Clin Orthop Relat Res (2009) 467:2098–2104
DOI 10.1007/s11999-009-0705-1

ORIGINAL ARTICLE

Tibial Torsion in Cerebral Palsy

Validity and Reliability of Measurement

Sang Hyeong Lee MD, Chin Youb Chung MD, Moon Seok Park MD, In Ho Choi MD, Tae-Joon Cho MD

Received: 28 July 2008 / Accepted: 6 January 2009 / Published online: 22 January 2009

Abstract Physical examinations of tibial torsion are used for preoperative planning and to assess outcomes of tibial osteomy in patients with cerebral palsy (CP). The thigh-foot angle (TFA) and transmalleolar axis (TMA) are commonly used, and the second toe test recently was introduced. However, the validity and reliability of the three methods have not been clarified. This study was performed to evaluate the validity and reliability of these physical measures. We recruited 18 patients (36 limbs) with CP. During reliability sessions, three raters with various levels of orthopaedic experience independently measured tibial torsion using the three different methods during one day before surgery. Validity was assessed by performing a correlation study between physical examination and two-dimensional computed tomographic (CT) findings. Interobserver reliability was greatest for the TMA followed by TFA and then by the second toe test with intraclass correlation coefficients of 0.92, 0.74, and 0.57, respectively. In terms of the concurrent validity, the correlation coefficients (r) for the CT measurements were 0.62, 0.52, and 0.55. When depicting tibial torsion by physical examination, all three methods had substantial validity, but test reliability and validity were highest for TMA measurements.

Level of Evidence: Level I, diagnostic study. See the Guidelines for Authors for a complete description of levels of evidence.

Each author certifies that he or she has no commercial associations (eg, consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.
Each author certifies that his or her institution has approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

S. H. Lee, C. Y. Chung, M. S. Park (✉)
Department of Orthopedic Surgery, Seoul National University Bundang Hospital, 300 Gumi-Dong, Bundang-Gu, Sungnam, Kyungki 463-707, Korea
e-mail: pmsmed@hanafos.com

I. H. Choi, T.-J. Cho
Department of Orthopedic Surgery, Seoul National University Children's Hospital, Seoul, Korea

Introduction

Since Staheli et al. described rotational profile, TFA and TMA have been used predominantly to measure tibial torsion [20, 21]. In patients with CP and myelomeningocele, to eliminate external tibial torsion, tibial derotation osteotomy frequently is performed [4–6, 19, 22]. Because physical measurements are obtained easily in outpatient, preoperative, and intraoperative settings, many surgeons rely on physical examinations. Furthermore, a new technique called the second toe test was proposed to measure tibial torsion [9]. Although physical examination has advantages, physical measurements have disadvantages such as lack of reproducibility.

Therefore, it is important to confirm a physical examination is as valid and reliable as an imaging study and which test is superior in determining tibial torsion.

The aim of this study was to determine the validity and reliability of physical measures of tibial torsion, namely, TFA, TMA, and the second toe test.

Materials and Methods

We enrolled 18 consecutive patients with CP meeting the following criteria: unilateral or bilateral involvement

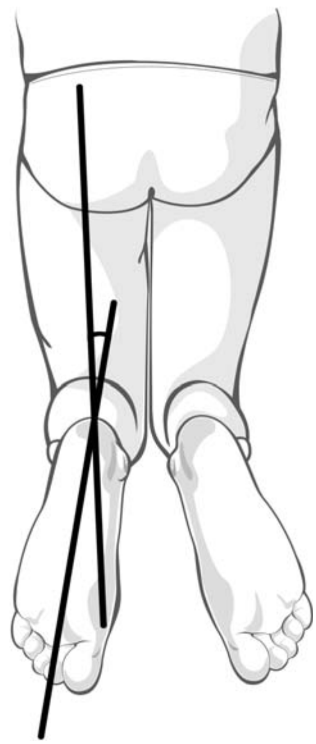

Fig. 1 Measurement of the TFA, which is the angle between the longitudinal axis of the thigh and the longitudinal axis of the foot, is shown.

