Research Report

VALIDITY AND RELIABILITY OF SELECTED OUTCOME MEASURES USED IN REHABILITATION FOR ANTERIOR CRUCIATE LIGAMENT RECONSTRUCTION: A LITERATURE REVIEW

Warren KJC, MHlthSc, PTRP; Chua CDM, PTRP; Tagala AA, PTRP; Cadiz BO, PTRP; Maglanque NP II

ABSTRACT

***Objectives*:** The aim of this study is to conduct a literature review of available evidence on selected outcome measures commonly used during rehabilitation following Anterior Cruciate Ligament Repair (ACL-R) with emphasis on their validity and reliability. ***Methodology*:** A comprehensive search for articles was performed through manual and electronic searches using predetermined keywords. Full text articles related to pre-identified outcome measures specifically, pain assessment scales, range of motion measurements, circumference measures, functional performance tests, and the Cincinnati Knee Rating System, were reviewed. ***Results*:** A total of 20 articles were included in the review. The outcome measures selected for the study were found to be valid and yielded moderate to high reliability. **Conclusion and Recommendations:** The outcome measures that were found to be valid and reliable were summarized into an evaluation form with a user's guide that outlines the testing procedures for use in the clinical setting.

Keywords: *outcome measures, Anterior Cruciate Ligament (ACL), validity, and reliability*

INTRODUCTION

Outcome measurement is the core of good clinical practice. As physiotherapy moves toward professional autonomy, practitioners must have a deeper understanding of specific assessment procedures. Outcome measurement is used to assess, evaluate, and justify clinical practice.[1] In sports physiotherapy, outcome measures determine the athlete's readiness to return to pre-injury activity level and minimize the risk of re-injury.[2]

With an increase in participation in athletic activities, a consequent rise in the number of sports-related injuries has been described in the literature.[3] Of these, injuries to the knee and its ligaments such as the anterior cruciate ligament (ACL) are the most common.[3] In the United States alone, ACL tears are estimated to occur in 1 in 3000[4] and there are approximately 100,000 ACL reconstructions (ACL -R) being carried out yearly. [4]

While surgery is important, rehabilitation following ACL-R is just as essential.[5] After surgical reconstruction of a torn ACL, the patient experiences a number of impairments and activity limitations as defined by the International Classification of Impairment, Disability and Handicap II (ICIDH-II, 2000).[6] A study by Risberg and colleagues[7] identified impairments following ACL reconstruction as anatomic (anterior displacement of tibia relative to femur) and physiologic (range of motion, muscle performance, pain) which can be measured by knee arthrometer, goniometer, and visual analog scales for pain. Activity limitations related to post ACL-R on the other hand can be assessed by scoring systems such as International Knee Documentation Committee (IKDC) evaluation, Cincinnati, and Tegner and Lysholm assessments which incorporate symptoms, signs, instrumented laxity testing and return-to-activity assessment.

At present, the use of outcome measures is left to the clinician's judgment as there is no consensus regarding which tests or combination of tests should be administered for a more comprehensive measure of outcome following ACL-R.[8] Therefore, when selecting specific tests to administer, the clinician should confront issues on reliability and validity in order to have meaningful and useful information. Regardless of which tests are chosen, it is crucial that they be standardized, reliable and valid.

Validity is defined as the ability of an instrument to actually measure what it is intended to measure.[8] There are different types of validity namely content, convergent, construct, and criterion-related. Content validity is the ability of an outcome measure to represent the variable of interest (also considered a theoretical form of validity).[8] Convergent validity, on the other hand, indicates that 'two measures assessing the same phenomenon measure the same construct, and yield similar results'.[9] When an instrument is able to represent a scientific hypothesis or the 'idea' underlying a measurement protocol, the instrument is said to demonstrate construct validity.[8] An assessment procedure is said to have criterion validity when it is able to demonstrate a close relationship to another instrument (criterion) or a specific scientific criteria.[8]

On the other hand, reliability refers to 'whether a measurement procedure or operational definition minimizes measurement error generating accurate and consistent measurements between repeated measures of the same variable'.[10] In sports physiotherapy, critical decisions such as return to competition of athletes is termed to be criteria-based (or based on set parameter) therefore "high" reliability of outcome measures to be used is vital.[2] The Intra-class Correlation Coefficient (ICC) is presently the recommended agreement for quantifying reliability.[2] Correlation coefficients of 0.50-0.70 are considered 'weak', 0.70-0.80 for 'moderate', 0.80-0.90 for 'high' and greater 0.90 as 'strong'.[2]

There are a variety of objective and subjective outcome measures that are currently being used following ACL-R. Tests should be 'inexpensive, take an acceptable length of time to administer, acceptable for the patient and convenient for the clinicians to use.[10] Thus, selected outcome measures that fit these criteria are of interest to this study including range of motion measures using goniometry and heel height measurements, circumference measures, functional performance tests, pain assessment scales, and Cincinnati Knee Rating System.

