioning" refers to all body functions, activities, and participation while the term "disability" includes the interaction between impairments and activity limitations or participation restrictions. Increasingly, clinicians and researchers are using standardized outcome measures to understand how impairments contribute to disability. In addition, greater attention is now being paid to the theoretical constructs underlying outcome measures.

The Constant-Murley score (CMS), developed in 1986,<sup>2</sup> is a commonly used shoulder-specific outcome measure. It is a summative scale that provides a global score based on weighted measures of physical impairments in range of motion (ROM) and strength, along with patient report of pain and activity limitation. Although the conceptual framework for development of the CMS has not been described, the emphasis on symptoms and functional difficulties would suggest that the CMS is a disability measure. The patient-report component of the CMS documents pain and difficulty in activities of daily living, work, sport, and sleep.<sup>3</sup> Discriminative and evaluative properties of these different metrics would likely vary. The "objective" clinical assessment involves having a health professional measure ROM and strength. A unique feature of the CMS is that it allows for age- and sex-related changes in the musculoskeletal system by converting the absolute score to the relative score to adjust for age and sex. This normalization is based on ROM and strength measurements of 900 healthy individuals.<sup>2</sup>

The CMS has been widely used in North America and Europe, with some information on its reliability.<sup>4-10</sup> Because of variations in application and scoring of this measure, information on its validity may not be comparable across studies.<sup>11-13</sup> A number of other additional issues might compromise the validity of the CMS, particularly when it is used in comparisons across different centres. Moseley originally described the strength measurement for the CMS through the use of either an unsecured cable tensiometer or a spring balance held at arm's length at 90° of forward flexion in the sagittal plane and 90° of abduction in the coronal plane.<sup>14</sup> Although standardization is important for minimizing

measurement error, Constant and Murley did not elaborate on the exact position of the hand, elbow, or shoulder or on the method for obtaining the strength score. For example, the plane of elevation, the duration of each measurement, and the number of repetitions are not specified. To improve consistency among clinicians and researchers, the European Society for Surgery of the Shoulder and Elbow (ESSSE) has provided guidance on the preferred use of the outcome scores.<sup>15</sup> Such recommendations are not uniformly applied, however, and potential variations have not been tested for equivalence.

The CMS is distinguished from other shoulder outcome measures that are primarily self-reported by the fact that the health professional-administered component accounts for 65% of the total score. For this reason, it is important to understand the relationship between the CMS and competing self-report shoulder scales. Since the measure of CMS strength is also unique, understanding how it correlates with simple isometric tests would be informative. In addition, as noted earlier, because of the lack of proper description of the CMS in previous studies, it is difficult to generalize the previous reliability and validity results. Therefore, the primary purposes of this study were to examine

1. the distribution of scores and internal consistency of the CMS,
2. cross-sectional convergent validity of the relative CMS total score with four self-report outcome measures,
3. the relationship between measured isometric strength in three directions and the strength score of the CMS, and
4. the extent to which different patient-report outcome measures explain variance in ROM and strength subscales and total score of the relative CMS.

The following hypotheses were tested:

1. Internal consistency analysis examined the relationship between subcategories of the patient-report component of the CMS. A low Cronbach's alpha correlation coefficient ( $\alpha < 0.70$ ) was expected, because each component of the patient-report score is intended to measure a different entity.
2. Analysis of convergent validity (the extent of correlation among measures at one point in time) examined the extent to which the total score of the relative CMS agreed with the result of other measures believed to be assessing the same attributes (symptoms and activity limitations). A priori predictions were made based on the conceptual model of the

CMS and results of previous studies.<sup>11,16</sup> We predicted that the correlation between total score on the CMS and the Western Ontario Rotator Cuff Index (WORC), the Rotator Cuff Quality of Life index (RC-QOL), the American Shoulder & Elbow Score (ASES), and the Upper Extremity Functional Index (UEFI) would be moderate ( $r=0.50-0.80$ ), because the CMS combines patient report with objective joint-specific measures, with some emphasis on pain and activity limitation.

