The reliability of commonly used arthroscopic classifications of ligamentum teres pathology

Brian M. Devitt*, Bjorn Smith, Robert Stapf, Suenghwan Jo and John M. O’Donnell

Department of hip arthroscopy, Hip Arthroscopy Australia, 21-23 Erin Street, Melbourne, Victoria 3121, Australia
*Correspondence to: B. M. Devitt. E-mail: bdevitt@hotmail.com
Submitted 22 August 2016; revised 15 December 2016; revised version accepted 2 February 2017

ABSTRACT

The importance of the ligamentum teres (LT) in the hip is increasingly being recognized. However, the incidence of LT tears in the literature is extremely variable. Although classification systems exist their reliability in classifying LT pathology arthroscopically has not been well defined. To determine the inter- and intra-observer reliability of two existing classifications systems for the diagnosis of LT pathology at hip arthroscopy. Second, to identify key pathological findings currently not included. Four experienced hip-arthroscopists reviewed 40 standardized arthroscopic videos. Arthroscopic findings of the LT were classified using the Gray and Villar (G&V) and descriptive classification (DC). Reviewers were asked to record other relevant pathology encountered. Inter- and intra-observer reliability was defined using Fleiss-Kappa and Cohen-Kappa statistics. Both classifications demonstrated fair inter-observer reliability. The intra-observer reliability for G&V was moderate-to-substantial and for DC was slight-to-moderate. An absolute agreement rate of 10% (G&V) and 37.5% (DC) was found. Differentiation between normal, and partial or low-grade tears was a common source of disagreement. The prevalence of LT pathology was 90%. Synovitis was the most common diagnostic finding that was not included in either classification system used in this study. Arthroscopic classification of LT pathology using the G&V and the DC demonstrated only fair inter-observer reliability. The major discrepancy in interpretation was between normal, and partial or low-grade tears. The presence of synovitis was not in either classification but was considered an important arthroscopic finding. Thorough arthroscopic scrutiny reveals the prevalence of LT pathology is higher than previously reported.

INTRODUCTION

The ligamentum teres (LT) has gradually assumed a greater role in hip arthroscopy as it has emerged as a potential generator of hip pain and its role in hip stability has been reassessed (1–7). The concept of microinstability has developed in recent times and the LT is believed to play a key stabilizing role in the kinematics and stability of the hip (5, 8–12). The LT has been found to not only contain free nerve endings but it has been shown to have a load-bearing capacity similar to that of the anterior cruciate ligament (11, 13).

Injury to the LT typically occurs when the hip is placed in extreme positions and is subjected to forcible loads. The injury pattern may be a complete rupture or partial tearing (4, 14–16). In addition, the LT, given its location as an intra-articular structure may be subjected to synovitis and degenerative change over time (17, 18). The prevalence of LT pathology has been reported to vary considerably between studies and has been reported as being as low as 4% and as high as 89% (3, 15, 19–21). Aside from traumatic events, it has been reported that increased age (>30 years) and a low lateral coverage index (LCI) (centre edge angle minus acetabular inclination) are risk factors for the development of LT pathology (16).

The inclusion of the LT in a list of differential diagnoses for hip pain has prompted clinicians to seek improved methods of reaching a diagnosis of injury. Specific physical examination techniques to elicit pain in the LT have been found to be both sensitive and specific, and have a high predictive value (22). MRI, while less sensitive has been
shown to also have some utility in reaching a preoperative diagnosis of injury (20, 23). MRI arthrography with traction has been shown to have greater levels of sensitivity and specificity for both low grade and high-grade tears compared with MRI alone (24). Notably, in all of these studies arthroscopy has been used as the gold standard to determine the presence or absence of LT pathology.

Therefore, the purpose of this study was to investigate the inter- and intra-observer reliability of arthroscopic assessment of the LT using the G&V and DC. The hypothesis was that neither classification would demonstrate substantial reliability. A secondary aim was to explore if there were any deficiencies in the current classification systems, and was there consensus amongst experienced surgeons on treatment options for LT pathology.

