Reliability and accuracy of the tape measurement method with a nearest reading of 5 mm in the assessment of leg length discrepancy

Jamaluddin S, Sulaiman A R, Kamarul Imran M, Juhara H, Ezane M A, Nordin S

ABSTRACT

Introduction: The aim of this study was to determine the reliability and accuracy of the tape measurement method (TMM) with a nearest reading of 5 mm in assessing leg length discrepancy (LLD).

Methods: This was a cross-sectional study conducted on 35 patients with LLD and 13 patients without LLD. Two blinded surgeons measured the lower limbs from the anterior superior iliac spine to the medial malleolus using TMM with a nearest reading of 5 mm. Computed tomography (CT) scanograms of the lower limbs of 22 patients were conducted by two blinded radiologists. Intraclass correlation coefficient (ICC) with 95 percent confidence interval was calculated to assess the interobserver reliability of TMM. The accuracy of TMM was assessed by comparison with CT as the gold standard.

Results: The interobserver reliability of LLD measurement using both TMM and CT scanogram was high, with ICCs of 0.924 and 0.971, respectively. No significant mean difference on paired sample *t*-test was observed for both TMM and CT scanogram. Compared to CT scanogram, TMM had good accuracy, with an ICC of 0.805. When the mean TMM readings by two observers were compared to those derived from CT scanogram, the ICC was found to be 0.847, with a mean difference of 1.95 (range −3.17 to 7.07) mm.

Conclusion: There was excellent agreement in the LLD measurements between the two surgeons using TMM, between the two radiologists using CT sonogram, and between the TMM and CT measurements. This study showed that one TMM with the nearest reading of 5 mm was reliable and accurate in measuring LLD.

Keywords: leg length discrepancy, tape measurement method

Singapore Med J 2011; 52(9): 681–684

Department of Orthopaedics, School of Medical Sciences, Universiti Sains Malaysia, Kota Bharu 16150, Malaysia

Jamaluddin S, MD, MMed
Surgeon

Sulaiman AR, MD, MMed
Surgeon and Lecturer

Department of Community Medicine

Kamarul Imran M, MD, MMed
Lecturer

Department of Radiology

Juhara H, MD, MMed
Radiologist and Lecturer

Ezane MA, MD, MMed
Radiologist and Lecturer

Department of Orthopaedics, Faculty of Medicine and Health Science, University Sultan Zainal Abidin, Jalan Sultan Mahmud, Kuala Terengganu 20400, Malaysia

Nordin S, MD, MS
Surgeon and Lecturer

Correspondence to: Dr Abdul Razak Sulaiman
Tel: (60) 9 767 6389
Fax: (60) 9 765 3370
Email: abdrazak@kb.usm.my

INTRODUCTION

Leg length discrepancy (LLD) can arise due to true bone shortening or secondary joint contracture. It can lead to limping, increased energy expenditure during walking, late degeneration of the knee and hip and lower back pain.[1-4] An accurate and reliable method of assessment would determine the management and prediction of future LLD. Radiographic measurement is usually used to obtain an accurate assessment for planning of specific limb lengthening. However, the risk of radiation exposure and high costs have led clinicians to depend on clinical measurement for primary or repeated assessment. The tape measurement method (TMM) is one such method that can easily be performed at the bedside and requires only a measurement tape. TMM is usually a preferred clinical measurement method when there is no shortening below the ankle joint and no influence of body load on the discrepancy. In most of the previous studies, the reliability and accuracy of TMM were assessed based on a nearest reading of 1 mm.[5-7] However, many clinicians now use TMM with a nearest reading of 5 mm. The aim of this study was to determine the reliability and accuracy of TMM using the nearest reading of 5 mm.

METHODS

This was a cross-sectional study carried out at Hospital Universiti Sains Malaysia for a period of 24 months, from January 2008 to January 2010. A total of 35 patients with LLD and 13 patients without LLD were recruited for the study. Patients with hip dislocation, hip adduction contracture or abduction contracture > 15° and asymmetrical knee deformity were excluded from

Table I. Descriptive statistics of patients who underwent LLD measurement by TMM only or both TMM and CT scanogram.