The objective of this study is to conduct a literature review of available evidence on selected outcome measures commonly used during the rehabilitation following ACL-R with emphasis on their validity and reliability.

This study could lead to the conduct of future researches involving standardization of post ACL-R rehabilitation assessment. It hopes to provide a basis for devising and recommending a knee evaluation form for ACL-R which includes a user's guide containing testing procedures and methods of administration recommended by the literature included in the study. The devised form can serve as an assessment tool to guide physiotherapists in the rehabilitation of patients who underwent ACL-R.

METHODOLOGY

The researchers have undertaken a comprehensive search of available studies using pre-determined keywords (Table 1).

Table 1. List of keywords
anterior cruciate ligament
anterior cruciate ligament reconstruction
anterior cruciate ligament assessment
outcome measures
knee assessment
knee rating scales
goniometer or goniometry
heel-height difference
functional tests
single hop test, triple hop test
limb girth measurement or muscle bulk measurement
Cincinnati Knee Rating Scale
pain assessment

The search was conducted through the use of libraries, popular journal holdings, and electronic databases without limitations on the year of publication. A summary of the methodology utilized for this study is depicted in Figure 1.

Only journals that investigated the validity and reliability of selected outcome measures on post ACL-R patients of any age group were included. Those with full text written in the English language were retrieved and used for the study. Studies collected were sorted according to its specific outcome measure being investigated such as pain assessment scales, range of motion measurements, circumference measures, functional performance tests, and Cincinnati Knee Rating System. Each category (e.g. pain assessment scales, goniometry and functional performance test) was assigned to one reviewer for a thorough appraisal of the journals included. The review focused on collecting information about validity and

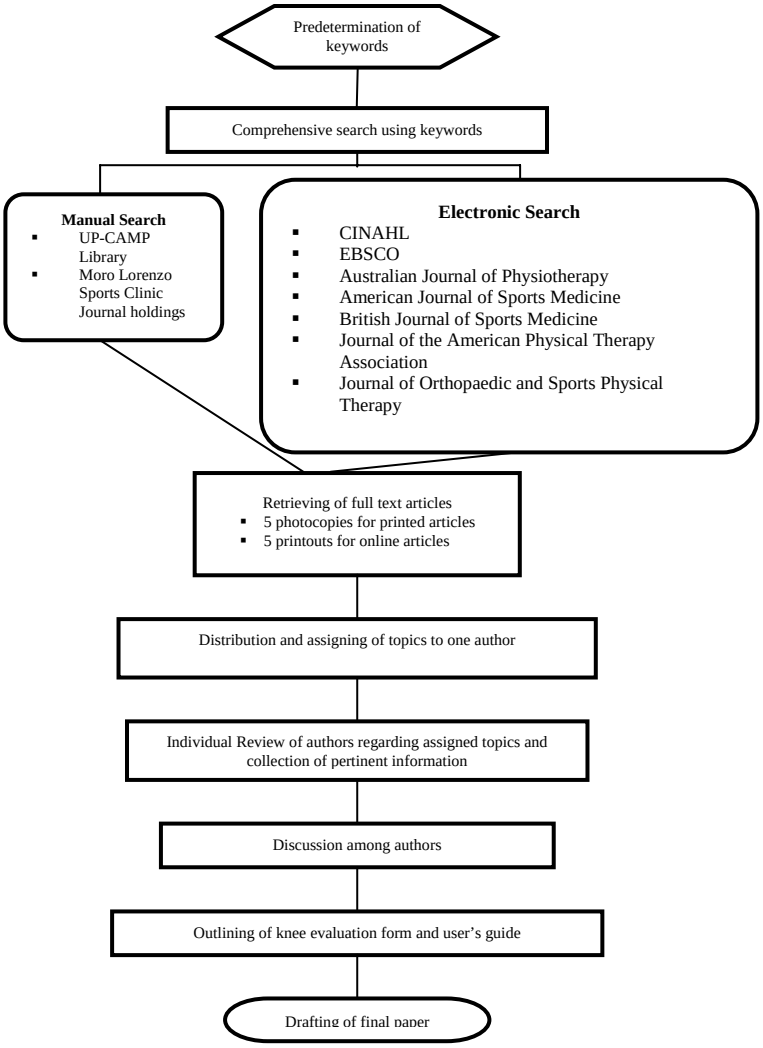

Figure 1. Flowchart for the methodology of the literature review.

reliability of the outcome measures being investigated. Information on ideal testing procedures, materials and equipment required and time taken to administer the test were also obtained. The reviewers then met to discuss the results of the reviews. Outcome measures with high ICC were identified and then their testing procedures recommended in order to standardize all testing procedures.

