



A systematic review on the validity and reliability of tape measurement method in leg length discrepancy

Behshid Farahmand¹, Esmaeil Ebrahimi Takamjani², Hamid Reza Yazdi³, Hassan Saeedi⁴, Mohammad Kamali⁵, Masumeh Bagherzadeh Cham*⁶

Received: 16 May 2018

Published: 26 May 2019

Abstract

Background: Leg length discrepancy measurement is an essential part of musculoskeletal clinical assessment, and tape measurement is a common clinical method. This study aimed to systematically review the results of the findings of studies on validity and reliability of the tape measurement method and the quality of reporting the literature on this topic.

Methods: A search was performed in PubMed, EBSCO, Science Direct, Web of Knowledge, Scopus, Embase, and Google Scholar using selected keywords from inception to December 2017. This systematic review was based on the PRISMA guideline. After a systematic selection process, the quality of the included studies was assessed independently by 2 reviewers using the Brink and Low Scale for quality assessment.

Results: A total of 11 studies were finally considered for this systematic review. Two studies were about the validity of (a measurement tool) studies and 4 were reliability analysis only. Validity and reliability analyses were simultaneously applied to 5 studies. Also, 9 out of 11 studies were deemed to be of high quality based on Brink and Low Scale. Studies showed high (ICC=0.7) to very high (ICC=0.9) levels of interrater and intrarater reliability. The validity of the methods ranged from low to very high depending on subjects.

Conclusion: Tape measurement method has acceptable reliability and validity in healthy people, but it does not have acceptable validity in measuring obese people and patients with musculoskeletal disorders. Thus, using a suitable method for LLD leg length discrepancy measurement seems to be necessary for obese and individuals with leg length discrepancy.

Keywords: Leg length discrepancy, Validity, Reliability

Conflicts of Interest: None declared

Funding: Iran University of Medical Sciences, Tehran, Iran

*This work has been published under CC BY-NC-SA 1.0 license.

Copyright© Iran University of Medical Sciences

Cite this article as: Farahmand B, Ebrahimi Takamjani E, Yazdi HR, Saeedi H, Kamali M, Bagherzadeh Cham M. A systematic review on the validity and reliability of tape measurement method in leg length discrepancy. *Med J Islam Repub Iran.* 2019 (26 May);33:46. <https://doi.org/10.14196/mjiri.33.46>

Introduction

Leg length discrepancy (LLD) is a common musculoskeletal disorder. Almost 70% of the general population suffers from LLD of up to 1 cm (1). LLD assessment is an important topic in evaluation of lower limb. Various leg length assessment procedures are commonly being used by several practitioners, such as physical therapists, orthopedics, technical orthopedics, chiropractic, and podiatrists (2, 3). LLD assessment is a challenging task among re-

searchers and clinicians. There is still controversy about the assessment of LLD measurements. Two general categories of methods have been used for LLD assessment: (1) imaging techniques, (2) clinical methods (4). Imaging techniques are the standard procedure for accurate LLD measurement, but they are costly, time-intensive, and expose patients to radiation (4). Therefore, clinical methods are more popular because of their availability, easy procedure,

Corresponding author: Dr Masumeh Bagherzadeh Cham, bagherzadehcham.m@iums.ac.ir

¹ Department of Orthotics & Prosthetics, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

² Department of Physiotherapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

³ Orthopedic Department, Firuzgar Hospital, Iran University of Medical Sciences, Tehran, Iran

⁴ Department of Orthotics & Prosthetics, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran

⁵ Rehabilitation Management Department, Iran University of Medical Sciences, Tehran, Iran

⁶ Neuromusculoskeletal Research Center, Department of Physical Medicine and Rehabilitation, Iran University of Medical sciences, Tehran, Iran

↑What is “already known” in this topic:

Tape measurement is a simple, affordable, and non-invasive method to evaluate limb length inequality in healthy people with leg length discrepancy more than 5 mm. The validity and reliability of this method decreased in obese people and in those with orthopedic disorder.