Measurement method	Mean LLD ± SD (mm)
TMM (n = 48)*	
Surgeon 1	24.10 ± 20.83
Surgeon 2	25.40 ± 23.58
CT (n = 22)†	
Radiologist 1	27.30 ± 21.53
Radiologist 2	27.80 ± 19.02
TMM (n = 22)†	
Surgeon 1	28.20 ± 18.10
Surgeon 2	30.20 ± 22.23

* Total no. of patients; † patients who underwent both TMM and CT scanogram.
LLD: leg length discrepancy; TMM: tape measurement method; CT: computed tomography; SD: standard deviation

the study. The sample size was estimated based on 'SAMPICC' on STATA statistical software (StataCorp LP, College Station, TX, USA). We expected the alternative hypothesis for intraclass correlation coefficients (ICCs) to be 0.8 and the null hypothesis to be 0.5, with the power set at 80% and Type I error set at 5%. The required sample size for our study was 22 patients.

Two blinded orthopaedic surgeons measured the lower limbs of all patients using TMM, with the patients positioned in supine. Measurement was done from the anterior superior iliac spine (ASIS) to the distal tip of the medial malleolus with the nearest 5 mm reading using standard woven tape measures. Any deformity below the ankle joint was not assessed in this study. ASIS was identified as the first bony prominence felt by palpation proximally along the inguinal ligaments. To ensure that both the ASIS were at the same level, the legs were moved in a parallel position. LLD was defined as the difference in the lengths of the two limbs. The differences in LLD observed in each patient were analysed for interobserver reliability of TMM.

Out of the 48 patients, 22 underwent discrepancy measurement with computed tomography (CT) scanogram using the LightSpeed Plus CT Scanner (GE Medical Systems, Milwaukee, WI, USA). An anteroposterior (AP) scout scanogram of the lower limb was captured while the patient was placed in the supine position. With both the ASIS at the same level, the patient's legs were fully extended and the feet were strapped in an upright position. The measurement tools on the Picture Archive and Communication System (PACS) workstation were used to measure the distance between the superior lip of the acetabulum to the distal end of the tibia (Fig. 1). Two blinded radiologists conducted the LLD measurements separately. The differences in LLD using CT sonogram

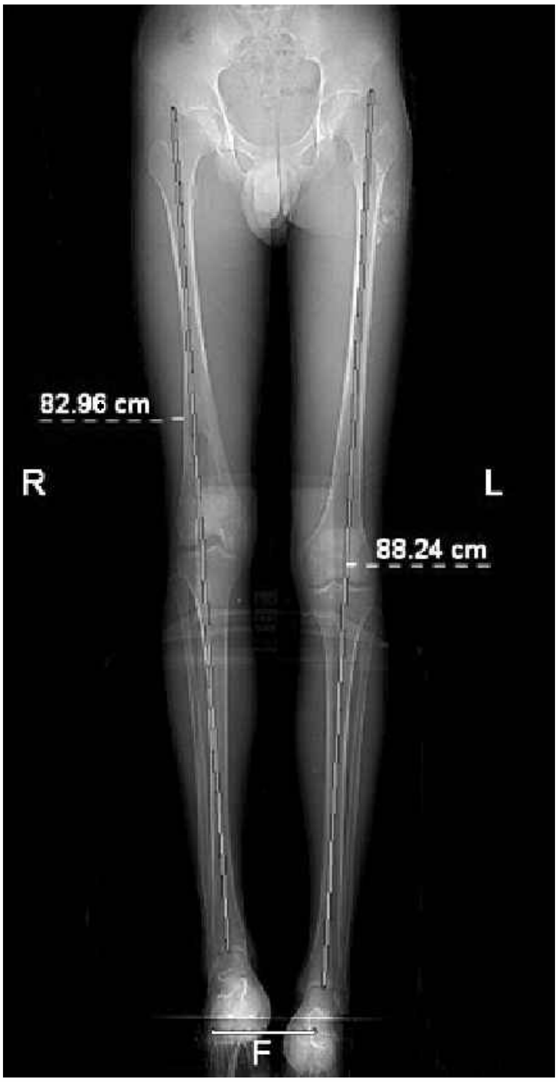

Fig. 1 CT image of the lower limb shows the limb length measurement between the superior lip of the acetabulum to the distal end of the tibia.

were analysed for interobserver reliability. This study was approved by the Human Ethical Committee of the School of Medical Sciences, Universiti Sains Malaysia.

