

[RESEARCH REPORT]

LYNNE PATTERSON STURGILL, PT, DPT, OCS¹ • LYNN SNYDER-MACKLER, PT, ScD, SCS, FAPTA²
TARA JO MANAL, PT, DPT, OCS, SCS³ • MICHAEL J. AXE, MD⁴

Interrater Reliability of a Clinical Scale to Assess Knee Joint Effusion

Joint effusion is an excessive accumulation of fluid within a joint capsule, indicating that the joint is inflamed or irritated. Knee joint effusion may be caused by injury, infection, or disease, or may be an acute consequence of surgical intervention.^{2,7,12}

Measurement of knee joint effusion is a component of a thorough physical examination of the knee. The amount of effusion and the speed of onset postinjury is helpful in establishing a diagnosis.¹² When an effusion is present, the function of the quadriceps can be impaired

and altered knee motion can occur.^{14,16} Knee joint effusion may mirror the health or irritability of the joint; therefore, monitoring effu-



sion is useful to determine appropriate progression of therapeutic exercises and activities and to monitor progress.

- **STUDY DESIGN:** Clinical measurement.
- **OBJECTIVE:** To determine the interrater reliability of a knee joint effusion grading scale in an outpatient orthopaedic physical therapy clinic.
- **BACKGROUND:** Knee joint effusion may indicate joint inflammation or irritation. Therefore, objective monitoring of effusion is important to decision making regarding patient prognosis and program progression. The clinicians in the authors' clinic use a modified stroke test to assess for knee joint effusion, which is operationally based on a 5-point grading scale.
- **METHODS:** Seventy-five patients (44 male, 31 female) receiving outpatient physical therapy for a unilateral knee problem, for whom effusion assessment was indicated, were tested. The subjects ranged from 16 to 65 years of age. Pairs of therapists graded the knee joint effusion using the clinical grading scale. A contingency table was constructed and analyzed using Cohen kappa

values to establish interrater reliability. Percent agreement was also calculated.

- **RESULTS:** The kappa value was 0.75, observed as a proportion of the maximum possible kappa, and the percent agreement was 73%. Fifty-four of 75 pairs of tests had perfect agreement. Only 5 had disagreement of 2 grades, and there were no disagreements of greater than 2 grades.

- **CONCLUSION:** These findings provide evidence to support the proposed clinical effusion grading scale as a reliable method to assess knee joint effusion between therapists in an outpatient orthopaedic physical therapy clinic in patients with unilateral knee dysfunction. Only 5 of 75 ratings resulted in disagreement that could result in different clinical decisions being made by the therapists. *J Orthop Sports Phys Ther* 2009;39(12):845-849. doi:10.2519/jospt.2009.3143

- **KEY WORDS:** hydroarthrosis, measurement, swelling, tibiofemoral

Despite the importance of knee joint effusion assessment, there is no "gold standard" evaluation method. In many instances, the assessment is a clinical observation,^{6,17} and the difficulty of accurately assessing the volume of the effusion in the knee has been reported.¹⁸ Fritz et al⁵ examined 2 clinical tests and interrater reliability of both tests was poor (fluctuation test, $\kappa = 0.37$; patellar tap test, $\kappa = 0.21$).

Circumferential measurements using a tape measure comparing the involved to the uninvolved extremity are also used to determine if joint swelling is present. The method has been touted to be a "cheap, fast, and easy-to-apply" measurement that is dependent on the accuracy of the clinician.⁹ It has been determined that circumferential measurements made 1 cm above the patella are more reliable than those made at midpatella.⁸ Inconsistencies in maintaining tape tension and measuring true circumference instead of ellipses, as well as challenges in measuring extremities with soft tissue variations (eg, synovial thickening or additional adipose tissue), may lead to measurement error.¹⁸

Magnetic resonance imaging and ultrasonography have a higher sensitivity in detecting joint effusion than clinical tests.^{6,15} Magnetic resonance imaging and ultrasonography are time consuming and costly, and demonstrate limited practical value for monitoring joint effusion in a