→What this article adds:

New valid, reliable, non-invasive, and cost-effective method is needed to evaluate leg length discrepancy in obese people and in those with orthopedic or neuromuscular disorder.

and low cost. Indirect and direct methods are used for clinical LLD assessment (4-6). In indirect methods, lift blocks of different thicknesses are put under the short leg and then the leveling of the pelvis is checked. In direct methods, tape measurement methods (TMM), a tape is used to measure the distance between the anterior superior iliac spine (ASIS) or anterior inferior iliac spine (AIIS) to the lateral or medial malleolus in supine position (4-8). Despite easy procedure, validity of this method is not known.

There is some controversy about validity and reliability of these methods. Several literatures on TMM have shown conflicting results. In their study, Friberg et al (9) concluded that ASIS to medial malleolus measurement is an inaccurate and imprecise measure of LLD, with a mean LLD error of 8.6 mm compared to radiographs, and a 1.1 mm intratester mean error. Authors of this study used a single measurement of ASIS to the medial malleolus. In agreement with this finding, Beattie et al (10) found validity estimates of LLD to have intra-class correlation coefficient (ICC) of 0.683, when utilizing a single measure using ASIS to medial malleolus measure. In addition, Gogia and Braatz (3) reported ICC of 0.99, with radiographs and intertester reliability of 0.98. Also, Hoyle et al (11) noted an intertester reliability ranging from 0.96 to 0.98 and an intratester reliability ranging from 0.89 to 0.95 for ASIS to medial malleolus measurement. The widely different results have been reported on the validity and reliability of the method. Thus, the aim of this study was to perform a systematic review on the validity and reliability of TMM as a common clinical method for LLD measurement, to evaluate the quality of these studies, and to investigate the factors affecting the reliability and validity of this method.

Methods

Search strategy

To systematically search the literature published from the date of inception to December 2017, PubMed, EBSCO, Science Direct, Web of Knowledge, Scopus, and Embase databases were explored. The keywords were as follow: ("leg length discrepancy" OR "limb length discrepancy" OR "leg length inequality" OR "leg length" OR "limb length") AND ("validity" OR "reliability" OR "inter-tester" OR "intra-tester") AND ("tape measure" OR "tape measurement" OR "clinical methods" OR "clinical assessment"). A word from each area was required to be in the text or the title-abstract-keyword of the study. An additional search of Google Scholar search engine was also performed. These searches were supplemented by hand searching the reference lists of the final articles found from the search.

Inclusion criteria of the studies:

- English language
 - Full-text availability
 - Measuring the validity and/or reliability of the tape measurement method should have been the primary aim of the study
 - Pearson's r , Cronbach α , and intraclass correlation coefficient should have been used for statistical analysis.
- Studies were excluded if they examined the validity

and/or reliability of tape measurement in total hip or knee disarticulation of patients and if they evaluated the accuracy and/ or precision of tape measurement.

Study selection

In this systematic review, all search procedures; selection, quality assessment, data extraction, and reading the articles were screened independently by 2 reviewers (BF & MB). In case of any difference of opinion between the reviewers, a third reviewer was asked to evaluate the article (MK). Initially, based on the inclusion criteria, studies were selected by reading the titles and abstracts. Then, full-text articles were explored to choose those that met the inclusion criteria.

Quality assessment

Ultimately, 11 articles were selected for final analysis. Two reviewers conducted an appraisal of the reporting quality of the 11 studies based on Kappa scores (12). Accordingly, a Kappa score of 0.92 or higher was regarded as acceptable.

The Brink and Louw Scale checklist was used for quality assessment (13). This 13-item scale has been developed by combining QUADAS (the Quality Assessment of Diagnostic Accuracy Studies) and QAREL (the Quality Appraisal of Diagnostic Reliability Studies) scales, as the selected studies could assess both reliability and validity of the tape measurement. Thus, this checklist can be more conveniently used as compared to either of QUADAS or QAREL independently (13, 14). The studies were considered as of high quality if their scores were higher than 60% (14, 15). Two reviewers independently assessed the quality of each study.

Data analysis

The intra-class correlation coefficient and Pearson's correlation coefficient were interpreted as follow: 0.00-0.29 as very low correlation, 0.30-0.49 as low correlation, 0.50-0.69 as moderate correlation, 0.70-0.89 as high correlation, and 0.90-1.00 as very high correlation (16).