3. It was hypothesized that the isometric strength component of the CMS would have a moderate correlation with isometric shoulder abduction and internal and external rotation in neutral position ( $r=0.50-0.80$ ).
4. The last hypothesis was that other shoulder measures should explain a moderate amount of variance of the CMS objective component.

## METHODS

### Participants

To examine the generalizability of the results to mild and severe rotator cuff disease, the study involved two groups of consecutive patients: (1) patients diagnosed with impingement syndrome or partial thickness rotator cuff tears (based primarily on clinical examination; some patients had undergone ultrasound or MRI), and (2) patients referred for rehabilitation following repair of full-thickness rotator cuff tears. Patients with referred pain, systemic inflammatory disease, or fractures of the upper extremity were excluded. All subjects provided informed consent. Approval for use of human subjects was obtained from the Research Ethics Board of the Sunnybrook Health Sciences Centre.

### Outcome Measures

Eight to 10 weeks after the initiation of physical therapy treatment, all patients were provided with two disease-specific outcome measures, the WORC index<sup>16</sup> and the RC-QOL index;<sup>17</sup> two shoulder-specific measures, the CMS<sup>2</sup> and the ASES score;<sup>18</sup> and one upper extremity outcome measure, the UEFI index.<sup>19</sup> Several studies have reported the extent to which these measures' scores have been shown to be reliable and valid.<sup>11,12,16-22</sup>

Although the framework for development of the WORC and RC-QOL indices was to measure quality of life (QOL) in different domains related to health, the focus of both instruments is to document symptoms, emotional difficulties, and activity limitations. Functional status is a health-related variable that is often confused with QOL. It is important to note that symptoms and functional difficulties secondary to rotator cuff disease can reduce QOL, and perhaps that is why they have been included in these types of questionnaires. However, not all individuals with reduced QOL

experience the same symptoms or functional difficulties.<sup>23</sup> This lack of reciprocal relationship between items that represent side effects of a disease (i.e., causal variables of QOL, such as pain, stiffness, and difficulty performing a task) and overall QOL suggests that the RC-QOL and WORC are predominantly measures of disability rather than QOL.<sup>23</sup> The UEFI is a pure functional measure for the upper limb, without any items directly related to symptoms. The ASES assigns 50 points to pain and 50 points to difficulty in performing upper-limb activities, work, and sport; it is consistent with the definition of a disability measure.

There are four CMS categories for pain: no pain (15 points), mild pain (10 points), moderate pain (5 points), and severe pain (0 points). Scores for activities of daily living (work and sports activities) vary from 0 (severe restriction) to 4 (no restriction). Sleep is categorized as not disturbed (2), grossly disturbed (1), or severely disturbed (0). Functional positioning is rated from 2 to 10 to signify waist-level to above-head movements as perceived by the patient. For the present study, strength and ROM were measured within the "pain-free" range, based on recommendations of ESSSE.<sup>15</sup> Assessing performance within the pain-free range is expected to provide a better representation of actual function.

The ROM and strength assessments were performed with the patient in sitting or standing position. No rotation of the upper body was allowed during the examination. Shoulder flexion, abduction, external rotation, and internal rotation (hand behind back) were measured separately. In the CMS system, there is precise information about how the points are calculated for flexion and abduction. The range is categorized in six mutually exclusive ranges of 0–30°, 31–60°, 61–90°, 91–120°, 121–150°, and 151–180°. Each range greater than 30° receives 2 points, with a maximum of 10 points. External rotation was measured depending on where the elbow and hand were in relation to the head; the hand was not allowed to touch the head. Two points were allocated for each of the following movements, with a minimum of 0 and maximum of 10 points: hand behind head with elbow forward; hand behind head with elbow back; hand on top of head with elbow forward; hand on top of head with elbow back; and full elevation from on top of head. Internal rotation was based on anatomical landmarks attained by the end of the thumb of the reaching hand. These landmarks are the greater trochanter (lateral thigh), the gluteal region, the lumbosacral junction, L3 (waist), T12, and T7 (interscapular area).

