Test-retest reliability of repeated knee laxity measurements in the acute phase following a knee trauma using a Rolimeter

Daniel Ericsson¹, Anna Hafsteinsson Östenberg¹, Erik Andersson², Marie Alricsson¹,²*

¹Department of Sports Science, Linnaeus University, Kalmar/Växjö, Sweden
²Swedish Winter Sport Research Centre, Department of Health Sciences, Mid Sweden University, Östersund, Sweden

INTRODUCTION

A noncontact injury with a rapid change in the direction is most often the cause of an anterior cruciate ligament (ACL) injury by “twisting the knee.” In most cases, the knee trauma results in pain and an inability to continue the activity. Knee effusion/hemarthrosis and decreased range of motion are immediate results following the trauma (Frobell et al., 2007; Micheo et al., 2010; Perera et al., 2013; Teske et al., 2010). Therefore, a correct diagnosis is essential in managing a patient with an ACL injury. It is crucial when creating a specific rehabilitation program and an important factor in minimizing the risk of recurrent episodes of instability and thereby lowering the risk of associated knee joint injuries such as meniscal tears and osteochondral fractures (Frobell et al., 2007; Micheo et al., 2010). The ACL diagnosis is mostly dependent on the history and physical examination (Howell et al., 2009; Panisset et al., 2012). In the initial evaluation of an acute knee trauma, the diagnostic value of the clinical test is low (Frobell et al., 2007; Renström, 2013). Frobell et al. (2007) showed a weak agreement between magnetic resonance imaging (MRI) and clinical tests in acute knee trauma. The risk of being misdiagnosed as having an “uncomplicated knee sprain” is relatively high, and a suitable re-evaluation plan in the clinical diagnosis of ACL injuries as well as a correct treatment are crucial, since it might decrease the development of osteoarthritis (Frobell et al., 2007). Previous studies have found the delays in getting a correct ACL diagnosis and the delay until referral to an orthopedic specialist to be very long, with a low diagnostic rate at the first clinical evalua-
tion and probably in all an underdiagnosed injury (Benjaminse et al., 2006; Bollen and Scott, 1996; Perera et al., 2013).

The Lachman test, pivot shift, and anterior drawer are three clinical tests commonly used to rule in or rule out an ACL injury (Howell et al., 2009; Micheo et al., 2010). The reliability of these tests, especially the pivot shift, has been questioned because of the variation in testing techniques and the subjectivity of the test (Kuroda et al., 2012). The examiner’s experience is of great importance in clinical testing without instruments, and the reliability has been questioned, at least regarding examiners with limited experience (Van Eck et al., 2013). The Lachman test and the anterior drawer test assess the anterior-posterior (AP) stability, AP laxity, and in the case of pivot shift, the anterior-lateral rotational stability (Howell et al., 2009; Micheo et al., 2010). In a study by Benjaminse et al. (2006) the Lachman test showed both high sensitivity and specificity in ACL injuries. It is viewed as one of the best tests when diagnosing ACL ruptures manually (Benjaminse et al., 2006; Howell et al., 2009; Perera et al., 2013; Renström, 2013).

Several mechanical testing devices have been developed to measure and quantify the AP laxity of the knee in order to detect ACL injuries (Pugh et al., 2009). The instruments are used as a more objective complement to the manual tests in the clinical evaluation of anterior knee laxity (Pugh et al., 2009). They are also used both preoperatively and postoperatively in the evaluation of ACL injuries as well as after reconstruction (Micheo et al., 2010; Panisset et al., 2012; Pugh et al., 2009; Van Eck et al., 2013). The combination of instrumented testing to verify a positive Lachman test provides the best findings for an ACL diagnosis, according to Howell et al. (2009). The most commonly used mechanical instruments in clinical practice are the KT 1000, KT 2000, Rolimeter, and Telos device (Temponi et al., 2015). In previous research, the KT 1000 Arthrometer has been the “golden standard” in measuring the anterior laxity of the knee and has shown good reliability and validity in measuring ACL deficiency, in several studies (Kuroda et al., 2012; Micheo et al., 2010; Pugh et al., 2009). The devices are suggested to be more precise in the later phases after a knee trauma and in chronic knee injuries with less pain and involuntary muscle contractions and should be combined with clinical tests and MRI in the evaluation of ACL injuries (Temponi et al., 2015).