Results

Study Selection

With a preliminary literature search, 496 abstracts were identified without excluding non-English articles. However, after applying the inclusion criteria, 11 articles were included in the review. Figure 1 presents the flow diagram of the selection process based on the PRISMA guidelines.

A total of 11 studies were finally considered for this systematic review (Table 1). Two studies were subjected to validity studies only (10, 17), while 4 were subjected to reliability analysis only (11, 18-20). Validity and reliability analyses were applied to 5 studies simultaneously (3, 7, 8, 21, 22).

Characteristics

Total numbers of participants in all samples was 458. The healthy sample included 174 individuals and LLD sample included 284 (Table 2). Five studies had solely evaluated healthy individuals (3, 7, 11, 18, 19), 3 focused

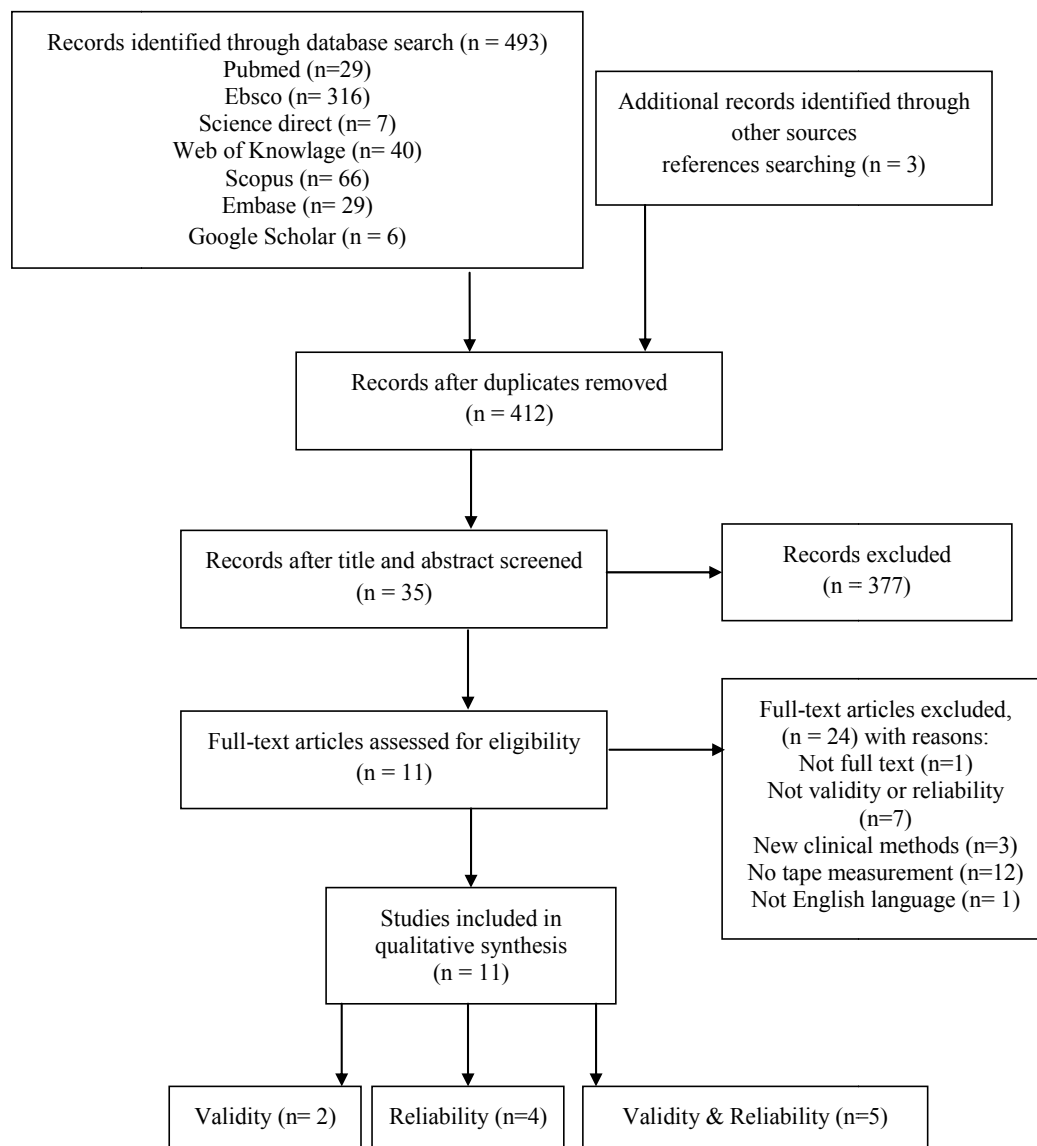


Fig. 1. PRISMA flow diagram

on LLD patients (17, 20), and 3 evaluated both healthy and LLD participants (8, 10, 21). A total of 129 females and 98 males were studied in 7 studies, and 3 studies did not report the gender of their participants (17, 20, 21).

Study characteristics

Among the studies that evaluated the validity of TMM, only Neelly et al (7) and Gogia et al (3) studies had used 2 examiners and in other studies only one examiner's evaluation was compared to the reference method (8, 10, 17, 21, 22). Moreover, only Beattie et al (10) study compared the validity of TMM in healthy (ICC: 0.359- 0.786) and patient (ICC: 0.770- 0.852) groups (Table 1).

In studies that measured reliability, all studies applied expert examiners and intertester and intratester reliability were 0.477- 0.991 and 0.679- 0.990, respectively (3, 7, 8, 11, 18-22). Only in Duff et al (21) study expert and non-expert examiners evaluated LLD and found between expert examiners intertester reliability of 0.49, between non-