The ESSSE<sup>15</sup> recommends the scapular plane instead of flexion and abduction planes for strength measurement using a mounted tensiometer. Since the assessment was performed in different clinic rooms where securing the tensiometer on the floor was not feasible, we modified the protocol by using an unsecured tensiometer. Therefore, strength was measured with a simple



**Figure 1** Strength measurement with a tensiometer in scapular plane.

tensiometer with the shoulder at 90° of elevation in the plane of the scapula and the elbow extended while the clinician pulled down on the tensiometer (see Figure 1). The maximum “pain-free” force that the patient could resist for 5 seconds as the examiner pulled down on the device was measured. One point was awarded for each pound lifted, to a maximum of 25 pounds. In the case of pain while holding the position, strength was given a score of zero.

The *absolute CMS* is based on a 100-point scale; the patient-report component is equal to 35 points, while measures of ROM and strength add up to a maximum of 65 points (40 points for ROM and 25 points for strength). The minimum absolute score is 0; 100 is the highest raw score, indicating full pain-free function. The *relative CMS* is calculated by normalizing for age and sex using the formula provided by Constant and Murley.<sup>2</sup> It is common practice to treat ordinal variables as interval-scaled variables when calculating the total score of patient-report measures of disability, and, despite controversy on the best analytical approach for rating scales, treating the total aggregate score as an interval-level measure does not introduce a major bias.<sup>24,25</sup> The original authors<sup>2,3</sup> and authors of all subsequent studies that have used the CMS as an outcome measure have made the assumption that equal distances between response choices represent equal distances on the pain or function dimensions. In the present study, therefore, we examined the subcategories as interval variables.

As a part of this study, isometric strength assessments of abduction, internal rotation, and external rotation were conducted using a hand-held dynamometer (ALIMED Inc., Dedham, MA 02027-9135) calibrated for each patient. Reliability of hand-held dynamometer scores has been reported to be high.<sup>26,27</sup> Assessment was performed with the patient sitting with the arm in a neutral position and the elbow tucked in against the body. The dynamometer was placed on the inner or

**Table 1** Demographic Frequency Distribution for 45 Participants: *N* (%).

Sex	Female	24 (53%)
	Male	21 (47%)
Dominant side	Right	37 (82%)
	Left	8 (18%)
Affected side	Right	31 (69%)
	Left	14 (31%)
Status	Surgical	33 (73%)
	Non-surgical	12 (27%)
Symptom characteristics	Night pain	28 (62%)
	Loss of motion	33 (73%)
	Weakness	34 (76%)
	Pain on movement	37 (82%)
Mechanism of injury	Insidious	13 (29%)
	Repetitive activities	11 (24%)
	Fall on an outstretched arm	6 (13%)
	Direct blow	5 (11%)
	Other	10 (22%)

outer aspect of the distal one-third of the forearm for internal and external rotation respectively. Abduction was assessed with the dynamometer placed in the outer aspect of the mid-upper arm. The average of three efforts (held for 3 seconds) was used for each movement.

### Statistical Analysis

Descriptive statistics were calculated for age, sex, affected and dominant side, mechanism of injury, symptoms, scores of pain-free range of motion and strength, and total scores for all measures (see Tables 1 and 2). Cronbach’s alpha was calculated to examine the extent of homogeneity and internal consistency between different patient-report subcategories. Pearson correlation coefficients and 95% confidence intervals were calculated to examine (1) convergent validity between the total scores of all measures and (2) correlations between isometric strength measurements (abduction, internal rotation, external rotation) and the CMS strength score on the affected side.

Separate linear regression analyses were conducted for total scores of the competing shoulder measures (ASES, UEFI, RC-QOL, WORC) as the only predictor variable in the model. Pain-free ROM, pain-free strength, and total objective CMS were used in separate analyses as the sole dependent variable. We calculated 95% confidence and prediction bands. Confidence bands are applied when estimating the mean value of the dependent variable for a specific independent variable value; prediction bands are applied when estimating the dependent variable value for an individual for a specific independent variable value. The SAS version 9.1 statistical package (SAS Institute, Cary, NC) was used for all analyses.