RESULTS AND DISCUSSIONS

A total of 20 articles were included in this study. Three articles utilized goniometry, 1 used heel height measurements, 5 investigated use of circumference measures, 5 focused on use of functional performance tests (FPTs), 4 studies utilized pain scale measures, and 2 studies employed the Cincinnati Knee Rating System (CKRS).

Objective Outcome Measures

Range of Motion Measurement

One of the more frequent complications of ACL reconstruction is loss of motion on the knee joint. In 1992, Harner, Irrgang, Paul, Dearwater & Fu found significant loss of motion after ACL-R accounting for 11.1% overall incidence in the study population. [11] The need for a valid and reliable assessment and measuring tool is of utmost importance for clinicians to assess the patient's status and to be able to design an appropriate plan of intervention as limitation of motion is an anticipated impairment in a significant number of patients after ACL-R.

Goniometry

Of the measurement procedures and equipment that can be used to measure joint range,

goniometry using a universal goniometer is commonly used in the Philippines owing to its ease of administration and the low cost of the instrument. Reliability is dependent on factors such as instruments used, methods of using the goniometer, passive or active movements and alignment with specified landmarks.[12 13]

Measuring active knee flexion in patients with knee joint restrictions using a universal goniometer showed high inter- and intra-tester reliability (*see Table 2*).[14 15 16]

Although the inter-tester reliability is high, the more recent study suggested that only one therapist should take measurements on a patient with knee deficiencies.[16] In the clinical setting where not only one therapist would take goniometric measurements of one patient, it is recommended that active knee ROM be taken at least three times and the mean be calculated for a higher inter-tester reliability.[13]

Several studies recommended detailed testing procedures for goniometry on the knee joint. In a study by Gogia, et al[14], the subjects were positioned lying on one side and placing the leg to be tested over a stabilizing board with pegs for standardization of the position. Rheault, et al[15] suggests having the patient lie on prone with both feet off the plinth and a roll of towel placed under the most distal part of the femur of the leg that is being tested, to allow for maximum knee extension. Patient is then asked to actively flex the testing knee.

The study by Brosseau, Balmer, Tousignant, O'Sullivan, Goudreault, Goudreault & Gringas.[16] investigated on the intra- and inter-tester reliability and criterion validity of the universal and parallelogram goniometers for measuring maximum active knee motions of patients with knee restrictions. In the said study, they measured active knee flexion while patient is in supine. To standardize measurement of maximum active knee flexion, Velcro™ straps were used to hold the knee in flexion during measurement. Landmarks used were the same as those recommended by Norkin & White.[17]

Of the three methods presented, the one by Brosseau et al[16] presented the best method of measuring active knee flexion because the other methods were tedious (construction of a stabilizing board) and contraindicated (open-chain kinematic knee flexion on prone) to post ACL-R patients.

Heel Height Measurement

Insidious knee joint range of motion limitation may not be easily detected using a universal goniometer.[12] It has been shown that computing for the arc tangent of the heel-height difference divided by the length of the lower limb (computed using the formula: subject's height in cm x 28%) will give a precise measurement of the degree of the knee flexion contracture.[12]

Due to limited access to journal articles regarding heel-height measurement, only one study was acquired on its reliability. This study concluded that 'heel-height difference is a valid method of documenting knee flexion contracture' due to a high degree of correlation shown by the heel-height measurement and the standard goniometry in measuring range of motion of the knee.