A side-to-side difference of anterior knee laxity equal to, or more than, 3 mm indicates an ACL deficiency when comparing the injured knee to the uninjured knee. This cutoff is accepted and used in several studies regarding anterior laxity measurement (Hatcher et al., 2005; Micheo et al., 2010; Temponi et al., 2015). The arthrometer, Rolimeter, (Balasch et al., 1999) is used in research, evaluation of knee laxity pre- and postoperatively and in some sports injury clinics in the diagnosis and evaluation of ACL injuries by quantifying the AP knee laxity. Measurements with the Rolimeter serve as a complement to the history of the injury and the physical examination with manual testing in the evaluation of the AP knee laxity. The validity and reliability (intratester and intertester) of the Rolimeter have been examined in several studies, with mostly good results in both independent studies (Hatcher et al., 2005; Muellner et al., 2001; Papandreou et al., 2005) and in comparative trials (Balasch et al., 1999; Ganko et al., 2000; Papandreou et al., 2005; Schuster et al., 2004). In these trials, the Rolimeter is mainly compared with the KT 1000 arthrometer and the manual Lachman test and have shown at least equal ability to measure anterior laxity. However, the set-up of the measurement procedure varies between previous studies, as the position of the leg support differs slightly. Different hand positions are used both for the stabilizing hand and the hand creating the anterior force during the test. In addition, the rotation of the lower leg is not stabilized in most of the previous studies involving Rolimeter measurements (Balasch et al., 1999; Ganko et al., 2000; Hatcher et al., 2005; Muellner et al., 2001; Papandreou et al., 2005; Schuster et al., 2004).

In previous studies, Rolimeter measurements have been conducted on subjects with either healthy knees or on patients with a known ACL injury a long time period after the actual trauma (Balasch et al., 1999; Ganko et al., 2000; Hatcher et al., 2005; Schuster et al., 2004). Different factors, such as pain, effusion, and involuntary muscle contractions have been postulated to make the evaluation of knee laxity, in the acute phase of trauma, difficult by creating pseudo stability and false negative test results (Howell et al., 2009; Panisset et al., 2012; Van Eck et al., 2013). Nevertheless, the Rolimeter and the KT 1000 are used in clinical practice for diagnostic evaluations in the acute phase following a knee trauma (Pugh et al., 2009; Van Eck et al., 2013).

To date, there is a lack of knowledge regarding the test-retest reliability of Rolimeter measurements conducted in the acute phase following knee trauma, where effusion, pain, and a decreased range of motion generally remain. Therefore, the objective of the present study was to examine the test-retest reliability of the Rolimeter during the close time-phases following an acute knee trauma, where the endpoint measurement was used as a reference value. A secondary objective was to compare the test-retest reliability of the measurement procedure between the injured and the noninjured knee.
MATERIALS AND METHODS

Subjects

The present study is a prospective study of repeated measurements in acute knee trauma. A total of 15 subjects (5 female and 10 male subjects with an mean ± standard deviation [SD] age of 22.7 ± 6.9 years), with an acute knee trauma, participated in the study. Baseline information/characteristics were collected regarding age, sex, date of trauma, sport/activity, level of activity, type of trauma (contact/noncontact), right/left knee injury, and dominant/nondominant knee. In addition, evaluation factors regarding swelling (circumference measurement), active extension deficit (measured with a goniometer), and collateral ligament stability (degree of instability 0–3, varus/valgus) were documented. This information was obtained at the acute measurement point (i.e., measurements at baseline).

All subjects who met the inclusion criteria and attended our clinic were asked to participate in the present study. Subjects met the inclusion criteria if they had an acute substantial knee trauma, had swelling/effusion, or a history of it, after the trauma, received an initial evaluation 0–7 days after the trauma, and were within the range of 15–45 years old. Subjects were excluded from participation if there was a suspicion of possible fracture, substantial joint surface injury, or contusion injury solely. Subjects with pronounced pain or other reasons making it impossible or inappropriate to measure the knee laxity were also excluded. Written permission and approval were obtained from all subjects who accepted participation in the study.