## RESULTS

Forty-five subjects, 21 men and 24 women whose average age was 56.84 years (minimum, maximum: 25,

**Table 2** Descriptive Data for Outcome Measures.

Variable	N	Mean	SD	Minimum	Maximum
ACMS	43	64.11	15.89	33.00	96.00
RCMS	43	80.98	20.02	41.25	126.66
Subjective	43	21.48	7.08	8.00	35.00
Objective	44	42.68	11.19	20.00	65.00
ROM	44	32.59	7.87	16.00	40.00
Strength	45	10.36	6.57	3.00	38.00*
ASES	44	70.80	17.89	28.33	100.00
RC-QOL	44	70.80	21.41	28.00	100.00
WORC	45	67.60	20.93	24.00	100.00
UEFI	43	65.21	17.93	20.00	96.00

ACMS = Absolute Constant-Murley score; ASES = American Shoulder & Elbow Score; ROM = range of motion; RCMS = Relative Constant-Murley score; RC-QOL = Rotator Cuff Quality of Life score; WORC = Western Ontario Rotator Cuff Index; UEFI = Upper Extremity Functional Index

\*Maximum strength was 38 lb, but all scores above 25 were adjusted to 25 for the CMS total score calculation.

82 years; SD: 15.8 years), participated in the study. Tables 1 and 2 provide descriptive demographic information and scores for all measures. A few of the maximum scores of the relative CMS (after being adjusted for strength) were still above the accepted score of 100, indicative of overcorrecting for age and sex factors. For example, the 90% quantile included high scores of 107%, 108%, and 126%.

Hypothesis 1 was accepted. The raw coefficient alpha was 0.60. Item correlations with the total subjective score varied from 0.37 to 0.50, indicating that sleep, work, recreational activities, and pain measure different aspects of patient-reported complaints. This means that the CMS patient-report subcategories are multidimensional without redundant questions.

Hypothesis 2 was accepted, because correlations between the relative CMS and the UEFI, RC-QOL, ASES, and WORC were moderate: 0.56 (0.31–0.73), 0.65 (0.42–0.79), 0.68 (0.48–0.82), and 0.75 (0.57–0.85) respectively. Correlations between patient-report CMS and the UEFI, RC-QOL, ASES, and WORC were higher at 0.66 (0.43–0.80), 0.82 (0.69–0.90), 0.70 (0.50–0.83), and 0.85 (0.74–0.92) respectively.

Hypothesis 3 was rejected for internal rotation, as the correlation was less than moderate at 0.40 (0.07–0.64) between isometric internal rotation and the CMS measure of strength. Correlations between the CMS measure of strength in the scapular plane and isometric abduction and external rotation were 0.59 (0.31–0.77) and 0.56 (0.28–0.75) respectively.

Hypothesis 4, which predicted that patient-report measures would explain a moderate amount of variance of the CMS, was not fully supported. Simple linear regression analyses showed that disability, as measured by patient-report measures, was more affected by impairment in functional ROM than by impairment in functional strength (see Table 3). The ASES and UEFI had the highest F-ratios, explaining 30 to 40% of variance in functional pain-free ROM and total score of the

**Table 3** Regression Analysis Statistics Explaining Variation in Pain-Free Range of Motion, Strength, and Total Objective Component of the CMS by Each Patient-Report Measure.

	F-Ratio	R <sup>2</sup>	Parameter Estimate* (CI)
Predictor Variable (ASES)			
ROM	15.39	28%	0.23 (0.11–0.36)
Strength	12.46	23%	0.15 (0.06–0.24)
Total objective CMS	24.43	38%	0.39 (0.23–0.55)
Predictor Variable (UEFI)			
ROM	16.96	30%	0.24 (0.12–0.36)
Strength	3.61	8%	0.09 (–0.01–0.18)
Total objective CMS	15.34	28%	0.33 (0.16–0.50)
Predictor Variable (WORC)			
ROM	9.08	18%	0.16 (0.05–0.27)
Strength	8.61	17%	0.11 (0.03–0.18)
Total objective CMS	14.68	26%	0.28 (0.13–0.42)
Predictor Variable (RCQOL)			
ROM	2.99	7%	0.09 (–0.01–0.21)
Strength	3.01	6%	0.06 (–0.01–0.14)
Total objective CMS	4.64	10%	0.17 (0.01–0.32)