The study showed that heel height measurement has a positive correlation coefficient of moderate reliability (r = .75 - .78) tested with the patient in prone and patella on or off the table. The study concluded that patient positioning does not necessarily affect the validity and reliability of the said measurement.[12]

Circumference Measures

Circumference measurement has been used to determine muscle atrophy or joint effusion.[18] Shaw et al. cited studies indicating the capability of knee circumference measurement in detecting postoperative effusion following ACL-R. These studies further suggested using the unimpaired contralateral leg as a control for standardization of measurements for each subject.[8] High intra-rater reliability was demonstrated when data from previous studies on both normal and post-ACL-R knees were combined.[8]

In relation to measurement of muscle atrophy,

Table 2. Reliability and Validity of the Universal Goniometer

First Author	Intra-tester Reliability	Inter-tester Reliability	*Validity*
Gogia	-	r = .98; ICC = .99	r = *.97 - .98; ICC = .98 - .99*
Rheault	-	r = .87	-
Brosseau	*ICC = .997 (flexion)* *ICC = .972 - .985 (extension)*	*ICC = .977 - .982 (flexion)* *ICC = .893 - .926 (extension)*	*(Criterion)* *r = .975 - .987 (flexion)* *r = .976 - .985 (extension)*

r = Pearson product-moment correlation coefficient
ICC = intraclass correlation coefficient

Shaw and colleagues determined poor correlation between thigh circumference measures and

isokinetic quadriceps testing (r = -0.04) and a negative correlation between thigh circumference measures and quadriceps CT cross-sectional area (r = -0.23).[8]

Variations in procedures for obtaining limb girth were also increasingly observable among published literature.[18] [19] [20] [21] In the study by Matthews and St-Pierre[20] on the time course of spontaneous recovery of muscle torques following arthroscopic meniscectomy, five bilateral circumferential measurements of the thigh were taken every 5 cm from the superior border of the patella with the subject in supine. Arangio and colleagues investigated the relationship of thigh muscle size and strength after ACL-R.[19] Measurement of thigh circumference was accomplished with the patient in supine and the knee in 10º flexion. Thigh girth was then taken at points 6.0 cm, 12.0 cm, and 18.0 cm, proximal to superior pole of the patella.[19] In 1998, Walsworth et al[21] conducted a study on the prediction of 10 RM for short-arc quadriceps exercise from hand-held dynamometer and anthropometric measurements were obtained by measuring once the circumference of the thigh 20 cm proximal to the tibial tuberosity. In the same year, Ross and Worrell[18] researched on the changes in thigh and calf girth following knee injury and surgery and measurements were taken at 10 and 20 cm proximal to the superior patellar pole after this distance was measured and marked using an ink pen on the subjects' anterior thigh.

Despite the varying procedures in which girth measurements were obtained, only few studies on their validity exist.[8] Among the number of circumferential measurement procedures mentioned above, only the study by Ross and Worrell attempted to conduct a reliability testing prior to the descriptive study.[18] For each of the three girth measurement sites, intra-session ICC's were found to be 0.99. In addition, inter-session reliability testing where subjects were measured 7 days apart, yielded ICC's equal to 0.98 for each measurement site, thus demonstrating excellent reliability for both intra- and intersession testing.[18]

Functional Performance Testing

Functional performance tests (FPT) are important outcome measures due to their value in identifying or confirming functional limitations, assessing the progress of rehabilitation and determining the individual's readiness to return to their pre-injury level of activity.[22] FPTs are popular because they 'require minimal space, equipment, time and personnel for its administration in the standard clinical context'.[8]

FPTs provide a gross measure of lower extremity function which covers clinical attributes such as joint stability, pain, muscle strength and power, while incorporating elements of neuromuscular coordination, proprioception and agility.[22] Furthermore, FPTs simulates the forces experienced during a sport-specific activity under controlled clinical conditions.[2] More importantly, traditional outcome measures (e.g. knee joint laxity and isokinetic quadriceps muscle strength) compared to FPT's have weak to moderate associations with functional tasks, and more often insignificant.[2]

FPTs described in literature include: single and triple hop tests for distance, timed hop tests, vertical jump tests, shuttle runs, cross-over hop tests, figure eight running and stairs hopple test and among these, single and triple hop tests for distance have been exhaustively investigated in terms of reliability and validity.[8]

Reliability of FPTs are widely investigated, particularly the single and triple hop tests for distance. Reliability measurements range from moderate to high Intra-class Correlation Coefficients (ICC) (*see Table 3)*.

However, Shaw et al[8] points out key issues regarding studies on FPTs such as variations in procedure, use of mean of multiple trials or use of best score, and types of subjects recruited. Juris, Phillips, Dalpe, Edwards, Gotlin, and Kane[23] state that during testing, particularly in maximal hop for distance test, leg and arm swing should be eliminated in order to accurately assess the functional capacity of a specific limb. They also recommended that this be imposed on other FPTs.