MATERIALS AND METHODS

A reliability study of arthroscopic findings of LT pathology was performed. A consecutive series of 40-patients included in the study underwent arthroscopic treatment for presumed intra-articular hip lesions. The senior author recorded a short video of a systematic diagnostic assessment of the LT, which included dynamic manoeuvres (internal and external rotation of the hip) to view the full extent of the ligament. In addition, an arthroscopic examination with an articulating probe (E-flex, Smith and Nephew, Andover) was also carried out to further assess the integrity of the LT.

A hip arthroscopic procedure was performed with the patient in the lateral position, using a mid-trochanteric and anterolateral portals (25). The standardized videos were obtained through the mid-trochanteric viewing portal using a 70° arthroscope, while the anterolateral portal was used for arthroscopic probing.

All videos included a standardized arthroscopic examination focusing on the anterior and posterior bundles of the LT in full internal and external rotation.

A consecutive series of 45 videos was taken prospectively with no specific inclusion or exclusion criteria in order to provide a true representation of a normal case mix seen in everyday practice. Cases were excluded for poor quality videos, inadequate visualization or inadequate probing of pathological lesions.

Included amongst the videos were several normal cases, partial LT tears, complete LT tears and degenerative tears. Forty patients were included in the study. Videos were generally 30–60 s in length.

Four experienced international surgeons with a speciality interest in hip arthroscopy, performing >100 hip arthroscopy cases per year participated in the study. The surgeons were informed in detail about the two classification systems. The videos were uploaded to a private YouTube site to facilitate remote viewing internationally. The surgeons (raters) were asked to classify the LT according to the Gray and Villar (G&V) (18) and descriptive classification (DC) system, proposed by Botser et al. (15) (Tables I and II). Given that these classification systems only determine pathological damage, the raters were instructed to choose normal if they deemed there was no pathology visible. In addition, the raters were asked to detail what other pathology was present that was not accounted for in each of the classification systems. Finally, they were invited to offer the treatment they would recommend in each case. The raters were not given any clinical or radiological information about the cases.

Inter-observer analysis was calculated by using each read by the individual reviewer; this included comparison of each of the 40 readings with each of the three other raters. The intra-observer error was calculated for two raters, which was based on comparison of the read of the videos on two separate occasions 8 weeks apart.

| Table I. The G&V classification system |
|--------------------------------------|
| Type       | Pathology                  |
| I          | Complete rupture           |
| II         | Partial rupture            |
| III        | Degenerative rupture (complete or partial) |

| Table II. The Descriptive Classification—Botser and Domb |
|---------------------------------------------|
| Grade | Pathology               |
| 1     | Low-grade tear (<50%)   |
| 2     | High-grade tear (>50%)  |
| 3     | Full-thickness tear (100%) |

Descriptive statistics are presented as numbers and frequencies as appropriate, with the absolute agreement defined as agreement between each of the four raters. Given that more than two raters were involved in both the assessment of the classification and the adequacy of the classification, the Fleiss Kappa statistic was estimated, with 95% CIs also presented. The Kappa values were interpreted as suggested by Landis and Koch (26). The intra-observer rating conducted by the two raters across the G&V and DC were then assessed by the Cohen’s Kappa statistic. All tests were two-tailed and a P values of <0.05
was considered to indicate statistical significance. Statistical analysis was performed using Stata version 12.1 (StataCorp LP, Texas, USA). The $j$-values were classified according to the following recommend categories: 0.81–1.0, excellent agreement; 0.61–0.80, substantial agreement; 0.41–0.60, moderate agreement; 0.21–0.40, fair agreement; and 0–0.20, slight agreement.

**RESULTS**

Table III shows the distribution of LT findings as classified by G&V and the DC. The overall prevalence of LT pathology was on average 90% (range 85–95%). The most common classification according to the G&V was partial rupture of the LT, which was diagnosed in on average 73.75% of cases (range 60–85%). The least common classification by G&V classification was a complete rupture, which was diagnosed in two cases (5%) by a single rater.