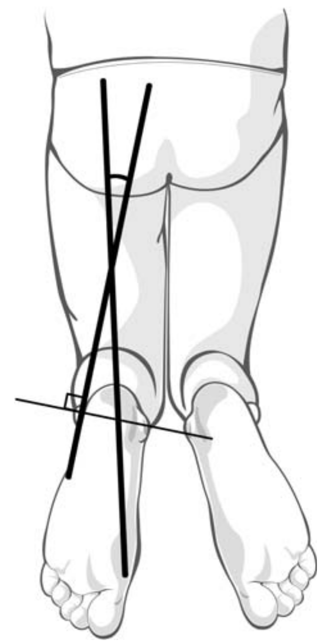

Fig. 2 Measurement of the TMA, which is the angle between the line of the longitudinal axis of the thigh and a line perpendicular to the axis connecting the most prominent portions of the medial and lateral malleolus, is shown.

planned for single-event multilevel surgery, including femoral derotation osteotomy or tibial derotation osteotomy, and the requirement for a CT scan to define torsional abnormalities. Prior precision analysis reveals the minimal sample size required as 36 limbs [1]. Our study had prior approval of our institutional ethics committees.

Consensus building was performed by all examiners before initiation of this study. Tibial torsion was measured using (1) TFA; (2) TMA; and (3) the second toe test. All methods were performed using a standard universal goniometer, which uses 1° increments with an arm length of 18 cm. Physical examinations were performed with patients in the prone position, and details of measurements were as follows: (1) TFA: With the patient's knee flexed 90°, the ankle in the neutral position, and the sole parallel to the floor, the angle between the longitudinal axis of the thigh and longitudinal axis of the foot were measured. A negative angle means internal rotation and a positive angle external rotation [21] (Fig. 1); (2) TMA: With the patient in the position described previously for TFA, the TMA was measured between the line of the longitudinal axis of the thigh and the line perpendicular to the axis that connected the most prominent portions of the medial and lateral malleolus [21] (Fig. 2); (3) second toe test: With the patient's knee fully extended, the lower extremity was rotated internally or externally until the second toe pointed directly toward the floor. Without changing the thigh rotation, the knee then was flexed 90°. The angle between the line vertical to the floor and the longitudinal axis of the shank was defined as the amount of tibial torsion [9] (Fig. 3).

Validities were assessed by comparing findings with those of a widely accepted method of measurement. Two-dimensional CT is viewed as an accurate means of obtaining quantitative measurements. CT (Mx8000IDT; Philips Medical Korea, Seoul, Korea) images at the levels of the proximal and distal tibia (within 2 cm of the articular surface) were selected. Physical examination validities were assessed using the concurrent validity method, which assesses validity versus a standard method. CT images were obtained on the same days as reliability sessions were performed and then archived. These archived images then were measured twice by two raters with an interval of 3 weeks (Fig. 4). The orders of the measurements were randomized using the method of random sampling without replacement The values from these four CT measurements were averaged to produce index values. Correlations were analyzed by averaging the physical examination findings of the three raters.

During the reliability session, interobserver reliability was examined by three raters with 21 (CY), 6 (MS), and 5 years (SH) orthopaedic experience. These three raters independently measured tibial torsion using the three different methods without knowledge of patients or the