Literature suggests two ways of computing the scores for FPTs. Greenberger and Paterno [24], Juris et al[23], Wilk et al[25] used the mean of scores as they were more concerned with limb symmetry. On the other hand, Clark et al[26], Hopper et al[22], and Risberg and Ekeland[7] chose to use the score of the last trial which is often the best score. Thus, it can be argued that warm-up and learning do not affect the score of each trial during the actual testing.

Verical FPTs measure leg power only in one direction (upward), while horizontal FPTs add a propulsion component to the vertical force or upward force that the knee would have to absorb[27]. Laboratory based kinematic and kinetic analyses have shown that in vertical FPTs, the knee

Table 3. Functional Tests: Reliability

First Authors	Outcome Measure	Subjects	ICC
Clark, N	Single hop for distance		
Bandy et al (1994)		Uninjured (n =18)	0.93
Bolga & Keskula (1997)		Uninjured (n =20)	0.96
Booher et al (1993)		Uninjured (n =18)	0.97
Brosky et al (1999)		ACL-R (n=15)	0.97
Greenberger & Paterno (1994b)		Uninjured (n =20)	0.96
Hu et al (1992)		Uninjured (n =30)	0.96
Kramer et al (1992)		ACL-R (n=38)	0.93
Paterno & Greenberger (1996)		Uninjured (n =20)	0.96
Paterno & Greenberger (1996)		ACL-R (n=13)	0.89
Worrell et al (1993)		Uninjured (n =26)	0.99
Bandy et al (1994)	Triple Hop for distance	Uninjured (n =18)	0.94
Bolga & Keskula (1997)		Uninjured (n =20)	0.95
Bandy et al (1994)	Crossover hop for distance	Uninjured (n =18)	0.90
Bolga & Keskula (1997)		Uninjured (n =20)	0.96
Goh & Boyle (1997)		Uninjured (n =10)	0.85
Hopper et al (2002)	6 m timed hop	ACL R (n=19)	0.96(reconstructed) 0.95(uninjured limb)
	Cross-over hop	ACL R (n=19)	0.98(reconstructed) 0.95(uninjured limb)
Clark et al (2002)	Adapted Cross-over hop	Uninjured (n=12)	0.94

contributes 49-56% of force required, while only 3.9% for horizontal FPT. Considering this, horizontal FPTs apparently do not demonstrate criteria-related validity. [2] In a study by Anderson and Foreman[28], they recommended the cross-over hop for distance which is a more sensitive FPT than other hop tests since it imposes both frontal and rotational forces on the knee in addition to the sagittal plane forces common with horizontal FPTs. Clark[2] in 1998, supported this finding when he concluded that crossover hop for distance demonstrated criteria-related validity.

Other authors attempted to test the validity of FPTs by investigating its relationship with traditional clinical outcome measures such as subjective knee scores, isokinetic testing[25] and knee extensor strength[24]. In the study by Wilk et al[25], they were able to identify a positive correlation between isokinetic tests with FPTs (e.g. single-leg hop testing) while subjective knee scores compare positively with ability to generate peak torque during isokinetic testing and single hop tests (timed hop tests and cross-over hop test) at specific peak torques only.

FPTs are easy to administer, it does not require sophisticated equipment except for a stopwatch. The rest of the materials are readily available in a regular physiotherapy clinic such as tape measure, athletic tapes, and markers. However Greenberger and Paterno[24], used additional material in their study. The researchers placed charcoal on the heel of the examinees' shoe during single leg hop for distance to provide markings on the floor as they land. This gave the assessors the chance to measure more accurately after each hop test.

Concerning data analysis, Limb Symmetry Index (LSI), has been described by literature to be valuable to clinicians due to its ability to 'quickly and easily' calculate injured limb deficits[2]. Using the uninjured limb as the control, LSI can be computed by dividing the score of the injured limb by the score of the uninjured limb then multiplying it by 100. 'Normal' knee function is defined by an LSI score of ≥ 85%. Concomitantly, Clark[2] provided three assumptions that clinicians recognize when applying this computations. Clark[2] cited the following assumptions: 1) the uninjured limb being the control is 'normal' in relation to the variables being tested in an FPT (e.g. joint laxity, muscle strength, proprioception and dynamic balance); 2) the control limb has not experienced detraining; 3) 'there is no effect of limb dominance on performance'. According to the author, the third assumption remains to be proven in the literature. Additionally, Clark[2] described that if the examinee has no previous pathology or detraining in the uninjured limb, the clinician can be confident in using it as a control when applying LSI during data analysis.