The most common reading using the DC was a low-grade tear (56.25%); however, this was highly variable with a range of 47.5–70%. The presence of a complete tear was diagnosed in one case (2.5%) by each of the raters.

**Inter-observer reliability**

G&V: The Fleiss-Kappa value was 0.39 (95% CI: 0.150–0.598), which was categorized as a fair agreement. The absolute agreement of all four raters was 10% (4/40 cases).

DC: The Fleiss-Kappa value was 0.384 (95% CI: 0.213–0.569), which was categorized as a fair agreement. The absolute agreement of all four readers was 37.5% (15/40 cases) (Table IV).

**Intra-observer reliability**

G&V: The Cohen-Kappa value of raters 2 and 4 was 0.737 (95% CI: 0.394–1.000) and 0.590 (95% CI: 0.307–0.833), respectively; these values were defined as substantial and moderate agreement, respectively.

DC: The Cohen-Kappa values for raters 2 and 4 were 0.393 (95% CI: 0.109 to 0.625) and 0.453 (95% CI: 0.195 to 0.666), respectively; the reliability was determined as slight and moderate, respectively (Table IV).

**Additional pathology noted**

All four reviewers added the presence of synovitis as a diagnosis not contained within the classification systems (Fig. 1). Two raters commented on the presence of impingement of the LT with the margin of the acetabular fossa (Fig. 2), while two other raters mentioned chondral damage on the femoral head as a key finding (Fig. 3).
Finally, two separate raters suggested attenuation and hypertrophy of the LT were important features to note.

**Treatment options**
All four raters suggested radiofrequency shrinkage or debridement of the LT for the treatment of partial and degenerative tears and low- and high-grade tears. Two reviewers recommended synovectomy in cases of synovitis, while the other two recommended radiofrequency ablation. Two of the four reviewers recommended LT reconstruction in the cases of high-grade partial tears or complete tears, one suggested debridement, while the final reviewer offered no treatment.

**DISCUSSION**
The main finding of this study was that the G&V and DC, used to diagnose LT pathology, demonstrated only fair inter-observer reliability amongst an international group of experienced hip arthroscopy surgeons. The prevalence of LT pathology, as determined by each of the four raters within this consecutive cohort of 40 patients, was very high with an average of 90% (range 85–95%). Partial rupture of the LT (73.8% of cases) and low-grade tears (56.3% of cases) were the most common grades of injury according to the G&V and DC, respectively. All reviewers advised that synovitis was an important pathology to note, and one that was not covered in either classification system. Radiofrequency ablation was the treatment option that was overwhelmingly recommended for partial tears, both low- and high-grade, and degenerative tears. This study demonstrates that significant subjectivity exists in the current classification systems for diagnosis of LT pathology.

Although a number of studies in the literature have commented on the presence of LT pathology, there has been a wide variation in results. Byrd and Jones reported prevalence rates of 15 and 25% in two separate studies focusing on a general patient population and an athlete specific population, respectively (4, 27). Although these results were very useful in highlighting the LT as an important source of pathology, they are limited in providing further detail on the specific damage encountered; in the study of hip arthroscopy in the athlete (27), the authors only commented on ‘LT damage’, whereas the general population study reported on rupture of the LT secondary to a known traumatic episode (4). Botser et al. (15), in a comprehensive study detailing specifically LT pathology...
and symptomatology in a large patient population, reported a prevalence rate of 51% in a cohort of 558 hip arthroscopies, using the G&V classification and introducing their own DC. A more recent study by Chahla et al. (21), reporting on a cohort of 2213 patients undergoing hip arthroscopy for FAI, revealed a prevalence rate of LT pathology of 89.5%. These figures are similar to this study of 40 consecutive patients, in which the prevalence of LT tears was on average 90% (range 85–95%). It is important to consider why this vast disparity in prevalence rates exists.

