Reproducibility of Radiographic Measurements in Assessment of Congenital Talipes Equinovarus

John Thometz, MD, Robert Manz, MD, Xue-Cheng Liu, MD, PhD, John Klein, PhD, and Barbara Manz-Friesth, DNS, RN

Abstract

Six commonly measured parameters in the assessment of congenital clubfoot were retrospectively assessed from standardized preoperative and intraoperative radiographs taken during operative complete subtalar release. These radiographic parameters were measured in 30 feet by 6 observers at 2 separate readings. The observers were orthopedic residents in different stages of training. Between-observers intraclass correlation coefficients (ICCs) were computed for each parameter.

All radiographic parameters were found to be reproducible across time and observers (range of preoperative intraobserver ICCs, 0.84-0.99; range of preoperative interobserver ICCs, 0.93-0.99), except for intraoperative anteroposterior (AP) talar–first metatarsal angle (intraoperative intraobserver ICC, 0.79) and lateral talocalcaneal angle (intraoperative interobserver ICC, 0.81). Differences in mean preoperative measurements between observers and time were tested by analysis of variance. There were no significant differences between observers and time in the 6 preoperative measurements ($P>.05$), except for intraoperative AP talar–first metatarsal angle, AP talocalcaneal angle, and degree of AP calcaneocuboid subluxation, which were significantly different ($P<.05$).

Our results support use of radiographs as a reliable method for guiding care in patients with clubfoot and as a reproducible method that physicians can use for comparisons.

Dr. Thometz is Chief of Pediatric Orthopedics, Department of Orthopaedic Surgery, Children's Hospital of Wisconsin, Milwaukee, Wisconsin, and Professor, Medical College of Wisconsin, Milwaukee, Wisconsin.
Dr. Manz is Resident, Medical College of Wisconsin, Milwaukee, Wisconsin.
Dr. Liu is Associate Professor, Department of Orthopaedic Surgery, Medical College of Wisconsin, and Director, MusculoSkeleton Functional Assessment Center, Children's Hospital of Wisconsin, Milwaukee, Wisconsin.
Dr. Klein is Professor and Director, Department of Population Health Biostatistics, Medical College of Wisconsin, Milwaukee, Wisconsin.
Dr. Manz-Friesth is Associate Professor, Indiana University School of Nursing, Indianapolis, Indiana.

Address correspondence to: John G. Thometz, MD, Department of Orthopaedic Surgery, Children's Hospital of Wisconsin, 9000 W Wisconsin Ave, PO Box 1997, Suite C360, Milwaukee, WI 53201 (tel, 414-337-7300; fax, 414-337-7337; e-mail, jthometz@chw.org).

Am J Orthop. 2009;38(12):617-620.

Since 1876, when Barwell[1] proposed using radiographs in the treatment of congenital talipes equinovarus, multiple angular measurements have been devised to assess clubfoot deformity. Both anteroposterior (AP) and lateral radiographs have been used to address all deforming components. Classically, hindfoot varus has been assessed with the AP talocalcaneal angle.[2-4] Similarly, the lateral talocalcaneal angle has helped in the assessment of hindfoot equinus.[3] Degree of calcaneocuboid subluxation has been judged using a grading classification developed by Thometz and Simons.[5] Severity of forefoot supination and adduction and hindfoot cavus is further assessed using the lateral talar–first metatarsal angle. Analytical radiographic assessment of the AP talar–first metatarsal angle has allowed interpretation of navicular subluxation before ossification.[6]

However, despite the continued instruction in present-day orthopedics of "normal" and "clubfoot-typical" radiographic signs, a large amount of controversy persists regarding the role and validity of radiographic findings in the assessment and initial treatment of congenital clubfoot. This likely stems from a multitude of issues, including the limitations of a two-dimensional assessment of a complex three-dimensional disease and the fact that the ossification centers of the cartilage anlages of the pediatric foot are often indistinct or not present during the first year of life.[7] Furthermore, there is a paucity of studies addressing the reproducibility of these radiographic angles.

With the development of the Ponseti techniques, the vast majority of clubfeet can be treated nonoperatively. However, there remains a severe-clubfeet subgroup of patients who require more extensive surgery and for whom radiographs can be helpful in deformity assessment. In addition, other congenital foot deformities may benefit from radiographic assessment.

In this study, we assessed the reproducibility of 6 commonly used radiographic parameters measured by 6 orthopedic residents on radiographs taken during the point in the clubfoot patient's life when this information would most significantly aid in the treatment of the disease—before and during surgery.