Subjective Outcome Measures

Pain Rating Scales

Clinicians recognize the multi-dimensional nature of pain. Of the different pain dimensions, emphasis is placed on patient-perceived pain intensity as an indicator of clinical change.[29] [30] The visual analog scale (VAS) and numeric pain rating scales (NPRS)

are the most commonly used self-report measures of pain intensity in clinical practice due to their ease of use and short duration of administration.[30]

VAS requires the patient to rate his/her perceived level of pain intensity by placing a mark on a 10-cm line with extreme anchors of 'no pain' to 'extreme pain'.[30] Scoring of the VAS is done by measuring the distance, in millimeters, from the 'no pain' anchor to the mark identified by the patient.[30] A major advantage associated with this tool is that it provides a 'continuous level data that can be analyzed using more powerful parametric statistics'.[8]

NPRS, on the other hand, involves asking the patient to quantify his/her degree of pain with a single number usually from zero to ten with zero being 'no pain' and ten corresponding to 'the worst pain ever experienced'.[8] Spandoni G, Stratford, P, Solomon P, Wishart L[31] found that older and less literate individuals have less difficulty completing this tool as compared to the VAS. Another advantage of NPRS is the flexibility of its administration. This method can be administered either verbally or in written form rendering it applicable even to visually impaired patients.

Criterion validity for the NPRS has not been established due to the absence of 'gold standards' for pain measurement.[30] However, both VAS and NPRS have been reported to be valid (concurrent and convergent) when correlated with each other or with other pain measurement instruments such as McGill Pain Questionnaire (MPQ) (*see Table 4.*).[30 32]

Test-retest reliability has been reported to be moderate to high (0.67-0.96) for NPRS and high for VAS (ICC=0.71-0.99).[30]

Cincinnati Knee Rating System

A number of subjective knee rating systems currently exists and is being used by clinicians in an attempt to measure or quantify current knee status. Among the many subjective knee-rating systems, the Cincinnati Knee Rating System (CKRS) is reported to most accurately determine the functional outcome of the knee.[32] CKRS is also unique in the sense that it has undergone meticulous validity and reliability testing.

In a study by Barber-Westin, Noyes, and McCloskey[33], they examined the validity and reliability of the CKRS in 350 subjects with uninjured, injured or ACL-R knees. They identified

Table 4. Validity and Reliability of NPRS and VAS (Kahl & Cleland, 2005)

Pain rating scale	Convergent Validity	Concurrent validity	Test-retest reliability
NPRS	0.79-0.95 (vs VAS)	-	mod-high (ICC=0.67-0.96)
VAS	0.30-0.95 (vs NPRS & MPQ)	Mod (0.71-0.78) (vs NPRS)	high (ICC=0.71-0.99)

Table 5. Summary of outcome measures and their reliability and validity

Outcome Measures	Reliability		Validity
Subjective			
Visual Analogue Scale for Pain	***Intersession:*** High		Convergent and Concurrent
Cincinnati Knee Rating System Score	***Intersession:*** High		Content and Construct
Objective			
Knee flexion (Goniometry)	***Intra Tester:*** High	***Inter Tester:*** High	No study available
Extension lag (Heel-height)	No study available		Convergent (Versus Goniometry)
Circumferential Measurement	***Intrasession:*** High	***Intersession:*** High	
Functional Performance Testing			
Single Hop Test for Distance	Moderate to High		Face; Convergent (Versus Isokinetic Strength Testing)
6m timed hop test	High		Face; Convergent (Versus Isokinetic Strength Testing)
Crossover hop test for Distance	High		Criteria-Related

that each item on the questionnaires demonstrate high test-retest reliability making it suitable for use in assessing groups of patients in different phases of recovery. The authors also reported that the questionnaire demonstrated good content validity, construct validity, and item-discriminant validity. This study used the updated version of the CKRS which included four categories namely: 1) symptom rating scales and patient perception scale; 2) sports activity scale, activities of daily living function scales, and sports function scales; 3) occupational rating scales and; 4) overall rating system. The last category was composed of subcategories that required the use of knee arthrometers. In this light, the authors recommend the use the older version of CKRS, which included only the first two categories.[34]In addition, Hopper et al[22] also determined the test-retest reliability of the CKRS in subjects approximately 12 months post ACL-R and reported a high ICC of 0.97.

Shaw et al[8] stated that even though CKRS permit computation of an overall score, objective items should be reported independently.

Summary of Outcome Measures
The following outcome measures were found to be valid and reliable (Table 5).