expert examiners reliability of 0.19, and between expert and non-expert examiners reliability of 0.01- 0.69. Hoyle et al (11) study measured intertester and intratester reliability of TMM in right and left foot separately. Also, Terry et al (20) measured intertester and intratester reliability of TMM in 2 positions: ASIS to medial maleolus (ICC: 0.80) and to lateral maleolus (ICC: 0.83).

There were 4 radiography studies (3, 17, 21, 22), 1 study with CT Scan (7), and 3 with scanogram (8, 19, 20). In Beattie et al (10) study, both radiography and mini scanogram were used as reference method (Table 2).

In most studies, leg length was measured from ASIS to medial maleolus (3, 7, 8, 10, 11, 18, 19, 21). In Badii et al (22) study, leg length was measured from ASIS to lateral maleolus and in Terry et al (20) study, it was measured from ASIS to both medial and lateral maleolus. However, only in lamp et al (17) study, leg length was measured from ASIS to knee joint and from knee joint to medial maleolus (Table 2).

Table 1. Summary of studies and validity and reliability of the included studies

Author	Validity	Inter-tester ICC	Intra-tester ICC
Badii et al, 2014	0.75	0.477 (0.253–0.706)	
Neelly et al, 2013	E1: 0.984 (0.973–0.990) E2: 0.978 (0.961–0.988)	0.991 (0.983–0.995)	E1: 0.990 (0.984–0.994) E2: 0.985 (0.976–0.991)
Jamaluddin et al, 2011	0.81 (0.59 to 0.91)	0.92 (0.87 to 0.96)	
Leard et al, 2009		T1: 0.955 T2: 0.963	E1: 0.991 E2: 0.993 0.679
Bretas et al, 2009		The medial malleolus: 0.80	The medial malleolus: 0.78
Terry et al, 2005		The lat malleolus: 0.83	The lat malleolus: 0.88
Duff et al, 2000	0.11 CI	EX1,2=0.49 EX1,NEX1=0.69 EX1,NEX2=0.17 EX2,NEX1=0.43 EX2,NEX2=0.01 NEX1,2=0.19	EX1=0.17 EX2=0.32 NEX1=0.29 NEX2=0.31 CI
Lampe et al, 1996	95% limit of agreement: -1.8/2.1 Predict value below the knee: 0.75 Above the knee:0.64		
Hoyle et al, 1991		Left:0.989 Right: 0.965	Left E1:0.954 E2: 0.904 Right E1: 0.895 E2:0.942
Beattie et al, 1990	First measurement Patient: 0.770 Normal: 0.359 All subjects:0.683 Second measurement Patient:0.803 Normal:0.786 All subjects:0.790 Mean measurement Patient:0.852 Normal:0.637 All subjects:0.793		
Gogia et al,1986	E1: 0.990 E2: 0.990	0.990	