CMS = Constant-Murley score; ASES = American Shoulder & Elbow Score; WORC = Western Ontario Rotator Cuff Index; UEFI = Upper Extremity Functional Index; RC-QOL = Rotator Cuff Quality of Life Score; ROM = range of motion

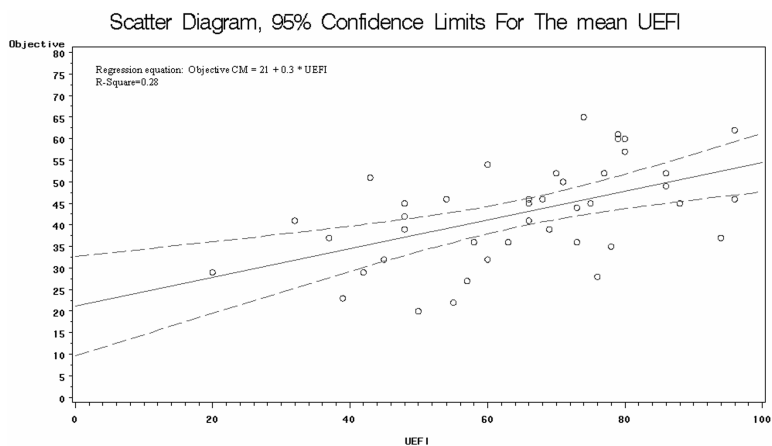
\*Parameter estimates are the estimates of predictor variables.

objective CMS. The ASES and WORC were affected by both ROM and strength. In our sample, however, the RC-QOL explained only 10% of variance in total objective score of the CMS. A scatter plot of one of the measures (UEFI) is provided as an example; Figures 2 and 3 demonstrate the confidence intervals and prediction bands for this measure.

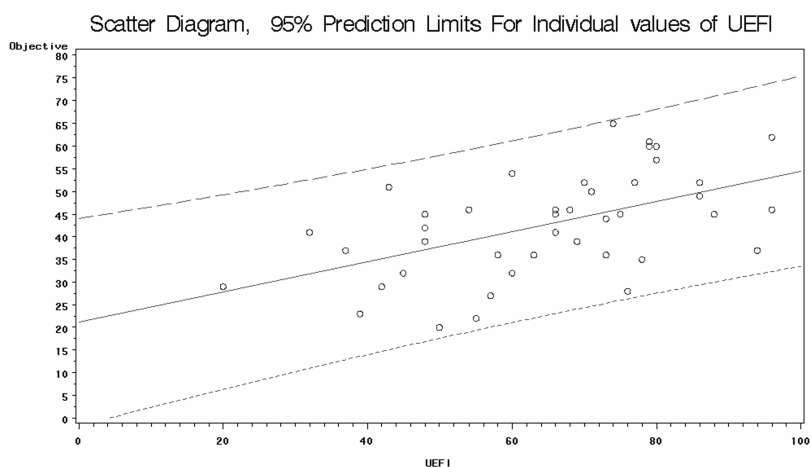
## DISCUSSION

The low internal consistency of the CMS patient-report subcategories indicates that the CMS measures different aspects of symptoms and daily activities. Results of the cross-sectional analysis indicate that the CMS has a moderate correlation with patient self-report measures, which is consistent with previous studies.<sup>11,16</sup> Moderate correlations between 90° of scapular plane elevation and isometric abduction and external rotation, which are known to indicate rotator cuff pathology,<sup>28–30</sup> may indicate the usefulness of the scapular plane strength measurement. However, further electromyography (EMG) studies are needed to examine this relationship more accurately.

Overall, the amount of variance explained by all other outcome measures with regard to the objective components of the CMS and its total score was low. The ASES, UEFI, and WORC had the highest value of R<sup>2</sup>, indicating that at least three self-report measures individually explained 30 to 40% of the variance in functional ROM and strength. The lower R<sup>2</sup> (0.10) of the RC-QOL may be related to the type of questions on this instrument, which address more intense and physically demanding activities such as shovelling snow, raking the lawn,



**Figure 2** Scatter diagram for the mean of the UEFI.



**Figure 3** Scatter diagram for the individual values of the UEFI.

and scrubbing pots and pans. The WORC and UEFI address more routine activities that may be more relevant to patients' actual impairment.