¹Staff Physical Therapist, University of Delaware Sports and Orthopedic Physical Therapy Clinic, Newark, DE; Adjunct Assistant Professor, Department of Physical Therapy, University of Delaware, Newark, DE. ²Alumni Distinguished Professor, Department of Physical Therapy, University of Delaware, Newark, DE. ³Director of Clinical Services, Department of Physical Therapy, University of Delaware, Newark, DE. ⁴Clinical Professor, Department of Physical Therapy, University of Delaware, Newark, DE. This study received exemption status from the University of Delaware Human Subjects Review Board because the data were originally collected as regular standard of care and analyses were performed on de-identified data. Address correspondence to Dr Lynne P. Sturgill, University of Delaware Physical Therapy Clinic, 053 McKinly Lab, Newark, DE 19711. E-mail: lsturkil@udel.edu

RESEARCH REPORT

clinical setting.⁹

Volumetry using water displacement as a means to assess limb volume is reliable.¹ Volumeters are not readily available, and the technique is time consuming due to filling the water tank, allowing time for the water level to stabilize after the patient is positioned, and changing the water between patients.¹ Optoelectronic imaging devices have been found to be as reliable as volumetry¹³; however, again, they are not readily available and difficulty with portability and expense may preclude clinical use.^{9,13}

In the University of Delaware Physical Therapy Clinics knee joint effusion is evaluated using a stroke test and quantified using a grading scale originally operationally defined by Michael J Axe, MD. This method has been considered a useful clinical monitor for establishing timelines for progression of exercise, functional activities, and testing in our clinics for the past 15 years. Neither reliability nor validity of this grading scale has been established. The purpose of this study was to determine the interrater reliability of this clinical measure of knee joint effusion in an outpatient orthopaedic physical therapy clinic. Establishing reliability is the first step in the validation process of this grading scale.

METHODS

Inclusion Criteria

SUBJECTS INCLUDED IN THIS STUDY were individuals referred to the University of Delaware outpatient physical therapy clinic for treatment of a unilateral knee dysfunction, for which effusion testing was deemed appropriate (eg, acute injury, postsurgery) by the treating therapist. No patient was excluded based on diagnosis, chronicity, or surgical status of the involved knee. This study received exemption status from the University of Delaware Human Subjects Review Board because the data were originally collected as regular standard of care, and analyses were performed on de-identified data.

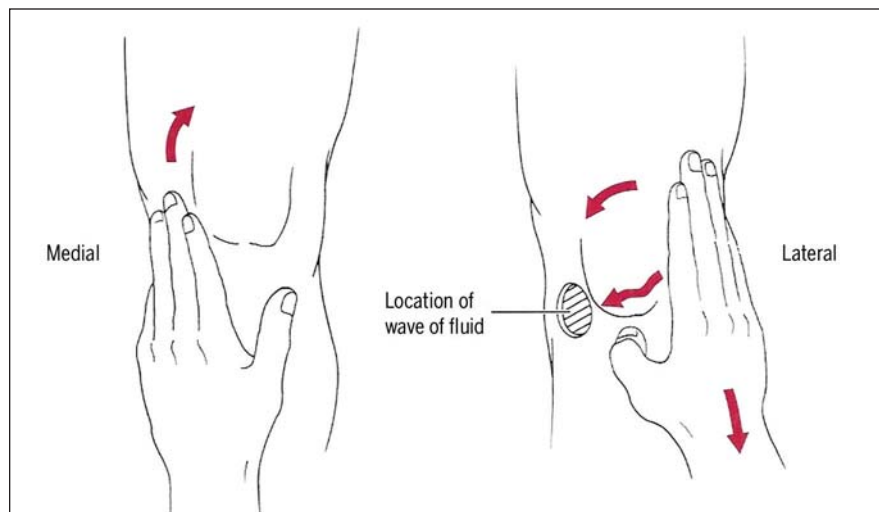


FIGURE. Diagram depicting the stroke test.¹² (A) Examiner strokes upwards from the medial joint line towards the suprapatellar pouch. (B) A downward stroke on the distal lateral thigh from the suprapatellar pouch towards the lateral joint line is performed; a wave of fluid is observed at the medial knee. Figure published with permission from *Orthopedic Physical Assessment*, 4th edition, David J. Magee, Knee chapter, page 726, Copyright Elsevier, 2002.