T: test session, E: examiner, EX: expert examiner, NEX: non-expert examiner

Table 2. Summary of the included studies

Author	Patients	Position test	Examiners	Reference methods
Badii et al, 2014	20 patients with LLD 10 M, 10 F	Supine Inferior aspect of the ASIS- Inferior aspect of the lateral malleolus	4 experienced examiners: rheumatologist, occupational therapist, physical therapist, and a senior rheumatology fellow	Radiographs
Neelly et al, 2013	30 community dwelling adults 11 M, 19 F 38.40±13.10 years	Supine Inferior aspect of the ASIS- Inferior aspect of the medial malleolus	2 expert physical therapist	CT scans
Jamaluddin et al, 2011	35 patients with LLD and 13 healthy 35 M, 13 F 22.60±15.80 years	Supine anterior superior iliac spine (ASIS) to the distal tip of the medial malleolus with the nearest 5 mm	2 professional orthopedic surgeons	CT scanogram
Leard et al, 2009	18 healthy 6 M, 12 F 23.70±3.60 years	Supine Inferior aspect of the ASIS- distal point of the medial malleolus	10 physical therapist student	-
Bretas et al, 2009	37 healthy 9M, 28F 50±18 years	Supine Inferior aspect of the ASIS- Inferior aspect of the medial malleolus	Trained examiner	CT scanogram
Terry et al, 2005	16 patients with LLD 9.97 years	Supine - Inferior aspect of the ASIS- Inferior aspect of the medial malleolus & - Inferior aspect of the ASIS- Inferior aspect of the lateral malleolus	4 observe: pediatric orthopedic surgeon, pediatric orthopedic fellow, fourth-year orthopedic resident, physical therapist	CT scanogram
Duff et al, 2000	13 patients with LLD and 12 healthy Age range 21-53	Supine Inferior aspect of the ASIS- Inferior aspect of the medial malleolus	Two expert podiatrist, two non-expert podiatrists	Radiographs

Quality of studies

The 2 reviewers had no disagreement about article se-

lection (kappa score 1). However, they assessed the quality of the articles differently. Nine out of the 11 studies

Table 2. Cntd

Author	Sample	Position	Examiner	Imaging
Lampe et al, 1996	190 children	Supine - Anterior Superior Iliac Spine to the medial knee joint line & - the medial knee joint line to the medial malleolus	One examiner	Orthoradiography
Hoyle et al, 1991	25 healthy 6 M, 19 F Age range 22-40	Supine Anterior Superior Iliac Spine to the medial malleolus	2 professional physical therapist	-
Beattie et al, 1990	10 patients with LLD 6M, 4F 34.10±11.20 9 healthy 26.50±3.70 3M, 6F	Supine Anterior Superior Iliac Spine to the medial malleolus	One examiner	Radiographs Mini scanogram
Gogia et al, 1986	30 healthy 12 M, 18 F Age range 18-65	Supine Anterior Superior Iliac Spine to the medial malleolus	2 expert physical therapist	Radiographs