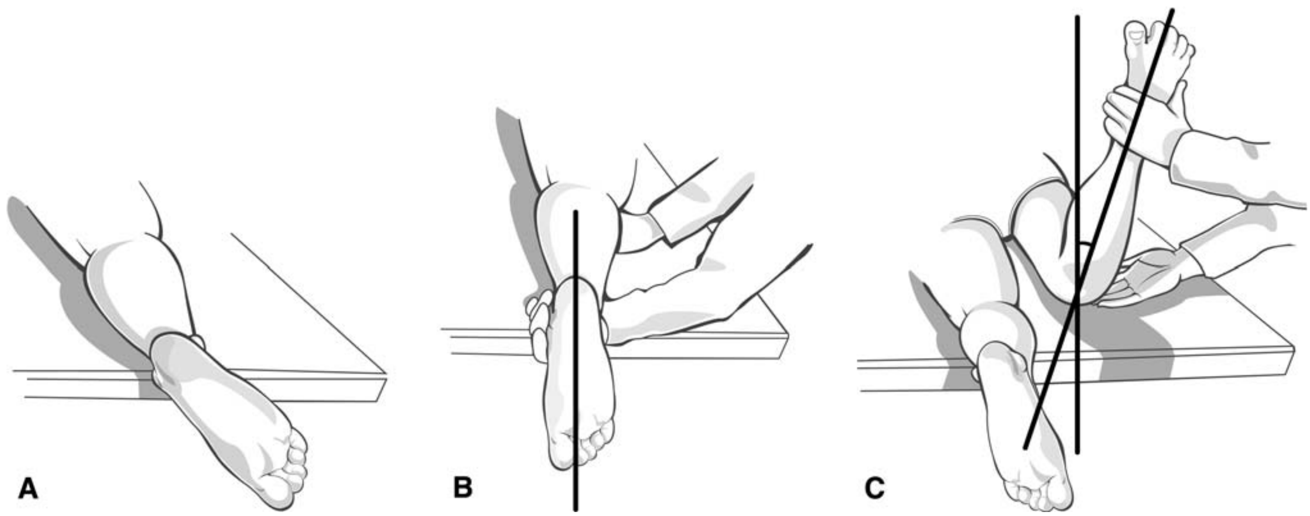

Fig. 3A–C Measurement using the second toe test is shown. (**A**) With the knee extended, the lower extremity was rotated internally or externally until the second toe pointed perpendicular to the floor. (**B**) Without changing the thigh rotation, the knee then was flexed 90°. (**C**) The angle between the line vertical to the floor and the longitudinal axis of the shank was then measured.

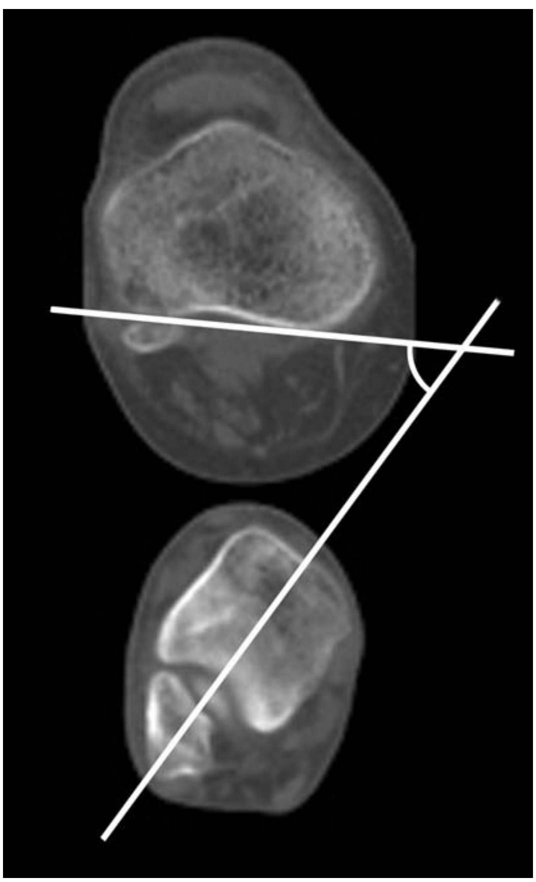

Fig. 4 Tibial torsion is defined as the angle between the line connecting both posterior condyles of the tibia and the line penetrating the medial and lateral malleolus. We regard this measurement as the reference standard for tibial torsion.

findings of other raters. Reliability sessions were conducted during a single day before surgery. Orders of measurement were assigned randomly and all data were collected by a research assistant who did not participate in the reliability sessions. An 18 × 3 matrix was acquired to determine interobserver reliability.

Prior precision analysis was performed to identify the minimal sample size required for the analysis. This study was designed to enable intraclass coefficients (ICCs) of reliability to be calculated at a target value of 0.8. In addition, we used the approximation suggested by Bonett [1]. Accordingly, when we set the 95% confidence interval to 0.2 for three raters, the minimal sample size was calculated to be 36 limbs. The ICC and their 95% confidence intervals were used to summarize the interobserver reliability. In our assumption, 36 limbs were representative of physical examinations and CT measurements (ie, the cases were considered a random factor) from a population with CP. The ability of a test to show interobserver reliability was evaluated using the one-way random effect model [15]. The ICC can be interpreted as follows: ICC = 1 means perfect reliability and ICC = 0 means opposite [8]. The normality of each variable was tested using a Kolmogorov-Smirnov test. Pearson's correlation coefficients were used to determine the relationship between the physical examinations and CT measurements. Linear regression was used to predict the true tibial torsion (CT values in this study) with physical examinations. The data were analyzed statistically using SPSS 11.0 (SPSS, Chicago, IL).