In our study, the maximum scores of the relative CMS overestimated the improvement in older patients as compared to the normal score of 100, raising questions about the formula proposed<sup>2</sup> for adjusting for age (see Table 1). Romeo et al.<sup>13</sup> found a significant effect of age on CMS only for female patients undergoing rotator cuff repair; other investigators have reported that normalizing for age and sex according to the original data by Constant and Murley may overestimate shoulder function in older individuals (40 and 60 years of age in women and men respectively).<sup>10</sup> Considering that our results support the above findings,<sup>10</sup> future studies are required, with larger samples, to investigate the impact of age and sex normalization as recommended by Constant and Murley.<sup>2</sup> Until such studies are performed, however, investigators should present both adjusted and unadjusted scores and consider any potential biases that may arise from overcorrection when comparing to other studies.

The Constant-Murley outcome measure is one of the most commonly used international shoulder-scoring scales consisting of both objective (administered and scored by a health professional) and patient self-report components. Conceptually, it may not be optimal to combine physical impairment scores with self-reported functional disability scores, because different concepts are pooled using arbitrary weightings that may or may not reflect actual function. The application of objective clinical assessment of ROM and strength is known to provide reliable and valid measures of physical impairment under certain circumstances. For example, the assessment of shoulder motion using goniometry<sup>31-34</sup> and of shoulder strength using hand-held dynamometers has been shown to produce reliable quantitative measures.<sup>35-38</sup> Conversely, the CMS outcome measure has been criticized for a lack of consensus on strength measurement and may be subject to additional sources of error variation. Similar concerns have been posed on the relevance of the self-report questions or ROM assessment. For example, classification of patients into

different ROM categories based on only 1° of difference (e.g., 120° vs. 121°) is unlikely to be valid. Furthermore, these ROM categories are arbitrary and do not necessarily reflect functional shoulder movements. Self-report of function is not specific to any particular activity and is therefore left open to interpretation by the patient. Patients who are not working or involved in sports activities may leave the “work or sports” questions blank because no “not applicable” option has been provided. In addition, some studies use a visual analogue scale (score of 0 to 10) while others use a categorical classification system (four categories) for measuring pain, adding to variation among studies. Given the problems with measurement scaling, standardization of testing protocols, device consistency and calibration, and construction and weighting of the scale itself, there is room for improvement in this measure.

This study provides preliminary information on the use of the CMS in patients with rotator cuff pathologies. Future cross-sectional and longitudinal studies with larger samples are needed to investigate the reliability, practicality, and validity of the CMS in patients with different levels and types of shoulder pathology.

## CONCLUSION

The Constant-Murley shoulder outcome measure has become widely accepted, although its development, standardization, and psychometric evaluation have not complied with accepted standards in this field. This study provides support for the use of the CMS measure in patients with rotator cuff disease but also recognizes the need for continued improvement in the application, reporting, and interpretation of this measure.

The patient self-report component of the CMS is a multidimensional construct, measuring different aspects of subjective complaints. In the present study, the strength component had a moderate correlation with function of the shoulder abductors and external rotators. Approximately 30 to 40% of variance in ROM and strength impairment was explained by the scores of the ASES, UEFI, and WORC, which predominantly measure symptom severity and/or activity limitations. The unexplained variance may be due, in part, to the lack of importance of the physical impairment to symptoms or activity limitations as measured by patient-report instruments.

Clinicians should decide on the best measure of disability depending on their own clinical needs. The potential advantage of self-reported functional measures is that they require less clinician time and minimal equipment, making it relatively easy to incorporate them in a variety of assessment situations. In addition, extensive

data collection by clinicians adds to the complexity of applying and interpreting the outcome measures. Selecting the most appropriate outcome measure is important, as it affects cost, time, and the burden placed on patients and clinicians. Patient self-report measures may be more practical for the majority of clinicians; however, as physical impairment and self-perceived function may represent different constructs, it is valuable to have reliable and valid measures of both.