Ethical approval was obtained from the Ethical Advisory Board in South East Sweden (Dnr EPK 311-2015) and the study was conducted in accordance with the Declaration of Helsinki for Human studies.

Measurements

The measurements were made with the arthrometer model, Rolimeter Aircast Europa, as seen in Fig. 1.

A firm leg support was provided for standardization of the knee angle between the different measuring points, to approximately 20°–30° of flexion. In the present study, the leg support developed for the KT 1000 Arthrometer was used and positioned slightly above the fold of the knee (Fig. 2). The supplied foot stand was also used to try to standardize the rotation of the legs by using a lateral support.

The Rolimeter was positioned proximally on the patella and distally on the tibia, fixed with a strap. To conduct the measurements, the moveable, sliding, reader-bar was centered on the tuberosity of the tibia. The frame of the device was then stabilized proximally by applying a gentle pressure with the thumb placed on the frame slightly below the reader bar and the other fingers placed on the patella part. The hand that applied the force was positioned dorsally on the calf, in line with the tuberositas tibiae and the reader bar. An anterior force was then applied with low velocity, and the test was executed using a maximum manual force. The white marker was used to measure the anterior translation of the tibia, in calibrated steps of 2 mm (Schuster et al., 2004).

Procedure

All measurements were conducted by one single tester, other than the author, who was blind to the overall results of the study and previous test results to avoid intertester differences. The tester...
in the present study had good knowledge of the Rolimeter and was experienced in the evaluation of ACL injuries both by clinical tests and instrumented testing. Measurements were made with a maximum manual force. The noninjured leg was examined first to avoid fear of pain or uncertainty of the testing procedure. Each subject was tested at three different time-points: baseline (0–7 days after the trauma), midpoint (3–4 weeks after the trauma), and endpoint (6–8 weeks after the trauma). At all three readings, three repeated measurements were made on each knee. The mean value of the two highest scores on each knee, at each of the readings, was then used in the statistical analysis. The measurements on the non-injured leg served as a control for the testing procedure in order to evaluate the reliability of the measurements over time.

Statistical analysis

The IBM SPSS Statistics ver. 24.0 (IBM Co., Armonk, NY, USA) was used to carry out our statistical analyses, and statistical significance was set at \( \alpha \leq 0.05 \). Data were checked for normality with the Shapiro–Wilks analysis and by visual inspection of Q-Q plots and are presented as mean±SD; the range is also presented in some of the cases. Two-way analysis of variance (ANOVA) \((2 \times 3)\) for repeated measures was applied for the comparison of the two knee conditions (i.e., injured vs. noninjured) and the 3 time-points of knee-laxity measurements. Bonferroni \( \alpha \) correction was applied to the ANOVA test. The sphericity assumption was tested using the Mauchly’s test. Eta squared \( (\eta^2) \) effect size (ES) was also reported for the ANOVA test. The systematic bias ±95% limits of agreement were evaluated for the baseline and midpoint versus the endpoint measurements of the injured and noninjured knee, respectively, by using Bland and Altman calculations (Bland and Altman, 1986) and using the endpoint measurement as a reference. The systematic bias for the respective comparison was tested with a one-sample \( t \)-test using a zero bias as the test value. Intra-class relationships between variables were assessed using linear regression and correlation analyses (Pearson). The strength of the correlation coefficients was classified according to Rowntree (1991). The relative and absolute typical errors for the comparisons were computed by taking the SD for the pair-wise mean differences (as a percentage and absolute values) divided by the square root of two.

RESULTS

No dropouts were present in this study. None of the subjects declined to participate and no one discontinued participation. All subjects had a substantial knee trauma with an effusion history after the trauma. In total, 12 knee injuries were noncontact injuries and 3 were contact injuries. Soccer accounted for 53% of the injuries (8 out of 15) and was clearly the dominant sport for trauma in the study (Table 1). The activity levels were predominantly at the moderate level \((n=8)\), but with two subjects competing at the elite level and three subjects competing in a youth league. Two of the subjects did not compete in any sport (Table 1).

All of the subjects had right knee dominance. In the present study, there were eight injuries to the dominant right knee and seven to the nondominant left knee. Twelve of the subjects in the study had no previous knee injury. Among the subjects, two of the subjects reported a previous injury to the same knee as the present injury, and one subject reported a previous injury to the noninjured knee.