Results

All eighteen patients (36 limbs) in this study were diagnosed with CP, and their average age was 11.9 years (Table 1). The range of tibial torsion in the CT measurements was from −8° to 66° (Table 2). The interobserver and intraobserver reliabilities of the CT study ranged from 0.92 to 1.00 in the manner of the ICC (Table 3).

The validities of physical examination were presented as concurrent validity. Correlations between CT and physical examination findings were all significant, although correlation was greatest for the TMA (Table 4). Linear regression produced the following equations for relations between CT and physical examination measurements (Fig. 5):

$$\mathrm{Y\ (tibial\ torsion\ by\ CT)} = 18 + 0.9\mathrm{x}\ \mathrm{(TFA)}$$

$$\mathrm{Y\ (tibial\ torsion\ by\ CT)} = 12 + 0.8\mathrm{x}\ \mathrm{(TMA)}$$

$$\mathrm{Y\ (tibial\ torsion\ by\ CT)} = 13 + 1.5\mathrm{x}\ \mathrm{(second\ toe\ test)}$$

Table 1. Summary of patients

Characteristics	Patients
Gender (male/female)	11/7
GMFCS level (I/II/III)	5/7/6
Age (years)	11.9 (4.3, 6–20)

Data are provided as means (standard deviation, range); GMFCS = gross motor function classification system.

Table 2. Summary of test data

Examination	Average	Range	Standard deviation
Thigh-foot angle	10.9	−15 to 32	8.9
Transmalleolar axis	19.7	−18 to 42	11.8
Second toe test	10.7	−11 to 25	6.6
CT measurement	28.5	−8 to 66	13.0

Table 3. Intraobserver and interobserver reliabilities of CT measurements

Reliability	Rater	Intraclass correlation coefficient	95% confidence interval
Intraobserver	First rater	0.98	0.97–0.99
	Second rater	0.99	0.99–1.00
Interobserver	First session	0.96	0.93–0.98
	Second session	0.96	0.92–0.98
Overall		0.97	0.95–0.98

Data are tested by one-way random effect model.

Table 4. Concurrent validities of physical examination tests and CT measurements

Examination	Correlation coefficient (r)	p Value
Thigh-foot angle	0.52	0.001
Transmalleolar axis	0.62	< 0.001
Second toe test	0.55	0.001

Normality of the variables was verified using a Kolmogorov-Smirnov test, and the correlation coefficients were calculated using the Pearson correlation.

Regarding the reliability of a physical examination, the interobserver reliability was greatest for the TMA followed by the TFA and then by the second toe test (Table 5).

Discussion

The main aims of this study were to determine if a physical examination is as valid and reliable as a CT study and which of three physical examinations was best in determining the tibial torsion. The study showed the TMA has greater validity and reliability.

Before discussing the clinical implications, it is important to address the limitations of the study. First, the strength of the validity relied on the concurrent validity. Although direct measurements using necropsy specimens are most accurate, these obviously are unavailable in a clinical situation. The concurrent validity can be defined as the correlation with a reference standard method, CT in the current study. The validity and reliability of the reference standard method should be documented before addressing the concurrent validity of physical examinations. Although the reliability of a CT study was tested (Table 3), the validity of CT was not tested in the current study. CT is widely used to measure tibial torsion. The validity of CT measurements of tibial torsion was determined in cadaveric studies [3, 11]. Its use has been advocated in numerous studies [2, 3, 7, 11, 12, 25]. Other methods such as MRI [17, 18, 24] and ultrasound [2, 10] also are valid tools for measuring the level of tibial torsion and can be used as reference standards. However, because CT was invented earlier, studies on the validity of MRI and ultrasound have used CT as the reference standard [2, 18]. The reference lines for CT in this study are the bimalleolar and posterior condylar axes (Fig. 4). For the distal reference line, some studies advocated the bimalleolar axis [7, 11] as the reference line. Some advocated other axes such as the center of the tibia [12] or dorsal border of the tibia [2]. For the proximal reference line, the transtibial [11] and posterior condylar axes [2, 7, 12] are widely used. However, these two reference lines are similar in the juxtaarticular area [7].