TABLE 1

EFFUSION GRADING SCALE OF THE KNEE JOINT BASED ON THE STROKE TEST

Grade	Test Result
Zero	No wave produced on downstroke
Trace	Small wave on medial side with downstroke
1+	Larger bulge on medial side with downstroke
2+	Effusion spontaneously returns to medial side after upstroke (no downstroke necessary)
3+	So much fluid that it is not possible to move the effusion out of the medial aspect of the knee

PROCEDURE

A STROKE TEST (FIGURE) IS PERFORMED with the patient in supine and with the knee in full extension and relaxed. Starting at the medial tibiofemoral joint line, the examiner strokes upward 2 or 3 times toward the suprapatellar pouch in an attempt to move the swelling within the joint capsule to the suprapatellar pouch. The examiner then strokes downward on the distal lateral thigh, just superior to the suprapatellar pouch, toward the lateral joint line. A wave of fluid may be observed within seconds on the medial side of the knee.¹²

Effusion of the knee joint is quantified using a 5-point scale (TABLE 1). A 0 grade is given when no wave is produced with the downward stroke. If the down-

ward stroke produces a small wave on the medial side of the knee, the effusion is graded as trace (ONLINE VIDEO 1); a larger bulge is given a 1+ grade (ONLINE VIDEO 2). If the effusion returns to the medial side of the knee without a downward stroke, the effusion is given a 2+ grade (ONLINE VIDEO 3). The inability to move the effusion out of the medial aspect of the knee is graded as a 3+.

Data for this study were collected in a university-based outpatient orthopaedic physical therapy clinic during a 5-month period. Nine therapists participated in the testing. All were board-certified clinical specialists in sports physical therapy and/or orthopaedic physical therapy, or were residents in an American Physical Therapy Association-accredited sports or orthopaedic residency program.

TABLE 2

A 5-BY-5 CONTINGENCY TABLE DEPICTING EFFUSION GRADES BETWEEN PAIRS OF THERAPISTS

	Rater 1				
	Zero	Trace	1+	2+	3+
Rater 2	Zero	3			
	Trace		6	2	2
	1+	1	4	24	2
	2+		2	4	15
	3+			3	6

All examiners received prior training in the application and grading of the effusion test as a component of their orientation to our clinic. An instruction sheet reviewing the operational definitions is readily available in the clinic. All therapists regularly used the scale, and no additional training was provided to the therapists prior to the collection of data. Over a 5-month period, when the primary treating therapist identified a patient for whom knee joint effusion testing was indicated, the therapist rated the effusion using the clinical grading scale and documented the grade on a score sheet. A second therapist then rated the knee joint effusion. Pairs of therapists were chosen by convenience (that is, their presence in the clinic at the time). Neither therapist viewed the other's testing or effusion score. The effusion grades of the respective pairs of therapists will be referred to as a set of ratings.

Data Analysis

A contingency table (TABLE 2) was constructed using the data collected between the pairs of therapists and was analyzed using Cohen's kappa values to establish reliability.³ Statistical analyses were conducted using SPSS for Windows, Version 10.0 (SPSS Inc, Chicago, IL) to establish level of agreement.

RESULTS

DATA WERE COLLECTED ON 75 patients, 31 female and 44 male. The average age of the patients was 31.4 years (SD, 13.4 years; range, 16-65 years). Thirty patients (40%) had undergone

knee surgery and 45 (60%) were nonsurgical cases.