CONCLUSION

Based on the literature review, the following outcome measures were found to be valid and

reliable: subjective measures such as the visual analogue scale (VAS) and the Cincinnati Knee Rating System (CKRS) score, and objective measures such as goniometry, circumferential measurement, functional performance testing such as single hop test for distance, 6m timed hop test and cross-over hop test for distance. These measures are inexpensive, take an acceptable length of time to administer, acceptable for the patient and convenient for the clinicians to use. Except for the new version of the CKRS which requires an arthrometer, all of these measures require minimal equipment and may be administered in local rehabilitation clinics.

LIMITATIONS AND RECOMMENDATIONS

There were a limited number of studies available for this review and this may have affected the results. Therefore, a superseding study, particularly a systematic review may be done which includes a more comprehensive search of additional data databases, personal contacts with experts, articles not written in the English language, and unpublished articles locally and internationally. More importantly, a complementary study is recommended to determine the patients' level of comfort, ease of use and the validity of the devised form. This may also include studies on the perceptions of therapists regarding the sensitivity of the form and their recommendations, so that adjustments can be made accordingly.

Currently, a limited number of studies exist regarding the ideal timeframe for tests to administer. In agreement with Shaw et al[8], while these measures are high in reliability and have good validity, it should be taken during an ideal timeframe. The authors therefore recommend further studies to investigate the ideal time of administration of these tests. These will help propel physiotherapy in the Philippines towards a more accurate, consistent, valid, standardized and evidence-based practice.

The selected outcome measures were outlined in an evaluation form devised by the authors, including a user's guide on testing procedures, patient positioning, instruments to be used and materials required. Further studies are warranted to assess the clinical usefulness of this tool as well as its ability to detect changes in patients' status during rehabilitation.

REFERENCES

1. Chartered Society of Physiotherapy. (2005). Effective practice: Outcome measures. Available from: http://www.csp.org.uk/director/effective practice/outcomemeasures.cfm.
2. Clark NC. Functional performance testing following knee ligament injury. Physical Therapy in Sports. 2001; 91-105.
3. Encyclopedia of sports medicine and science [homepage on the internet]. Sydney: Internet Society for Sport Science.; c1998 [updated 2005 April 19; cited November 2005]. Available from: http://www.sportsci.org/encyc/aclinj/aclinj.html.
4. Freedman KB, D'Amato MJ, Nedeff DD, Kaz A, Bach BRJ. Arthroscopic anterior cruciate ligament reconstruction: A meta-analysis comparing patellar tendon and hamstring tendon autografts. American Journal of Sports Medicine. 2003;31: 2-11.
5. Bowditch, M. Anterior cruciate ligament rupture and management. Trauma. 2001; 3:249–261.
6. World Health Organization. ICIDH-2: International Classification of Functioning,