Data was entered, cleaned and analysed using the Predictive Analysis Software version 18 (SPSS Inc, Chicago, IL, USA). We described LLD using the mean and standard deviation. In order to determine the agreement between two TMM measurements, two CT measurements and between a TMM and CT measurement of LLD, the ICCs were calculated. The value of a single measure ICC ≥ 0.75 indicated excellent agreement. Paired *t*-tests were performed to compare the differences in the mean measurement of LLD between the two surgeons using TMM, and between the TMM and CT measurements. The mean differences were checked, and a p-value > 0.05 indicated no significant differences in the measurements. A p-value < 0.05 was considered to be statistically significant (two-sided).

Table II. Comparison of leg length discrepancy measurement obtained by different observers and via different methods.

Comparison	ICC; 95% CI	Mean difference; 95% CI (mm)	p-value*
TMM† (n = 48)	0.92 (0.87 to 0.96)	−1.35 (−3.87 to 1.17)	0.29
CT¶ (n = 22)	0.97 (0.93 to 0.99)	−0.59 (−2.81 to 1.63)	0.59
TMM vs. CT†† (n = 22)	0.81 (0.59 to 0.91)	0.93 (−4.68 to 6.54)	0.74
Mean TMM vs. CT¶¶ (n = 22)	0.85 (0.67 to 0.93)	1.95 (−3.17 to 7.07)	0.44

* paired *t*-test; † Surgeon 1 vs. Surgeon 2; ¶ Radiologist 1 vs. Radiologist 2; †† Surgeon 1 vs. Radiologist 1; ¶¶ Surgeon 1 & 2 vs. Radiologist 1

ICC: intraclass correlation coefficient; CI: confidence interval; TMM: tape measurement method; CT: computed tomography

RESULTS

A total of 48 patients were included in the analysis. The mean age of the patients was 22.60 ± 15.80 years. 35 (72.9%) patients were male and 13 (27.1%) were female. The mean age of the 22 patients (15 male and seven female) who underwent CT sonogram was 14.00 ± 5.46 years.

Interobserver reliability of LLD measurement using TMM was analysed for 48 patients (Table I). The calculated ICC with 95% confidence interval (CI) between two observers using TMM was 0.92, which indicated a high correlation of interobserver reliability (Table II). The mean difference calculated using paired sample *t*-test was –1.35 (range –3.87 to 1.17) mm, with a p-value of 0.29 (Table II). Interobserver reliability of LLD measurement using CT was analysed for 22 patients (Table I). The calculated ICC was observed to be 0.97 with 95% CI, indicating a high correlation of interobserver reliability. The mean difference using paired sample *t*-test was –0.59 (range –2.81 to 1.63) mm, with a p-value of 0.59 (Table II).

Since the interobserver reliability of both TMM and CT was high, we compared the findings of TMM conducted by the first surgeon to the CT findings by the first radiologist on 22 patients (Table I). The correlation was good, with an ICC of 0.81 (Table II). When the mean of two TMM readings by the two surgeons were compared with the CT measurements by the first radiologist, the ICC was 0.85, and the mean difference using paired sample *t*-test was 1.95 (range –3.17 to 7.07) mm (Table II).

DISCUSSION

An ideal measurement method should be reliable and accurate. Reliability is the variation between observers and within a single observer in obtaining the measurement, whereas accuracy refers to the variation in measurement using a technique compared with the actual measurement.[8]

In this study, we defined LLD as a condition of unequal length of the lower limbs, which may originate anywhere from the hip to the ankle joints. Since the measurement is performed from the ASIS to the medial malleolus, it is very difficult to obtain a consistent reading with a nearest reading of 1 mm, as both the ASIS and medial malleolus are smooth prominent bones under the skin. TMM with a nearest reading of 5 mm in this study was found to have better interobserver ICC (0.924) compared to TMM with a nearest reading of 1 mm in other studies that had ICCs of 0.668 and 0.83.[6,7] Beattie et al found that interobserver ICC improved to 0.910 when the mean of two readings by each examiner was compared.[7] The finding in our study was based on only a single reading by each examiner. In this study, we did not assess the intraobserver reliability; however, other studies have shown that intraobserver reliability (ICC = 0.807) was better than interobserver reliability (ICC = 0.668).[7]