Nonsurgical diagnoses included anterior cruciate ligament (ACL) injury, osteoarthritis (patellofemoral and/or tibiofemoral), posterior cruciate ligament injury, synovitis, medial meniscal tear, patellofemoral pain, and medial collateral ligament injury. Surgeries included ACL reconstruction, meniscectomy, proximal-distal patellar realignment, proximal patellar realignment, and chondroplasty.

The overall kappa value was 0.61, indicating substantial agreement between raters.¹⁰ The 95% confidence interval for the kappa values was 0.54 (lower limit) and 0.81 (upper limit), with an observed kappa of 0.68 and a standard error of 0.07. The maximum attainable kappa was 0.90, and the observed kappa as proportion of maximum possible was 0.75. Fifty-four of 75 sets of ratings had perfect agreement. Sixteen of 75 differed by 1 grade. Only 5 sets of ratings had a disagreement of 2 grades, and there were no disagreements of more than 2 grades. Percent agreement was 73%.

DISCUSSION

THE RESULTS OF THIS STUDY PROVIDE evidence that the clinical grading scale used in our clinic is a reliable method of assessing knee joint effusion. The scale offers an objective means of communicating about the status of the patient's knee joint among clinicians. The data provided by assessing effusion with this scale are a factor in clinical decision making regarding exercise progression

and functional testing.

The results of this study show substantial agreement between raters. Only 6% of the sets of ratings (5/75) in this study resulted in disparity of 2 grades. Only 1 of these patients was a postsurgical case (postchondroplasty), while the other 4 had recently diagnosed ACL injuries. Four of the grading differences involved the combination of 1 rater grading the effusion as a trace, while the other grading the effusion as a 2+. Two of these cases involved the same pair of raters on different dates. In the fifth case, involving a patient with an ACL injury, one therapist gave the effusion a 1+ grade while the other therapist graded the effusion as a zero.

In the authors' clinics, general rules regarding the care of patients with knee joint effusion have been established. Patients' exercise, agility, or functional program is not progressed when the effusion is graded as a 2+ or more. If a 2+ effusion persists, despite therapeutic measures to decrease the swelling (eg, limb elevation, effusion massage, compression wrapping with felt donut, decreased weight bearing) the referring physician is contacted regarding initiation of anti-inflammatory medications or potential aspiration of the joint. If the referring doctor is not easily accessible geographically, or, if the patient is self-referred, we consult with a physician who regularly performs knee aspiration. If an effusion increases more than 2 grades or appears when previously absent, activity is decreased to the level prior to the change in effusion. Activity is gradually resumed when the effusion has decreased to the previous level. The presence of no more than a trace effusion is one requirement for individuals with ACL injury to be able to participate in a screening process to determine if they are able to compensate neuromuscularly for a deficient ACL.⁴

The effusion grades and the resulting clinical decisions implemented are based on the absence of overt signs and symptoms of inflammation or infection. The presence of escalating pain, redness, an increase in local tissue temperature or fever with or

RESEARCH REPORT

without changes in the effusion grade may be indicative of a more serious condition and warrants immediate medical referral.

In light of these rules, the same clinical decision would have been made with regard to patient progression in 94% of the patients. As the clinical guidelines have not been validated, the implications of choosing the seemingly incorrect treatment path are unknown. Examining the operational definitions (TABLE 1), it is understandable that the difference between a trace ("small wave on downstroke") and a 1+ ("larger bulge") may occur. The lack of specific objectivity between these 2 grades leaves room for individual interpretation, thus the potential for decreased agreement; although this only occurred on 6 occasions in the current study. The errors appeared to be random. No patterns were found; no single mistake identified. If the errors were systematic or pervasive, the kappa scores would have been lower. In all cases, rating pairs of therapists who disagreed on the effusion grade for 1 or more patients were in agreement most of the time; therefore, it does not seem that misinterpretation of testing or grading was the problem. Thus these errors may not have been something that retraining or further training could have improved. In addition, if the number of categories were truncated, it is possible there would have been better agreement between raters; however, fewer categories would decrease the specificity of quantifying the effusion and thus potentially decrease the clinical usefulness. This effusion grading tool is robust, demonstrating very good reliability among many raters with varied years of experience. The results are especially relevant, considering that this study was performed in a busy clinical setting.