Methods and Materials

We reviewed the hospital records and radiographs of 19 patients treated with operative complete subtalar release at the Children's Hospital of Wisconsin in Milwaukee between 1985 and 2000. Neuromuscular disorders (eg, arthrogryposis multiplex congenita, myelomeningocele) were excluded

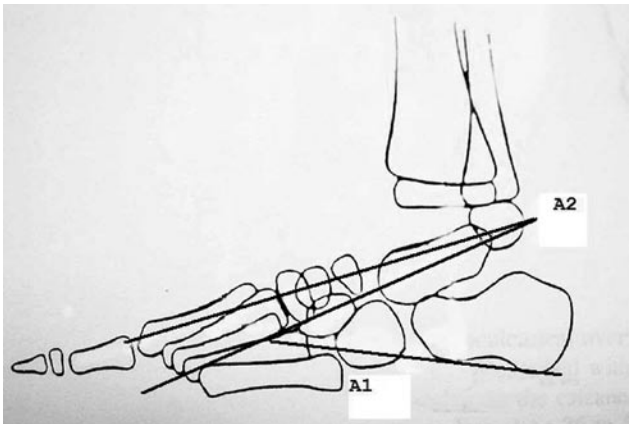

Figure 1. On lateral radiograph of foot, measurements of talocalcaneal angle (A1) and talar–first metatarsal angle (A2).

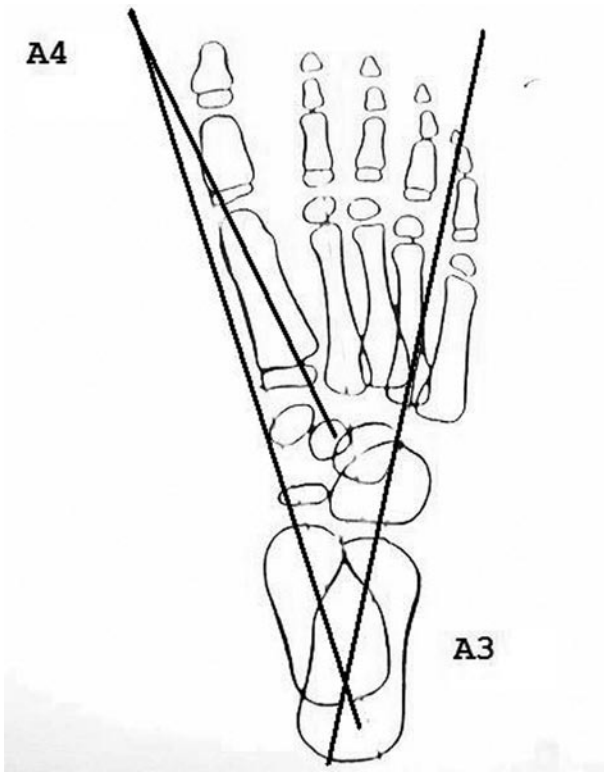

Figure 2. On anteroposterior radiograph of foot, measurements of talocalcaneal angle (A3) and talar–first metatarsal angle (A4).

from analysis. As 11 of the 19 patients had bilateral talipes equinovarus, a total of 30 clubfeet were assessed radiographically in this study. Each patient underwent AP and lateral radiographic imaging both before and during surgery using the standardized technique described by Simons[6]; thus, there were 2 sets of radiographs for each patient. Mean patient age at initial radiographic assessment and at operative intervention for the 30 clubfeet was 6 months and 9 months, respectively.

Clear undeveloped film was superimposed over each of the fixed radiographs on a radiographic viewer to provide an easily erasable or replaceable medium for recording measurements. All parameters were recorded with a standardized pencil and measured with a standardized goniometer. Marks were completely erased between readings. Six observers participated in the study. All were orthopedic residents who had had at least 1 year of orthopedic training and had undergone a short tutorial on proper techniques for radiographic assessment in the setting of clubfoot. Six different parameters were recorded from each radiograph by each observer at 2 sittings separated by at least 2 days. Therefore, a total of 720 measurements (360 preoperative, 360 intraoperative) was taken on the 30 clubfeet.