Botser et al. (15) suggested that the substantial increase they reported from previously reported data was due to an increased awareness of LT pathology as well as the inclusion of low-grade tears. This is most certainly the case in this study, given that defining a pathological grade was the primary aim of the arthroscopic assessment. In addition, a very detailed assessment of the LT was carried out in each video, which included dynamic manoeuvres (maximal internal and external rotation of the hip) and probing of the LT bundles with an articulated probe. It is interesting to note that in the study by Chahla et al. (21), a similarly detailed assessment of the LT was made, which is likely to increase the rate of detection of pathology. This level of assessment is probably not always performed routinely.

A further reason for the high prevalence of LT pathology may be explained by the lack of a definition of what constitutes a normal LT in either classification. This is evidenced in this study by the fact that there was only absolute agreement on two cases as having as being normal. Furthermore, the range of values of normal reads varied from 5 to 15%. One of the attractive aspects of the G&V and DC is their simplicity, but the problem is that neither classification gives any guidance about what constitutes the minimum ‘pathology’ to be called abnormal.

It is obvious from the low inter-observer reliability that surgeons interpret tears differently, but why is this the case? The major discrepancy in interpretation was found between what constituted normal, and partial or low-grade tears. The reason for this may be explained by the structure of the LT. It is formed as consolidation of mesenchymal tissue, which takes its broad origin at the transverse ligament of the acetabulum to its foveal insertion (2, 3). A layer of synovium typically covers the ligament. The presence of the synovium can aid but also hinder the diagnosis of LT pathology. Synovial folds may mimic partial tears but they may also hide a significant partial tear beneath. However, the presence of periligamentous synovitis may herald an injury to the LT or point specifically to a site of impingement against the acetabular fovea. Also, it is not always possible to view the anterior bundles of the LT without performing dynamic manoeuvres. The ability to adequately probe the ligament, particularly with a non-articulated device, is often restricted by the prominence of the femoral head. All of these factors would suggest that a cursory glance is probably not sufficient in establishing a thorough diagnosis of LT pathology. The authors would recommend, at the very least, performing dynamic external and internal rotation manoeuvres to visualize both the anterior and posterior bundles of the LT. If there is significant synovitis, it may be beneficial to use a probe to look beneath the synovium to assess the integrity of the ligament. This is supported by the fact that each of the raters defined the presence of synovitis as a significant finding, but one that was not included in either classification.

Recent literature has suggested that there are a number of factors that increase the likelihood of developing LT pathology. Domb et al. (16), in their study of 462 hips, examined the relationship between non-traumatic LT tears and acetabular radiographic architecture. Using the LCI they determined that patients with a low LCI (<19°) were 1.74 times more likely to have LT tears than those high LCI hips (>34°) (16). Chahla et al. (21) revealed that there was a higher prevalence of LT tears in women, patients with a lower centre edge angle, and isolated cam or pincer deformity. There was also an association with hip laxity and chondral defects of the femoral head. Notably, chondral defects of the femoral head were considered an important arthroscopic finding by two raters in this study. Further, O’Donnell et al. (22) have also previously demonstrated a novel ‘LT test’ is effective in assessing the presence of LT tears with moderate-to-high interobserver reliability. The association with hip capsular laxity would suggest that an examination of joint hypermobility might also be a relevant preoperative assessment (21). The key to making a diagnosis is to first consider it. Therefore, an awareness of these risk factors for injury should prompt the surgeon to scrutinize the LT for pathological damage.

This study has identified that the DC and G&V classification systems have gaps. A recent, more detailed classification system has been proposed by Porthos Salas and O’Donnell (28) in an attempt to fill some of these gaps. Indeed, many of the additional findings that the raters in this study detailed are listed, including synovitis, impingement of the LT on the acetabular wall, and femoral head lesions. However, the classification is more complex with six grades as opposed to three, does not offer a definition of normal and the clinical utility has yet to be determined.