The reliability of this effusion scale ($\kappa = .61$ and $\kappa = .75$ observed as a proportion of the maximal possible kappa) is far better than that of the fluctuation test ($\kappa = .37$) and the patellar tap test ($\kappa = .21$).⁵ Circumferential measurements, while sometimes reliable, offer no clinically useful information on which to base decisions.⁸ Rating effusion, as in this study,

was easily implemented into a patient's visit; no special equipment was necessary, as with magnetic resonance imaging, ultrasonography, optoelectronic imaging, and volumetry. In addition, the method proposed is more cost effective than imaging or volumetry, as only minutes of hands-on time are required to perform the testing. Ease of application of this technique is an advantage; this method was readily incorporated into a typical physical therapy clinic visit by similarly trained physical therapists. No additional instruction was required. The results are generalizable to an outpatient orthopaedic population that includes patients with both nonsurgical and surgical conditions. No published measure of knee joint effusion offers the practicality, reliability, and clinical usefulness of this method.

There are some limitations to this study. All participating therapists were clinical specialists or eligible to test for clinical specialization, and therefore the reliability may not generalize to those who are not similarly trained. Classically, Cohen's kappa is calculated between a single pair of raters. While some¹¹ have suggested the overall kappa is considered to be less informative than if separate kappas are computed for each rater pair, the absence of systematic error and the consistency among pairs of raters make it unlikely that this is a meaningful difference.

CONCLUSION

THE CLINICAL EFFUSION RATING scale provides a reliable means of assessing knee joint effusion among therapists in an outpatient orthopaedic physical therapy clinic in patients with knee dysfunction. Seventy of the 75 paired tests resulted in effusion grades that would result in the same clinical decisions being made by different therapists. Further research is necessary to establish validity of this measure, as well as to determine the implications of a particular effusion grade or change in grade on functional testing and exercise and agility progression. ●

KEY POINTS

FINDINGS: A 5-point scale is presented to grade knee joint effusion through visual observation of fluid mechanics after performance of a stroke test. The results of this study provide evidence that the grading scale presented is a reliable means of assessing knee joint effusion.

IMPLICATION: This method to assess and quantify knee joint effusion is a simple measure that can assist in clinical decision making.

CAUTION: Further research is necessary to determine the validity of this measure.

ACKNOWLEDGEMENTS: *The authors would like to acknowledge the effort and support of the physical therapy staff at the University of Delaware Sports and Orthopedic Physical Therapy Clinic, especially Airelle Hunter-Giordano and Noel Goodstadt.*

REFERENCES

1. Bednarczyk JH, Hershler C, Cooper DG. Development and clinical evaluation of a computerized limb volume measurement system (CLEMS). *Arch Phys Med Rehabil.* 1992;73:60-63.
2. Brophy RH, MacKenzie CR, Gamradt SC, Barnes RP, Rodeo SA, Warren RF. The diagnosis and management of psoriatic arthritis in a professional football player presenting with a knee effusion: a case report. *Clin J Sport Med.* 2008;18:369-371. <http://dx.doi.org/10.1097/MJT.0b013e31815fa7a6>
3. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20:37-46.
4. Eastlack ME, Axe MJ, Snyder-Mackler L. Laxity, instability, and functional outcome after ACL injury: copers versus noncopers. *Med Sci Sports Exerc.* 1999;31:210-215.
5. Fritz JM, Delitto A, Erhard RE, Roman M. An examination of the selective tissue tension scheme, with evidence for the concept of a capsular pattern of the knee. *Phys Ther.* 1998;78:1046-1056; discussion 1057-1061.
6. Hauzeur JP, Mathy L, De Maertelaer V. Comparison between clinical evaluation and ultrasonography in detecting hyarthrosis of the knee. *J Rheumatol.* 1999;26:2681-2683.
7. Johnson MW. Acute knee effusions: a systematic approach to diagnosis. *Am Fam Physician.* 2000;61:2391-2400.
8. Kirwan JR, Byron MA, Winfield J, Altman DG, Gumpel JM. Circumferential measurements in the assessment of synovitis of the knee. *Rheumatol Rehabil.* 1979;18:78-84.