Rolimeter measurement results

In total, 270 Rolimeter readings were made, ranging between 4–15 mm for the injured knee and 4–10 mm for the noninjured knee. All subjects had an average knee effusion of 1.3 cm (0.5–2 cm). Extension deficit was present in 11 of the 15 cases, with a mean deficit of 8° (0°–25°) at the baseline measuring point. The average typical errors for the three repeated anterior displacement scores measured at baseline, midpoint and endpoint, respectively, were 8% (0.4 mm), 7% (0.3 mm), and 6% (0.3 mm) for the injured knee, and 0% (0 mm), 6% (0.2 mm), and 6% (0.2 mm) for the noninjured knee. The anterior displacement scores, as an average of the two highest readings at each time-point, for the injured and the noninjured knee at the three different time-points are presented in Table 2. The Rolimeter scores were significantly different for the injured versus the noninjured knee (on average

<table>
<thead>
<tr>
<th>Table 1. Subdivisions of sport/activity and competitive level for the subjects in the study ((n=15))</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sport/activity</strong></td>
</tr>
<tr>
<td>Soccer</td>
</tr>
<tr>
<td>Floorball</td>
</tr>
<tr>
<td>Handball</td>
</tr>
<tr>
<td>Ice-hockey</td>
</tr>
<tr>
<td>Other activity</td>
</tr>
<tr>
<td><strong>Competitive level</strong></td>
</tr>
<tr>
<td>Elite</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Youth league</td>
</tr>
<tr>
<td>No sport</td>
</tr>
</tbody>
</table>

http://www.e-jer.org
Injured knee ≥ F Side-to-side difference F 9.3 ± 2.3 0.208 6.7 ± 1.8 0.11

Time-point effect F 9.7 ± 2.2 0.02 2.7 ± 2.9 0.775

Interaction effect F 2.8 ± 2.4

The results of the current study showed a high agreement in the test-retest reliability for the noninjured knee. The mean value for the anterior displacement was very similar at the three time-points of measurement, with the exact same range at all three occasions (4–10 mm). Repeated measures showed no significant changes over time for the measurements on the noninjured leg and were in line with previous studies (Hatcher et al., 2005; Muellner et al., 2001). The measurement reliability was almost perfect, with a very high agreement for the two different time-points versus endpoint (typical error of 0.2 mm and r ≥ 0.98).

Although the observed test-retest reliability was slightly lower for the injured knee compared to the noninjured knee, the mean AP displacement showed very small differences between the three time-points. In total, a high level of the test-retest reliability was revealed in the present study, which highlights the possibility of identifying ACL injuries during the acute phase following knee trauma. Almost identical correlations were found when analyzing the relationship between baseline to endpoint and midpoint to endpoint. Both the intraclass correlations were very high (r = 0.93 and r = 0.92), and the systematic biases were almost negligible (i.e., very close to zero). However, typical errors of ~0.7 mm together with limits of agreements that ranged between 4 mm, indicate some impreciseness of the measurement procedure on the injured knee, with the potential of setting a wrong diagnosis when side-to-side differences are close to the cutoff level (3 mm) for classifying a person as being ACL injured.

Only a few previous studies have evaluated the test-retest reliability of the Rolimeter measurement procedure. In a study by Muellner et al. (2001) 10 healthy subjects were measured two times in order to assess the intraobserver reliability of the Rolimeter method. Three different observers were used, and the study showed intraobserver reliability with correlation coefficients ranging between 0.55–0.72. The test-retest correlations observed in the current study were considerably higher for both the injured and the noninjured knee. Furthermore, Hatcher et al. (2005) evaluated the intratester reliability on noninjured and ACL deficient patients and showed a moderate/good correlation between the first and second measurements. In addition, Schuster et al. (2004) mea-

Table 2. Mean ± standard deviation anterior displacement (mm) for the injured and the noninjured knees at the three different time-points of Rolimeter readings (n = 15)