But, why is a classification needed at all? A classification may be descriptive, prognostic, and it may guide treatment.
Based on the findings of this study, it would appear that neither classification offers much additional guidance to treatment. The recommendations for treatment were remarkably similar between each of the raters. They all proposed debridement or the use of radiofrequency ablation for the treatment of partial and degenerative tears and low- and high-grade tears. In the setting of synovitis, they all recommended some form of debridement or radiofrequency ablation. If the LT was deemed normal, no treatment was proposed. It was of some interest to note that two surgeons suggested a LT reconstruction, which likely reflects their expertise; one surgeon recommended it twice for complete tears, while the other recommended it for a high-grade partial tear on three occasions. This finding would reflect the lack of consensus regarding the indications for LT reconstruction and, importantly, the absence of long-term results. The current literature is deficient in this area and is restricted to surgical technique and short-term case series with small numbers (7, 29–32). Further studies on the long-term effectiveness of LT reconstruction are certainly required.

The authors acknowledge that there are several potential limitations to this study. Classification of LT pathology by viewing recorded intraoperative videos fails to provide the reviewer with tactile feedback. The classification of disease based on arthroscopic videos may be a source of bias. Although the reviewers in this study were aware of the purpose of the study, they were not provided with other clinical or radiological information about the patients. This may have heightened their awareness of potential for LT pathology, particularly the presence of abnormal bony morphology. The arthroscopic visualization and probing could have been potentially influenced by the pathological abnormalities present. Nonetheless, the videos were standardized in adhering to a consistent diagnostic methodology.

CONCLUSION
Arthroscopic classification of LT pathology by four experienced hip arthroscopists using the G&V and the DC systems demonstrated only fair inter-observer reliability. The major discrepancy in interpretation was found between what constituted normal and partial or low-grade tears. The presence of synovitis was identified as an important arthroscopic finding, which was not contained in either classification. If arthroscopy is to be regarded as the gold standard in diagnosing LT pathology, dynamic manoeuvres and probing are recommended. Thorough arthroscopic scrutiny reveals the prevalence of LT pathology is higher than previously reported.

ACKNOWLEDGEMENTS
The authors would like to acknowledge the assistance of the surgeons in rating the arthroscopic videos and the help of Mr Mark Tacey in performing the statistical analysis.

CONFLICT OF INTEREST
None declared.

REFERENCES
1. O’Donnell JM, Pritchard M, Salas AP, Singh PJ. The ligamentum teres–its increasing importance. J Hip Preserv Surg 2014; 1:3–11.
2. Bardakos NV, Villar RN. The ligamentum teres of the adult hip. J Bone Joint Surg Br 2009; 91:8–15.
3. Rao J, Zhou YX, Villar RN. Injury to the ligamentum teres. Mechanism, findings, and results of treatment. Clin Sports Med 2001; 20:791–9, vii.
4. Byrd JW, Jones KS. Traumatic rupture of the ligamentum teres as a source of hip pain. Arthroscopy 2004; 20:385–91.
5. Cerezal L, Kassarian J, Canga A et al. Anatomy, biomechanics, imaging, and management of ligamentum teres injuries. Radiographics 2010; 30:1637–51.
6. Cerezal L, Arnaiz J, Canga A et al. Emerging topics on the hip: ligamentum teres and hip microinstability. Eur J Radiol 2012; 81:3745–54.
7. Haviv B, O’Donnell J. Arthroscopic debridement of the isolated Ligamentum Teres rupture. Knee Surg Sports Traumatol Arthrosoc 2011; 19:1510–3.
8. Kivlan BR, Richard Clemente F, Martin RL, Martin HD. Function of the ligamentum teres during multi-planar movement of the hip joint. Knee Surg Sports Traumatol Arthrosoc 2013; 21:1664–8.
9. Martin HD, Hatem MA, Kivlan BR, Martin RL. Function of the ligamentum teres in limiting hip rotation: a cadaveric study. Arthroscopy 2014; 30:1085–91.
10. Martin RL, Kivlan BR, Clemente FR. A cadaveric model for ligamentum teres function: a pilot study. Knee Sports Traumatol Arthrosoc 2013; 21:1689–93.
11. Wenger D, Miyani F, Mah A, Oka R. The mechanical properties of the ligamentum teres: a pilot study to assess its potential for improving stability in children’s hip surgery. J Pediatr Orthop 2007; 27:408–10.
12. Wenger DR, Mubarak SJ, Henderson PC, Miyani F. Ligamentum teres maintenance and transfer as a stabilizer in open reduction for pediatric hip dislocation: surgical technique and early clinical results. J Child Orthop 2008; 2:177–85.
13. Leunig M, Beck M, Stauffer E et al. Free nerve endings in the ligamentum capitis femoris. Acta Orthop Scand 2000; 71:452–4.
14. Schaumkel JV, Villar RN. Healing of the ruptured ligamentum teres after hip dislocation—an arthroscopic finding. Hip Int 2009; 19:64–6.
15. Botser IB, Martin DE, Stout CE, Domb BG. Tears of the ligamentum teres: prevalence in hip arthroscopy using 2 classification systems. Am J Sports Med 2011; 39(Suppl):1175–25S.
16. Domb BG, Martin DE, Botser IB. Risk factors for ligamentum teres tears. Arthroscopy 2013; 29:64–73.
17. Amenabar T, O'Donnell J. Successful treatment of isolated, partial thickness ligamentum teres (LT) tears with debridement and capsulorrhaphy. *Hip Int* 2013; **23**:576–82.