In the radiological method of assessing LLD, factors like soft tissue and uncooperative patients are eliminated. Proper and consistent identical bony points could be selected to ensure better interobserver reliability among different examiners. We found a high interobserver reliability (ICC = 0.96), a finding that was supported by Leitzes et al's study, which reported an ICC of > 0.99.[9] They also found that the accuracy of CT in measuring LLD has an absolute difference of 0.68 mm when compared to the measurement of cadaveric femoral length using an electronic calliper as the gold standard.[9] Apart from accuracy, the radiation dose of a CT Scout view has been found to be 50–100 times lower than that of conventional radiography.[10] Based on these facts, CT measurement by one radiologist was chosen to be the gold standard for measurement of LLD in our patients.

Higher interobserver reliability is expected using TMM with a nearest reading of 5 mm compared to TMM with a nearest reading of 1 mm; however, it could be less accurate. In this study, we found that one TMM measurement by a single observer had good correlation with the gold standard CT (ICC 0.805) measurement.

This finding was almost similar to that of Beattie et al's study, where a single TMM measurement with a nearest reading of 1 mm had an ICC of 0.770, which improved to 0.852 when the mean of two TMM readings was used to compare with those obtained using radiological measurement.[11] Beattie et al used radiographic measurement as the gold standard, while another study showed that it had excellent reliability and accuracy, similar to that of CT measurement.[9,11] Hanada et al found that clinical measurement by block method using standing radiograph of the pelvis underestimated the LLD measurement by an average of 5.1 mm.[5] We found that TMM may under- or overestimate the LLD by 4.7–6.6 mm; however, this was not statistically significant.

We conclude that there is excellent agreement between the TMM measurements of the two surgeons, between the CT measurements of the two radiologists, and between the TMM and CT measurements. Also, the differences in mean measurements in the above three comparisons were not statistically significant. This study has shown that one TMM measurement with a nearest 5 mm reading was reliable and accurate in measuring LLD.

REFERENCES

1. Abdulhadi HM, Kerrigan DC, LaRaia PJ. Contralateral shoe-lift: effect on oxygen cost of walking with an immobilized knee. Arch Phys Med Rehabil 1996; 77:670-2.
2. Dixon AS, Campbell-Smith S. Long leg arthropathy. Ann Rheum Dis 1969; 28:359-65.
3. Kaufman KR, Miller LS, Sutherland DH. Gait asymmetry in patients with limblength inequality. J Pediatr Orthop 1996; 16:144-50.
4. Friberg O. Clinical symptoms and biomechanics of lumbar spine and hip joint in leg length inequality. Spine (Phila Pa 1976) 1983; 8:643-51.
5. Hanada E, Kirby RL, Mitchell M, Swuste JM. Measuring leg-length discrepancy by the "iliac crest palpation and book correction" method: reliability and validity. Arch Phys Med Rehabil 2001; 82:938-42.
6. Terry MA, Winell JJ, Green DW, et al. Measurement variance in limb length discrepancy: clinical and radiographic assessment of interobserver and intraobserver variability. J Pediatr Orthop 2005; 25:197-201.
7. Beattie P, Rothstein JM, Kopriva L. The clinical reliability of measuring leg length. Abstract Phys Ther 1988; 68:588.
8. Huurman WW, Jacobsen FS, Anderson JC, Chu WK. Limb-length discrepancy measured with computerized axial tomographic equipment. J Bone Joint Surg Am 1987; 69:699-705.
9. Leitzes AH, Potter HG, Amaral T, et al. Reliability and accuracy of MRI scanogram in the evaluation of limb length discrepancy. J Pediatr Orthop 2005; 25:747-9.
10. Helms CA, McCarthy S. CT scanograms for measuring leg length discrepancy. Radiology 1984; 151:802.
11. Beattie P, Isaacson K, Riddle DL, Rothstein JM. Validity of derived measurements of leg-length differences obtained by use of a tape measure. Phys Ther 1990; 70:150-7.