9. Labs KH, Tschoepel M, Gamba G, Aschwanden M, Jaeger KA. The reliability of leg circumference assessment: a comparison of spring tape measurements and optoelectronic volumetry. *Vasc Med*. 2000;5:69-74.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
11. Maclure M, Willett WC. Misinterpretation and misuse of the kappa statistic. *Am J Epidemiol*. 1987;126:161-169.
12. Magee DJ. *Orthopedic Physical Assessment*. Philadelphia, PA: Elsevier Health Sciences; 2006.

13. Man IO, Markland KL, Morrissey MC. The validity and reliability of the Perometer in evaluating human knee volume. *Clin Physiol Funct Imaging*. 2004;24:352-358.
14. Palmieri-Smith RM, Kreinbrink J, Ashton-Miller JA, Wojtys EM. Quadriceps inhibition induced by an experimental knee joint effusion affects knee joint mechanics during a single-legged drop landing. *Am J Sports Med*. 2007;35:1269-1275. <http://dx.doi.org/10.1177/0363546506296417>
15. Seltzer SE, Weissman BN, Finberg HJ, Markisz JA. Improved diagnostic imaging in joint diseases. *Semin Arthritis Rheum*. 1982;11:315-330.
16. Spencer JD, Hayes KC, Alexander IJ. Knee joint

effusion and quadriceps reflex inhibition in man. *Arch Phys Med Rehabil*. 1984;65:171-177.

17. Waddell DD, Marino AA. Chronic knee effusions in patients with advanced osteoarthritis: implications for functional outcome of viscosupplementation. *J Knee Surg*. 2007;20:181-184.
18. Wright RW, Luhmann SJ. The effect of knee effusions on KT-1000 arthrometry. A cadaver study. *Am J Sports Med*. 1998;26:571-574.



MORE INFORMATION
WWW.JOSPT.ORG

PUBLISH Your Manuscript in a Journal With International Reach

JOSPT offers authors of accepted papers an **international audience**. The *Journal* is currently distributed to the members of the following organizations as a member benefit:

- APTA's Orthopaedic and Sports Physical Therapy Sections
- Sports Physiotherapy Australia (SPA) Titled Members
- Physio Austria (PA) Sports Group
- Comit  de Fisioterapia Esportiva do Estado do Rio de Janeiro (COFEERJ)
- Canadian Physiotherapy Association (CPA) Orthopaedic Division
- Suomen Ortopedisen Manuaalisen Terapian Yhdistys ry (SOMTY)
- Hellenic Scientific Society of Physiotherapy (HSSPT) Sports Injury Section
- Chartered Physiotherapists in Sports and Exercise Medicine (CPSEM) and Chartered Physiotherapists in Manipulative Therapy (CPMT) of the Irish Society of Chartered Physiotherapists (ISCP)
- Gruppo di Terapi Manuale (GTM), a special interest group of Associazione Italiana Fisioterapisti (AIFI)
- New Zealand Sports and Orthopaedic Physiotherapy Association
- Portuguese Sports Physiotherapy Group (PSPG) of the Portuguese Association of Physiotherapists
- Orthopaedic Manipulative Physiotherapy Group (OMPTG) of the South African Society of Physiotherapy (SASP)
- Swiss Sports Physiotherapy Association (SSPA)
- Association of Turkish Sports Physiotherapists (ATSP)

In addition, *JOSPT* reaches students and faculty, physical therapists and physicians at **more than 1,400 institutions** in the United States and around the world. We invite you to review our Information for and Instructions to Authors at www.jospt.org and submit your manuscript for peer review at <http://mc.manuscriptcentral.com/jospt>.