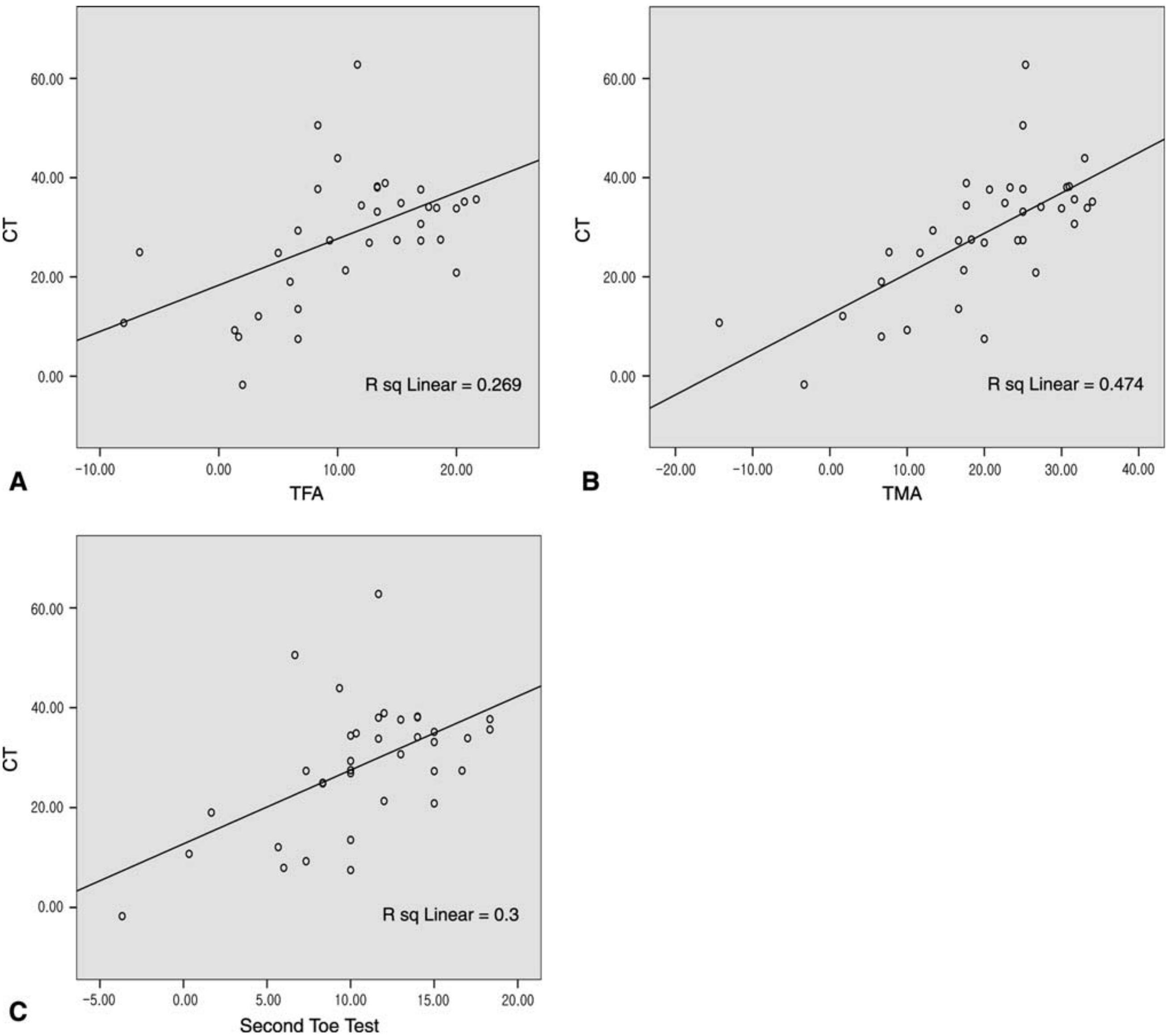

Fig. 5A–C (**A**) The relationship between CT measurements and the TFA is shown. (**B**) The linear relationship between CT measurements and the TMA and (**C**) between CT measurements and second toe test findings are shown.