18. Gray AJ, Villar RN. The ligamentum teres of the hip: an arthroscopic classification of its pathology. *Arthroscopy* 1997; **13**:575–8.

19. Byrd JW, Jones KS. Diagnostic accuracy of clinical assessment, magnetic resonance imaging, magnetic resonance arthrography, and intra-articular injection in hip arthroscopy patients. *Am J Sports Med* 2004; **32**:1668–74.

20. Devitt BM, Philippon MJ, Goljan P et al. Preoperative diagnosis of pathologic conditions of the ligamentum teres: is MRI a valuable imaging modality? *Arthroscopy* 2014; **30**:568–74.

21. Chahla J, Soares EA, Devitt BM et al. Ligamentum Teres Tears and Femoroacetabular Impingement: Prevalence and Preoperative Findings. *Arthroscopy* 2016; **32**:1293–7.

22. O’Donnell J, Economopoulos K, Singh P et al. The ligamentum teres test: a novel and effective test in diagnosing tears of the ligamentum teres. *Am J Sports Med* 2014; **42**:138–43.

23. Philippon MJ, Devitt BM, Ho CP et al. Corrections to our article “preoperative diagnosis of pathologic conditions of the ligamentum teres: is MRI a valuable imaging modality?”. *Arthroscopy* 2014; **30**:1219–20.

24. Cerezal L, Carro LP, Llorca J et al. Usefulness of MR arthrography of the hip with leg traction in the evaluation of ligamentum teres injuries. *Skeletal Radiol* 2015; **44**(11):1585–95.

25. Amenabar T, Piriz J, Mella C et al. Reliability of 3 Different Arthroscopic Classifications for Chondral Damage of the Acetabulum. *Arthroscopy* 2015; **31**:1492–6.

26. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**:159–74.

27. Byrd JW, Jones KS. Hip arthroscopy in athletes. *Clin Sports Med* 2001; **20**:749–61.

28. Porthos Salas A, O’Donnell JM. Ligamentum teres injuries - an observational study of a proposed new arthroscopic classification. *J Hip Preserv Surg* 2015; **2**:258–64.

29. Garabekyan T, Chadayammuri V, Pascual-Garrido C, Mei-Dan O. All-arthroscopic ligamentum teres reconstruction with graft fixation at the femoral head-neck junction. *Arthrosc Tech* 2016; **5**:e143–7.

30. Lindner D, Sharp KG, Trenga AP et al. Arthroscopic ligamentum teres reconstruction. *Arthrosc Tech* 2013; **2**:e21–5.

31. Philippon MJ, Pennock A, Gaskill TR. Arthroscopic reconstruction of the ligamentum teres: technique and early outcomes. *J Bone Joint Surg Br* 2012; **94**:1494–8.

32. Simpson JM, Field RE, Villar RN. Arthroscopic reconstruction of the ligamentum teres. *Arthroscopy* 2011; **27**:436–41.