Table 3. Results of quality assessment of studies using Brink and Louw Scale

Brink and Louw Scale	Badii et al., 2014	Neelly et al., 2013	Jamaluddin et al., 2011	Lear et al., 2009	Breath et al., 2009	Terry et al., 2005	Duff et al., 2000	Lampe et al., 1996	Hoyle et al., 1991	Beattie et al., 1990	Gogia et al., 1986
Was the sample of subjects' representative?	+	+	+	+	+	+	+	+	+	-	+
Was the sample of raters' representative?	+	+	+	+	+	+	+	+	+	-	+
Was the reference standard explained?	+	+	+	Not	Not	+	+	+	-	+	+
Were raters blinded to the findings of other raters?	+	+	+	+	+	+	+	-	+	Not	+
Were raters blinded to their own prior findings?	+	+	Not	+	Un	un	+	Not	Un	Not	Not
Was the order of examination varied?	Not	Not	Not	Not	Not	+	+	Not	Not	Not	Not
Is the time period between reference standard and index test short enough?	Un	Un	Un	Not	Not	+	Un	Un	Not	Un	+
Was the time interval between repeated measures appropriate?	Un	Un	Un	+	+	+	+	Un	Un	+	+
Was the reference standard independent to the index test?	+	+	+	Not	Not	+	+	+	Not	+	+
Was the execution of the index test described in sufficient detail?	+	+	+	+	+	+	+	+	+	+	+
Was the execution of the reference standard described in sufficient detail?	+	+	+	Not	Not	+	+	-	Not	+	+
Were withdrawals from the study explained?	Un	Un	Un	Un	+	Un	Un	Un	Un	+	Un
Were the statistical methods appropriate?	+	+	+	+	+	+	+	-	+	+	+
Total score	9	9	8	7	7	11	11	5	5	7	10

Not = Not applicable, Un = Unclear, (+) = Yes, (-) = No.

were deemed to be of high quality (score > 60%). One of the studies subjected to validity studies was found to be of high quality (10). Three out of 4 reliability studies were found to be of high quality (18-20). Also, all studies in which combined reliability and validity studies were performed were found to be of high quality (3, 7, 8, 21, 22).

Details of the scoring process are presented in Table 3. The main areas of weakness were insufficient interrater and intrarater blinding, lack of variation in testing order, the time period between reference standard and index test, and withdrawal from the study.

Discussion

The aim of this study was to review the validity and reliability of TMM for LLD measurement. Results of this study showed that TMM is a valid and reliable tool for LLD measurement in healthy people with no excess weight or musculoskeletal disorder. Moreover, an expert examiner measured LLD from ASIS to medial or lateral malleolus twice or more, and the mean measurements were considered. Even though radiography is the standard procedure for LLD measurement, it is too expensive, exposes the subject to radiation, and requires special equipment. Therefore, based on the results of this study, it seems that TMM serves as a safe method for LLD in clinics and research centers.

TMM refers to direct measurement of LLD using bony landmarks (4, 5). In this method, the person is in supine position with his/her lower limb being in anatomical position (extended hip and knee with the ankles in neutral position) (3, 10). Different studies have checked different distances (eg, ASIS to medial malleolus, ASIS to lateral malleolus, ASIS to medial knee joint, medial knee joint to medial malleolus, umbilicus to medial malleolus, or sternum to malleolus) (17, 20, 22, 23).

The results of the present study showed that the distance between ASIS to the medial malleolus provides the highest validity, while the distance between ASIS to the lateral malleolus merely provides acceptable validity (17, 20). Other measurement methods provided significantly lower validities. Results of a study by Terry et al (20) demonstrated that the values of ICC for intrarater reliability of the ASIS to the medial malleolus and lateral malleolus was 0.78 and 0.88, respectively, and those for interrater reliability of the same distances was 0.80 and 0.83, respectively.

In these studies, validity and reliability of TMM were investigated in healthy individuals and patients. The patients had a history of lower limb fractures, Blount disease, Tibia or Femur Hemimelia, hemihypertrophy, osteomyelitis, osteoarthritis, low back pain, tibia dysplasia, obesity, or other deformities. The results also revealed that validity and reliability of the TMM for LLD measurement were acceptable in healthy participants (0.97-0.98) (7) compared to the patients (ICC=0.33-0.39) (19). However, Beattie et al (10) found relatively different results (Table 1). Significantly lower validity and reliability were observed in obese participants (ICC=0.22) (19).

In the meantime, TMM uses bony landmarks for LLD measurement, making it difficult to detect accurate location of the ASIS in obesity. Asymmetry in other segments (thighs, knees, ankles) due to swelling, muscle atrophy, contracture, or obesity can change the direction of the tape and alter the results. Pelvic obliquity may not actually cause LLD but can affect measurement (10, 19).