Table 5. Interobserver reliabilities of physical examinations

Examination	Intraclass correlation coefficient	95% confidence interval
Thigh-foot angle	0.74	0.55–0.86
Transmalleolar axis	0.91	0.85–0.95
Second toe test	0.57	0.25–0.76

Data are tested by one-way random effect model.

The second limitation is the setting of the measurements was in a preoperative and outpatient setting. It is possible physical examinations with patients under anesthesia could improve reliability and validity [9]. Furthermore, the second toe test and TFA after foot deformity correction might have shown improved reliability and validity [9]. Although validity and reliability in a preoperative and outpatient setting are important in decision-making, readers should consider these limitations when interpreting these results.

After Staheli et al. described tibial torsion [21] by TFA and TMA, numerous authors examined these measurements [7, 11, 13, 14, 23]. The validity of TFA or TMA have been described [16, 23]. One study compared the physical examination and CT findings [23], whereas the other compared the physical examination findings with direct tibial torsion measurements in cadavers [16]. However, these two studies showed disagreement with the

Table 6. Studies of validity of the transmalleolar axis

Study	Number of subjects	Reference standard	Range of torsion	Correlation coefficient (r)	Clinical use
Stuberg et al. [23]	17	CT study	10°–42°	*	Advocate
Milner and Soames [16]	10	Direct measurement	21°–47°	0.33	Decline
Current study	18	CT study	−8°–66°	0.62	Advocate

* Data were not provided in the original article; CT = computed tomography.

validity of the TMA (Table 6). The study which showed poor validity of the TMA used the reference standard with direct measurement of a cadaver [16]. Although they released the soft tissue structure to facilitate knee and ankle motion, it is difficult to perform a physical examination on a cadaver. Our study adds the results of the second toe test and shows physical examinations have reasonable validity.

To achieve adequate reliability, there are two prerequisites. First, the subject matrix (number of subjects × number of raters) should be large enough. For practical reasons, we considered that raters be limited to three for physical examinations for patient compliance reasons. Thus, we limited raters to three and recruited sufficient subjects. Second, the range of variables should be wide enough. In a narrow range of variables, one can easily assume that the reliability is better than a wide range; that is, the correlation can be higher when the subjects are homogenous and lower when the subjects are heterogeneous. Moreover, the correlation can be overestimated in a homogenous population. This study showed significant variability (−8° to 66°), and the heterogeneity is believed to be sufficient.

The physical examinations used in the current study have unique advantages and disadvantages. The second toe test has three disadvantages. First, this test has two measurement reference lines, ie, the longitudinal axis of the lower leg and the line vertical to the floor. However, in patients with bowed legs, it is difficult to define the longitudinal axis of the lower leg. Thus, in patients with bowed legs, tibia vara, or tibia valga [9], the longitudinal axis should be used as a measurement reference for the proximal tibia rather than the whole tibia. Moreover, when performing measurements using the second toe test, the knee must be flexed and the foot progression angle must be maintained. However, in older and obese patients, flexing the knee and holding the thigh firmly without rotation may be difficult. Foot deformities, varus, or valgus affect test results because one of the reference lines is the longitudinal axis of the foot, which must be maintained throughout the tests. The TFA is influenced by foot deformity, ankle instability, and knee instability, and the longitudinal axis of the thigh cannot be defined easily in obese patients. In obese patients, TMA measurements are problematic because it is difficult to define the longitudinal axis of the thigh. Moreover, knee rotatory instability can cause unreliable results. However, in transmalleolar measurements, foot deformities and ankle instability do not interfere with results.

When measuring tibial torsion in an outpatient or preoperative setting, the test validity and reliability were greatest for the TMA method.

Acknowledgments We thank SungJu Kim for statistical advice and Mi Seon Ryu for assistance in data collection.