Additional research, including direct comparison of the CMS with actual measures of physical impairment and functional performance in longitudinal studies, is required to determine optimal methods of assessing outcomes following interventions for shoulder disorders. Further studies should examine the discriminative and evaluative properties of this measure in larger samples of patients with shoulder pathology.

## KEY MESSAGES

### What Is Already Known on This Subject

The Constant-Murley score (CMS) is one of the most commonly used international shoulder scoring scales. The CMS combines physical impairment scores with self-reported functional disability scores using arbitrary weightings. Previous studies have not explored the relationship between different components of the CMS and competing self-report measures. In addition, because the CMS has not been applied consistently in these studies, it is difficult to generalize the results of the reliability and validity of this measure.

### What This Study Adds

The present study examined the internal consistency and cross-sectional convergent validity of the CMS as described by the European Society for Surgery of the Shoulder and Elbow (ESSSE). Our results suggest that the three commonly used self-report measures, the ASES, UEFI, and WORC, individually explain a low proportion of the variance in functional ROM and strength. Our study provides support for the use of the CMS in patients with rotator cuff disease but also recognizes the need for continued improvement in how this measure is applied, reported, and interpreted.

## REFERENCES

1. World Health Organization [WHO]. WHO International Classification of Functioning, Disability and Health. Geneva: The Organization; 2001.
2. Constant CR. Age related recovery of shoulder function after injury [thesis]. Cork, Ireland: University College; 1986.
3. Constant CR. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br.* 1997;79:695–6.