Disability and Health. Pre-final Draft,2000. Geneva.
7. Risberg MA, Ekeland A. Assessment of functional tests after anterior cruciate ligament surgery. The Journal of Orthopedic and Sports Physical Therapy. 1994;19:212-217.
8. Shaw T, Chipchase LS, WIlliams MT. A user's guide to outcome measurement following ACL reconstruction. Physical Therapy in Sport. 2004; 5: 57-67.
9. Currier DP. Elements of Research in Physical Therapy. 3rd ed. Baltimore, MD: Williams & Wilkins.
10. Kahl C. & Cleland JA. Visual analogue scale, numeric pain rating scale and the McGill Pain Questionnaire: An overview of psychometric properties. Physical Therapy Reviews. 2005;10:123-128.
11. Harner CD, Irrgang JJ, Paul J, Dearwater S, Fu FH. Loss of motion after anterior cruciate ligament reconstruction. American Journal of Sports Medicine. 1992; 20: 499-506.
12. Schlegel TF, Boublik M, Hawkins RJ, Steadman JR. Reliability of heel-height measurement for documenting knee extension deficits. American Journal of Sports Medicine. 2002;30:479-482.
13. Gajdosik RL, Bohannon RW. Clinical measurement of range of motion: Review of goniometry emphasizing reliability and validity. Physical Therapy. 1987;67:1867-1872.
14. Gogia PP, Braatz JH, Rose SJ, Norton BJ. Reliability and validity of goniometric measurements at the knee. Physical Therapy.1987; 67:192-195.
15. Rheault W, Miller M, Nothnagel P, Straessle J,Urban D. Intertester reliability and concurrent validity of fluid-based and universal goniometers for active knee flexion. Physical Therapy.1988;68:1676-1678.
16. Brosseau L, Balmer S, Tousignant M, O'Sullivan JP, Goudreault C, Goudreault M., Gringras S. Intra- and intertester reliability and criterion validity of the parallelogram and universal goniometers for measuring maximum active knee flexion and extension of patients with knee restrictions. Archives of Physical Medicine and Rehabilitation.2001;8:396-402.
17. Norkin CC, White DJ. (2003). The knee. In: Norkin CC, White D, editors. Measurement of joint motion: A guide to goniometry. Philadelphia: F.A. Davis Company; 2003. P221-240.
18. Ross CM, Worrell TW. Thigh and calf girth following knee injury and surgery. Journal of Orthopaedic & Sports Physical Therapy. 1998;27:9-15.
19. Arangio GA, Chen C, Kalady M, Reed JF. Thigh muscle size and strength after anterior cruciate ligament reconstruction and rehabilitation. Journal of Orthopaedic & Sports Physical Therapy. 1997;5: 238-243.
20. Matthews P, St-Pierre DMM. Recovery of muscle strength following arthroscopic menisectomy. Journal of Orthopaedic & Sports Physical Therapy.1996;1:18-26.
21. Walsworth M, Schneider R, Schultz J, Dahl C, Allison S, Underwood F et al. Prediction of 10 repetition maximum for short-arc quadriceps exercise from hand-held dynamometer and anthropometric measurements. Journal of Orthopaedic & Sports Physical Therapy. 1998;28:97-104.
22. Hopper DM, Goh SC, Wentworth LA, Chan DYK, Chau JHW, Wootton GJ et al. Test-retest reliability of knee rating scales and functional hop tests one year following anterior cruciate ligament reconstruction. Physical Therapy in Sports. 2002;3:0-18.
23. Juris PM, Phillips EM, Dalpe C, Edwards C, Gotlin RS, Kane DJ. A dynamic test of lower extremity function following anterior cruciate ligament reconstruction and rehabilitation. The Journal of Orthopedic and Sports Physical Therapy.1997; 26:184-191.
24. Greenberger HB, Paterno MV. Relationship of knee extensor strength and hopping test performance in the assessment of lower extremity function. The Journal of Orthopedic and Sports Physical Therapy. 1995;22:202-206.
25. Wilk KE, Romaniello WT, Soscia SM, Andrews JR. The relationship between subjective knee scores, isokinetic testing, and functional testing in the ACL-reconstructed knee. The Journal of Orthopedic and Sports Physical Therapy.1994; 20:60-73.
26. Clark NC, Gumbrell CJ, Rana S, Traole CM, Morrissey MC. Intra-tester reliability and measurement error of the adapted crossover hop for distance. Physical Therapy in Sports.2002; 3:143-151.
27. MaulderP, Cronin J. Horizontal and vertical jump assessment: reliability, symmetry,discriminative and predictive ability. Physical therapy in Sport. 2005;6:74-82.
28. Anderson M, Foreman T. Return to competition: functional rehabilitation. In:Zachazewski J, Magee D, Quillen W, editors. Athletic Injuries and rehabilitation. Philadelphia,PA; 1996. P229-261.
29. Jensen MP, McFarland CA. Increasing the reliability and validity of pain intensity measurement in chronic pain patients. Pain. 1993; 55:195-203.
30. Kahl C, Cleland JA. Visual analogue scale, numeric pain rating scale and the McGill Pain

Questionnaire: An overview of psychometric properties. Physical Therapy Reviews. 2005;10:123-128.

31. Spandoni G, Stratford, P, Solomon P, Wishart L. The evaluation of change in pain intensity: a comparison of the P4 and single-item numeric pain rating scales. Journal of Orthopaedic and Sports Physical Therapy. 2004; 34: 187-193.
32. Mawdsley RH, Moran KA, Connif LA. Reliability of two commonly used painscales with elderly patients. Journal of Geriatric physical therapy 2002. Available from: http://findarticles.com/p/articles/mi_qa4055/is_200201/ai_n9064372/pg_6
33. Barber-Westin SD, Noyes FR, McCloskey JW. Rigorous statistical reliability, validity and responsiveness testing of the Cincinnati Knee Rating System in 350 subjects with an injured, injured or anterior cruciate ligament-reconstructed knees. The American Journal of Sportsmedicine. 1999;27: 402-416.
34. Magee DJ. Knee. In Magee, DJ, editor. Orthopedic Physical Assessment. Philadelphia, PA: Saunders; 2002. P.661-763.