References

1. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med.* 2002;21:1331–1335.
2. Butler-Manuel PA, Guy RL, Heatley FW. Measurement of tibial torsion: a new technique applicable to ultrasound and computed tomography. *Br J Radiol.* 1992;65:119–126.
3. Clementz BG, Magnusson A. Assessment of tibial torsion employing fluoroscopy, computed tomography and the cryosectioning technique. *Acta Radiol.* 1989;30:75–80.
4. Dias LS, Jasty MJ, Collins P. Rotational deformities of the lower limb in myelomeningocele. Evaluation and treatment. *J Bone Joint Surg Am.* 1984;66:215–223.
5. Dodgin DA, De Swart RJ, Stefko RM, Wenger DR, Ko JY. Distal tibial/fibular derotation osteotomy for correction of tibial torsion: review of technique and results in 63 cases. *J Pediatr Orthop.* 1998;18:95–101.
6. Dunteman RC, Vankoski SJ, Dias LS. Internal derotation osteotomy of the tibia: pre- and postoperative gait analysis in persons with high sacral myelomeningocele. *J Pediatr Orthop.* 2000;20:623–628.
7. Eckhoff DG, Johnson KK. Three-dimensional computed tomography reconstruction of tibial torsion. *Clin Orthop Relat Res.* 1994;302:42–46.
8. Fleiss JL. *Statistical Methods for Rates and Proportions.* New York, NY: John Wiley & Sons; 1981:218.
9. Gage J. *The Treatment of Gait Problems in Cerebral Palsy.* London, England: Mac Keith Press; 2004:363.
10. Hudson D, Royer T, Richards J. Ultrasound measurements of torsions in the tibia and femur. *J Bone Joint Surg Am.* 2006;88:138–143.
11. Jakob RP, Haertel M, Stussi E. Tibial torsion calculated by computerised tomography and compared to other methods of measurement. *J Bone Joint Surg Br.* 1980;62:238–242.
12. Jend HH, Heller M, Dallek M, Schoettle H. Measurement of tibial torsion by computer tomography. *Acta Radiol Diagn (Stockh).* 1981;22:271–276.
13. Laasonen EM, Jokio P, Lindholm TS. Tibial torsion measured by computed tomography. *Acta Radiol Diagn (Stockh).* 1984;25:325–329.

14. le Damany PG. Technique of tibial tropometry. 1903. *Clin Orthop Relat Res.* 1994;302:4–10; discussion 2–3.
15. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods.* 1996;1: 30–46.
16. Milner CE, Soames RW. A comparison of four in vivo methods of measuring tibial torsion. *J Anat.* 1998;193:139–144.
17. Schneider B, Laubenberger J, Jemlich S, Groene K, Weber HM, Langer M. Measurement of femoral antetorsion and tibial torsion by magnetic resonance imaging. *Br J Radiol.* 1997;70:575–579.
18. Schneider B, Laubenberger J, Wildner M, Exne V, Langer M. NMR tomographic measurement of femoral ante-torsion and tibial torsion [in German]. *Rofo.* 1995;162:229–231.
19. Selber P, Filho ER, Dallalana R, Pirpiris M, Nattrass GR, Graham HK. Supramalleolar derotation osteotomy of the tibia, with T plate fixation: technique and results in patients with neuromuscular disease. *J Bone Joint Surg Br.* 2004;86:1170–1175.
20. Staheli LT. In-toeing and out-toeing in children. *J Fam Pract.* 1983;16:1005–1011.
21. Staheli LT, Corbett M, Wyss C, King H. Lower-extremity rotational problems in children: normal values to guide management. *J Bone Joint Surg Am.* 1985;67:39–47.
22. Stefko RM, de Swart RJ, Dodgin DA, Wyatt MP, Kaufman KR, Sutherland DH, Chambers HG. Kinematic and kinetic analysis of distal derotational osteotomy of the leg in children with cerebral palsy. *J Pediatr Orthop.* 1998;18:81–87.
23. Stuberg W, Temme J, Kaplan P, Clarke A, Fuchs R. Measurement of tibial torsion and thigh-foot angle using goniometry and computed tomography. *Clin Orthop Relat Res.* 1991;272: 208–212.
24. Tamari K, Tinley P, Briffa K, Breidahl W. Validity and reliability of existing and modified clinical methods of measuring femoral and tibiofibular torsion in healthy subjects: use of different reference axes may improve reliability. *Clin Anat.* 2005;18:46–55.
25. Widjaja PM, Ermers JW, Sijbrandij S, Damsma H, Klinkhamer AC. Technique of torsion measurement of the lower extremity using computed tomography. *J Comput Assist Tomogr.* 1985;9: 466–470.