4. Bankes MJ, Emery RJ. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br.* 1997;79:696.
5. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley shoulder assessment. *J Bone Joint Surg Br.* 1996;78:229–32.
6. Cook KF, Roddey TS, Olson SL, Gartsman GM, Valenzuela FF, Hanten WP. Reliability by surgical status of self-reported outcomes in patients who have shoulder pathologies. *J Orthop Sport Phys.* 2002;32:336–46.
7. Brinker MR, Cuomo JS, Popham GJ, O'Connor DP, Barrack RL. An examination of bias in shoulder scoring instruments among healthy collegiate and recreational athletes. *J Shoulder Elbow Surg.* 2002;11:463–9.
8. Machesney MR, Emery RJH. The method for the objective measurement of muscular strength and presentation of reference ranges for the Constant-Murley score. *J Shoulder Elbow Surg.* 1994;00:S26.
9. Johansson KM, Adolphson LE. Intraobserver and interobserver reliability for the strength test in the Constant-Murley shoulder assessment. *J Shoulder Elbow Surg.* 2005;14:273–8.
10. Yian EH, Ramappa AJ, Arneberg O, Gerber C. The Constant score in normal shoulders. *J Shoulder Elbow Surg.* 2005;14:128–33.
11. Holtby R, Razmjou H. Measurement properties of the Western Ontario Rotator Cuff outcome measure: a preliminary report. *J Shoulder Elbow Surg.* 2005;14:506–10.
12. MacDermid JC, Ramos J, Drosdowech D, Faber K, Patterson S. The impact of rotator cuff pathology on isometric and isokinetic strength, function, and quality of life. *J Shoulder Elbow Surg.* 2004;13:593–8.
13. Romeo AA, Mazzocca A, Hang DW, Shott S, Bach BR Jr. Shoulder scoring scales for the evaluation of rotator cuff repair. *Clin Orthop Relat Res.* 2004;427:107–14.
14. Moseley HF. *Shoulder lesions.* 2nd edn. Edinburgh: Churchill Livingstone; 1972. p. 28–9.
15. Scoring systems for the shoulder: circular to members of British Shoulder and Elbow Society [document on the Internet]; 1997 [cited 2007 Nov 5]. Available from: [http://www.axelina.com/ENG/kvalitet/kva\\_score.htm](http://www.axelina.com/ENG/kvalitet/kva_score.htm).
16. Kirkley A, Alvarez C, Griffin S. The development and evaluation of a disease-specific quality-of-life questionnaire for disorders of the rotator cuff: the Western Ontario Rotator Cuff Index. *Clin J Sport Med.* 2003;13:84–92.
17. Hollinshead RM, Mohtadi NG, Vande Guchte RA, Wadey VM. Two 6-year follow-up studies of large and massive rotator cuff tears: comparison of outcome measures. *J Shoulder Elbow Surg.* 2000;9:373–81.
18. Richards RR, An K, Bigliani LU, Friedman R, Gartsman GM, Gristina AG, et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg.* 1994;3:347–52.
19. Stratford PW, Binkley JM, Stratford DM. Development and initial validation of the upper extremity functional index. *Physiother Can.* 2001;53:259–67.
20. Razmjou H, Bean A, van Osnabrugge V, MacDermid JC, Holtby R. Cross-sectional and longitudinal construct validity of two rotator cuff disease-specific outcome measures. *BMC Musculoskelet Disord.* 2006;7:26.
21. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg.* 2002;11:587–94.
22. MacDermid JC, Drosdowech D, Faber K. Responsiveness of self-report scales in patients recovering from rotator cuff surgery. *J Shoulder Elbow Surg.* 2006;15:407–14.
23. Fayers PM, Hand DJ, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual Life Res.* 1997;6:393–406.
24. Tennant A, Hillman M, Fear J, Pickering A, Chamberlain MA. Are we making the most of the Stanford Health Assessment Questionnaire?. *Br J Rheumatol.* 1996;35:574–8.
25. Kaplan RM, Ernst JA. Do category rating scales produce biased preference weights for a health index? *Med Care.* 1983;21:193–207.
26. Leggin BG, Neuman RM, Iannotti JP, Williams GR, Thompson EC. Intrarater and interrater reliability of three isometric dynamometers in assessing shoulder strength. *J Shoulder Elbow Surg.* 1996;5:18–24.
27. Hayes K, Walton JR, Szomor ZL, Murrell GA. Reliability of 3 methods for assessing shoulder strength. *J Shoulder Elbow Surg.* 2002;11:33–9.
28. Diehr S, Ison D, Jamieson B, Oh R. Clinical inquiries: what is the best way to diagnose a suspected rotator cuff tear? *J Fam Pract.* 2006;55:621–4.
29. Murrell GA, Walton JR. Diagnosis of rotator cuff tears [published erratum appears in *Lancet.* 2001;357:1452]. *Lancet.* 2001;357:769–70.
30. Harryman DT II, Mack LA, Wang KY, Jackins SE, Richardson ML, Matsen FA III. Repairs of the rotator cuff: correlation of functional results with integrity of the cuff. *J Bone Joint Surg Am.* 1991;73:982–9.
31. Green S, Buchbinder R, Forbes A, Bellamy N. A standardized protocol for measurement of range of movement of the shoulder using the Plurimeter-V inclinometer and assessment of its intrarater and interrater reliability. *Arthritis Care Res.* 1998;11:43–52.
32. Hayes K, Walton JR, Szomor ZR, Murrell GA. Reliability of five methods for assessing shoulder range of motion. *Aust J Physiother.* 2001;47:289–94.
33. MacDermid JC, Chesworth BM, Patterson S, Roth JH. Intratester and intertester reliability of goniometric measurement of passive lateral shoulder rotation. *J Hand Ther.* 1999;12:187–92.
34. Youdas JW, Carey JR, Garrett TR, Suman VJ. Reliability of goniometric measurements of active arm elevation in the scapular plane obtained in a clinical setting. *Arch Phys Med Rehabil.* 1994;75:1137–44.
35. Fokter SK, Cicak N, Skorja J. Functional and electromyographic results after open rotator cuff repair. *Clin Orthop Relat Res.* 2003;415:121–30.
36. Itoi E, Minagawa H, Sato T, Sato K, Tabata S. Isokinetic strength after tears of the supraspinatus tendon. *J Bone Joint Surg Br.* 1997;79:77–82.
37. Plotnikoff NA, MacIntyre DL. Test–retest reliability of glenohumeral internal and external rotator strength. *Clin J Sport Med.* 2002;12:367–72.
38. van Meeteren J, Roebroek ME, Stam HJ. Test–retest reliability in isokinetic muscle strength measurements of the shoulder. *J Rehabil Med.* 2002;34:91–5.