<table>
<thead>
<tr>
<th>Time-points</th>
<th>Injured knee</th>
<th>Noninjured knee</th>
<th>Side-to-side difference</th>
<th>Two-way ANOVA</th>
<th>F statistic</th>
<th>P value</th>
<th>ES (η²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>9.3 ± 2.3</td>
<td>6.7 ± 1.8</td>
<td>2.6 ± 2.5</td>
<td>Condition effect*</td>
<td>F(1,14) = 16.9</td>
<td>0.001</td>
<td>0.55</td>
</tr>
<tr>
<td>Midpoint</td>
<td>9.7 ± 2.2</td>
<td>6.9 ± 1.8</td>
<td>2.8 ± 2.4</td>
<td>Time-point effect</td>
<td>F(2,28) = 1.7</td>
<td>0.208</td>
<td>0.11</td>
</tr>
<tr>
<td>Endpoint</td>
<td>9.5 ± 2.7</td>
<td>6.9 ± 1.8</td>
<td>2.7 ± 2.9</td>
<td>Interaction effect</td>
<td>F(2,28) = 0.3</td>
<td>0.775</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Side-to-side difference, defined as the difference between the injured and the noninjured knee. F, P value, and effect size (ES) were obtained by a two-way analysis of variance (ANOVA) (2 × 3: knee conditions vs. time-points). Significant two-way ANOVA effect for condition* (i.e., injured versus noninjured knee).

DISCUSSION

The objective of the present study was to determine the test-retest reliability of the Rolimeter method in the acute phase of knee trauma and comparing the reliability between the injured and the noninjured knee. The repeatability of the measurement process was also investigated.

The present study involved subjects with an acute substantial knee trauma and the use of the Rolimeter in measuring the AP knee displacement in the early stages after the trauma. The study was conducted under the influence of pain, effusion, decreased mobility, and/or anxiety, which makes the evaluation of the knee laxity more difficult than testing healthy knees or in chronic ACL deficient knees (Hatcher et al., 2005; Panisset et al., 2012).

The results of the current study showed a high agreement in the test-retest reliability for the noninjured knee. The mean value
Fig. 3. Bland-Altman plots (left) and corresponding intra-class correlations (right) for the anterior knee-displacement (i.e., laxity) scores, using Rolimeter readings, at different time-points after acute knee trauma: (A) injured knee, endpoint vs. baseline, (B) injured knee, midpoint vs. baseline, (C) noninjured knee, endpoint vs. baseline, and (D) noninjured knee, midpoint vs. baseline. Measurements at baseline, midpoint and endpoint were performed 0–7 days, 3–4 weeks, and 6–8 weeks after the trauma, respectively. Bland-Altman plots represent the mean difference in the anterior knee displacement (i.e., systematic bias) ±95% (1.96 SD) limits of agreement (LOA) between the time-points. Dashed lines in the intraclass correlations represent the lines of identity.
measured noninjured and ACL injured subjects with the Rolimeter and reported intraclass correlations that ranged between 0.95–0.98 for the three different observers, when comparing the first versus the second measurement. Altogether, these results are in line with the test-retest reliability results reported in the present study.

Several previous studies involving Rolimeter measurements as well as other instrumented knee measurement systems address the possible errors and difficulties in the standardization of the test procedure (Daniel et al., 1985; Hatcher et al., 2005; Muellner et al., 2001). In the present study, an attempt was made to standardize the test procedure and thereby increase the reproducibility of the repeated measurements; therefore, a leg support to the KT 1000 arthrometer was used. This device has a firm leg support and also an adjacent foot plate with a lateral foot support, which helps to keep the rotation of the leg in a more consistent position when making measurements at different time points. Fiebert et al. (1994) found differences in the anterior displacement when measuring with the KT1000, in different rotation positions of the lower leg. In the present study, the knee support was placed just above the fold of the knee to prevent the knee support from putting any anterior translation force to the lower leg and thereby possibly creating a false low displacement value compared to the femur when measuring with the Rolimeter. Maximum force, which has been demonstrated in previous studies to give more reliable results for both the KT 1000 and Rolimeter measurements, was used (Balasch et al., 1999; Panisset et al., 2012; Torzilli et al., 1991; Van Eck et al., 2013). The position of the feeler was also controlled to be centered over the tibial tuberosity. The proximal hand of the tester was placed with the thumb just distally to the feeler bar, on top of the frame, and the other fingers on the patella part of the Rolimeter. This technique was used to control the position of the Rolimeter to avoid the risk of a medial or lateral slide, which would alter the measurement position. The distal hand, creating the anterior force to the tibia was held with all fingers dorsally onto the calf, based on previous recommendations by Ganko et al. (2000) and Schuster et al. (2004).