Incorrect detection of bony landmarks and inappropriate placement of the tape along the lower limb can introduce errors into the measurement, and therefore the examiner's skill and experience in LLD measurement via TMM affects reliability and validity of the results (21, 23). In most studies, LLD measurement via TMM is practiced by expert examiners, such as physiotherapists, occupational

therapist, rheumatologist, orthopedic surgeons, or podiatrists (3, 7, 8, 11, 22). One study evaluated the effect of tester's skill on intertester and intratester reliability of TMM and found a direct relationship between examiner's skill and reliability of the results (21).

TMM was compared to a reference method to evaluate its validity. For this purpose, the examiner measured LLD via TMM for 1 or more than 1 time. The results showed that to enhance the validity of TMM, the tester must measure LLD for at least 2 times and average the results before comparing it to the reference method (10). Beattie et al (10) demonstrated the validity of TMM at the first and second measurements to be 0.359 and 0.786, respectively. They also found the highest validity with dual measurements to be 0.852. The results also indicated that, in general, the second measurement was more valid than the first, perhaps due to increased examiner's skill. However, Neely et al (7) and Jamaluddin et al (8) found that one-time measurement still provides acceptable validity.

Validity of TMM for LLD measurement has also been shown to depend on the length of the 2 limbs. Accordingly, higher errors have been reported for difference of less than 5 mm, while significantly lower errors are reported for cases where the difference exceeds 5 mm (9, 10). Beattie et al (10) found that any difference of smaller than 5 mm in leg length can increase the rate of error in determining the shorter length by the examiner on 4 out of 9 individuals, and the examiner will not make any mistake if the difference exceeded 5 mm.

Different studies have compared TMM with standard procedures, such as radiography, CT scan, and scanography (4). In LLD measurement, using CT scan images, the same bone landmarks as those used in TMM are used and measurements are made in supine position (24). In the CT scan method, biomechanical or structural asymmetry of the leg and ankle is not shown due to weight bearing. In studies where TMM was compared to CT scan, the validity of LLD measurements was higher than that provided by standing radiographs (7, 8, 20).

The results of a reliability study showed that when both examiners were adequately skillful, intertester and intratester reliability were acceptably high. Leard et al (18) reported an intertester reliability of 0.95-0.96 and intratester reliability of 0.99. Also, Gogia et al showed intertester reliability of 0.99. Duff et al (21) showed that intertester error exceeds that of intratester reliability. Jamaluddin et al (8) and Brêtas et al (19) found an intertester reliability of 0.807 and an intratester reliability of 0.668.

The current literature review has several limitations. First, as with any systematic review, it is possible that some related articles could not be identified, such as unpublished work or conferences articles. Second, only articles in English were included. Third, studies that evaluated accuracy and precision were not included. Finally, type of participants, examiners, and landmarks that were used for the measurement and the reference methods were not uniform.

Conclusion

In general, according to studies in which validity and re-

liability of the TMM were evaluated, this method has been shown to exhibit acceptable reliability and validity among healthy individuals. However, it fails to show acceptable validity for obese individuals or those engaged with orthopedic problems. The examiner should have the necessary skill and expertise to find bony landmarks and perform LLD measurement properly. TMM may provide unrealistic results for those with pelvic tilt or other problems. Thus, it is necessary to use a method which is suitable for both the healthy, the obese, and patients. Also, the method should be done in standing position, should eliminate the need of x-rays, and should not require examiner skill and expertise.

Acknowledgments

The authors would like to thank Iran University of Medical sciences, Tehran, Iran, for financially supporting this study.

Conflict of Interests

The authors declare that they have no competing interests.