The testers experience has previously been shown to play an important role for the reliability of the measurements. In the study by Muellner et al. (2001) three different observers were used when examining the intraobserver reliability, and the study showed intraclass correlation coefficients of $r=0.55$ (between first and second reading) for the inexperienced tester and $r=0.65$ and 0.72 for the experienced testers. In this study, an experienced tester performed all the measurements which has to be considered when interpreting the results and comparing with previous studies.

In the current study, three measurements were made on both knees at each respective time-point, and the mean score of the two highest values were used for data analysis. In previous studies, the mean score of all three measurements has mostly been used (Balasch et al., 1999; Ganko et al., 2000; Muellner et al., 2001). Hatcher et al. (2005) have argued that only the highest score should be used because the only way that the value could be less is due to involuntary muscle contractions or other patient-related factors. In visual analysis of the data in the present study, there were occasional Rolimeter scores that differed from the other two values at the same knee at the same time point. In these cases, the first attempt differed, and the second and third readings were closer and higher. This gives strength to the use of the two highest values in determining the AP displacement in acute knee injuries. The endpoint measurements were used as reference scores in this test-retest design. The endpoint was 6- to 8-week posttrauma and considered to be a point in time comparable to previous studies on healthy and known ACL deficient subjects, which has shown good reliability results for accurately quantifying anterior knee laxity (Balasch et al., 1999; Ganko et al., 2000; Hatcher et al., 2005; Muellner et al., 2001; Schuster et al., 2004). These endpoint measurements were considered to be performed at a time-point where the effusion, extension deficit, pain, and other factors related to the acute phase of the trauma did not markedly affect the measurement results.

The present study provides important knowledge regarding the test-retest reliability of repeated measurements with the Rolimeter in the acute phase of a substantial knee trauma and the subsequent changes, or lack of changes, in knee laxity during the first 6–8 weeks after trauma. The results show rather small differences in the anterior displacement between the different time-points for both the injured and the non-injured knee. One could assume that pain, limited knee extension, effusion, or enhanced muscle activity would confound the measurements, at least during the acute time-point. However, it seems that similar results can be generated during the first week after trauma when comparing with the endpoint measurements conducted 6–8 weeks following the trauma. In addition, none of the subjects demonstrated any pain provoked by the measurements that interfered with the position of the leg support, and the position seemed to be relaxed with a knee angle exceeding the limitation of extension. Also, the effusion obviously did not affect the measurements. Therefore, all these factors that substantially affect the active and passive angular movement after a trauma may not restrict the anterior translation of the knee to any greater extent.
In the present study, 15 subjects were evaluated, with ten male and five female subjects, and soccer was the dominant sport in generating the injury. Although the test-retest reliability results support the use of the Rolimeter in quantifying anterior knee laxity also during the acute phase following a substantial knee trauma, the results have to be interpreted with some caution. For instance, the side-to-side difference between the injured and the noninjured knee was not displayed or discussed in the current study and related to the fact that one subject reported previous injury to the contra lateral knee and two subjects to the injured knee. Moreover, the predictive value of the Rolimeter, in detecting ACL injuries, was not assessed since data on the ultimate diagnosis were not accessible and hence not possible to analyze. Finally, further research should combine acute measurements with the Rolimeter and MRI for determining and highlighting the predictive value of the Rolimeter in detecting ACL injuries in the acute phase, following a substantial knee trauma.

In summary, the Rolimeter method was observed to be reliable, and it showed very high test-retest reliability for the noninjured knee, and more importantly, the test-retest reliability was also high for the injured knee. Although a noticeably higher reliability was observed for the measurements on the noninjured knee, the reliability results on the injured knee were also high but indicate a risk for misclassification when side-to-side differences are close to the cutoff level. However, instrumented measurements should always be used and interpreted in combination with a thorough injury history and manual clinical tests. Altogether, these results support previous studies that suggest the Rolimeter measurement method as being reliable for determining the AP displacement of the knee and that the current method also is reliable during the acute phase directly following a knee injury.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

REFERENCES