References

1. Knutson GA. Anatomic and functional leg-length inequality: A review and recommendation for clinical decision-making. Part I, anatomic leg-length inequality: prevalence, magnitude, effects and clinical significance. *Chiropr Osteopat*. 2005;13:12.
2. Cooperstein R. Leg checking: Why and why not? *DC Tracts*. 2012;24(2):4-11.
3. Gogia PP, Braatz JH. Validity and Reliability of Leg Length Measurements. *J Orthop Sports Phys Ther*. 1986;8(4):185-8.
4. Sabharwal S, Kumar A. Methods for Assessing Leg Length Discrepancy. *Clin Orthop Relat Res*. 2008;466(12):2910-22.
5. Brady RJ, Dean JB, Skinner TM, Gross MT. Limb length inequality: clinical implications for assessment and intervention. *J Orthop Sports Phys Ther*. 2003;33(5):221-34.
6. Gurney B. Leg length discrepancy. *Gait & Posture*. 2002;15(2):195-206.
7. Neelly K, Wallmann HW, Backus CJ. Validity of measuring leg length with a tape measure compared to a computed tomography scan. *Physiother Theory Pract*. 2013;29(6):487-92.
8. Jamaluddin S, Sulaiman AR, Imran MK, Juhara H, Ezane MA, Nordin S. Reliability and accuracy of the tape measurement method with a nearest reading of 5 mm in the assessment of leg length discrepancy. *Singapore Med J*. 2011;52(9):681-4.
9. Friberg O, Nurminen M, Korhonen K, Soininen E, Manttari T. Accuracy and precision of clinical estimation of leg length inequality and lumbar scoliosis: comparison of clinical and radiological measurements. *Int Disabil Stud*. 1988;10(2):49-53.
10. 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(3):150-7.
11. Hoyle DA, Latour M, Bohannon RW. Intraexaminer, interexaminer, and interdevice comparability of leg length measurements obtained with measuring tape and metrecom. *J Orthop Sports Phys Ther*. 1991;14(6):263-8.
12. McHugh ML. Interrater reliability: the kappa statistic. *Biochemia Medica*. 2012;22(3):276-82.
13. Brink Y, Louw QA. Clinical instruments: reliability and validity critical appraisal. *J Eval Clin Pract*. 2012;18(6):1126-32.
14. Cohen L, Kobayashi S, Simic M, Dennis S, Refshauge K, Pappas E. Non-radiographic methods of measuring global sagittal balance: a systematic review. *Scoliosis Spinal Disord*. 2017;12:30.
15. Adhia DB, Bussey MD, Ribeiro DC, Tumilty S, Milosavljevic S. Validity and reliability of palpation-digitization for non-invasive kinematic measurement - a systematic review. *Man Ther*. 2013;18(1):26-34.
16. Munro BH VM. *Statistical methods for health care research*: Philadelphia: Lipincott Williams & Wilkins; 2005. p. 239-58.
17. Lampe HH, Swierstra BA, Diepstraten AFM. Measurement of limb length inequality: comparison of clinical methods with orthoradiography in 190 children. *Acta Orthop Scand*. 1996;67(3):242-4.
18. Leard JS, Crane BA, Ball KA. Intrarater and interrater reliability of 22 clinical measures associated with lower quarter malalignment. *J Manipulative Physiol Ther*. 2009;32(4):270-6.
19. Brêtas DA, Nogueira JVF, Carneiro MV, de Souza RA, Simao AP. Analysis of intra-examiner reliability of the tape method measure to leg length discrepancy. *Fitness & Performance Journal (Online Edition)*. 2009;8(5):335-41.
20. Terry MA, Winell JJ, Green DW, Schneider R, Peterson M, Marx RG, et al. Measurement variance in limb length discrepancy - Clinical and radiographic assessment of interobserver and intraobserver variability. *J Pediatr Orthop*. 2005;25(2):197-201.
21. Middleton-Duff T, George K, Batterham A. The reliability and validity of the 'Tape' and 'Block' methods for assessing anatomical leg-length discrepancy. *Phys Ther Sport*. 2000;1(3):91-9.
22. Badii M, Wade AN, Collins DR, Nicolaou S, Kobza BJ, Kopec JA. Comparison of lifts versus tape measure in determining leg length discrepancy. *J Rheumatol*. 2014;41(8):1689-94.
23. Woerman AL, Binder-Macleod SA. Leg length discrepancy assessment: accuracy and precision in five clinical methods of evaluation. *J Orthop Sports Phys Ther*. 1984;5(5):230-9.
24. Harris I, Hatfield A, Walton J. Assessing leg length discrepancy after femoral fracture: clinical examination or computed tomography? *ANZ J Surg*. 2005;75(5):319-21.