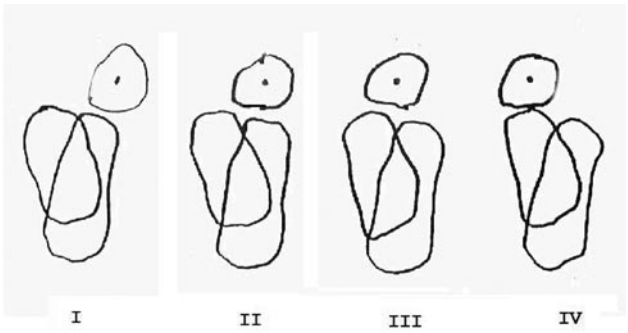

Figure 3. On anteroposterior radiograph of foot, classification of calcaneocuboid subluxation (I-IV).

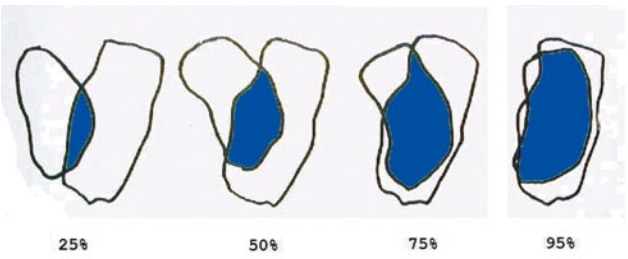

Figure 4. On anteroposterior radiograph of foot, percentage of overlap between talar and calcaneal bone (25%-95%).

The parameters assessed were lateral talocalcaneal angle (A1) and lateral talar–first metatarsal angle (A2) (Figure 1); AP talocalcaneal angle (A3) and AP talar–first metatarsal angle (A4) (Figure 2); calcaneocuboid subluxation (Figure 3); and talocalcaneal overlap (Figure 4). All measurements were recorded in values of degrees, except calcaneocuboid subluxation and talocalcaneal overlap, which were classified into number grading systems, one of which was developed by Thometz and Simons.[5] Observers were encouraged to mark the midpoint of the ossific nucleus at each end of the bone before drawing the longitudinal axis on bones that are often difficult to accurately assess, such as the talus. Furthermore, observers were asked to draw the lateral talocalcaneal angle using the plantar surface of the calcaneus as the axis. A poster depicting proper methods for measuring parameters was available for immediate reference in the vicinity of the reading area.

Intraobserver reliability and interobserver reliability of each angle were then checked. Intraclass correlation analysis was used to assess general reliability, and 2-way analysis of variance (ANOVA), performed with SAS software (SAS Institute Inc, Cary, NC) was used to more closely assess individual variability and standard deviation.

Table. Intraobserver and Interobserver Reliability for Measuring Talipes Equinovarus on Radiographs

Radiographic Measurement	Intraclass Correlation Coefficients (95% Confidence Intervals)			
	Preoperative		Intraoperative	
	Intraobserver	Interobserver	Intraobserver	Interobserver
Anteroposterior talar–first metatarsal angle	0.97 (0.88-0.99)	0.99 (0.98-0.99)	0.79 (0.16-0.94)	0.84 (0.72-0.91)
Anteroposterior talocalcaneal angle	0.84 (0.43-0.97)	0.93 (0.89-0.96)	0.89 (0.60-0.98)	0.88 (0.79-0.94)
Calcaneocuboid subluxation	0.90 (0.72-0.97)	0.95 (0.91-0.97)	0.85 (0.46-0.97)	0.87 (0.78-0.93)
Talocalcaneal overlap	0.98 (0.91-0.99)	0.99 (0.98-0.99)	0.95 (0.82-0.99)	0.98 (0.96-0.99)
Lateral talocalcaneal angle	0.90 (0.76-0.97)	0.95 (0.92-0.98)	0.81 (0.29-0.98)	0.93 (0.88-0.96)
Lateral talar–first metatarsal angle	0.99 (0.95-0.99)	0.99 (0.98-0.99)	0.94 (0.79-0.99)	0.97 (0.95-0.98)

Results

All data from the radiographic analysis were first analyzed with a 2-way fixed-effects model to establish an intraclass correlation coefficient (ICC) for each angle.[8] Intraclass correlation analysis allows one to assess general reliability between 2 variables; in this case, each angle was first analyzed across time and then across observers. Coefficients are values between 0 and 1, with values closer to 1 representing stronger agreement between the raters. A value of 0.7 was predetermined to establish a level of high correlation.[8] The Table demonstrates both intraobserver and interobserver reliability for all angles with corresponding ICCs and 95% confidence intervals (95% CIs). All preoperative radiographic angles were found to have strong interobserver and intraobserver reliability, with ICCs all higher than 0.7 (range, 0.84-0.99). Of the preoperative parameters, only AP talocalcaneal angle had a 95% CI (0.43-0.97) that included an ICC of 0.7. Intraoperative assessment similarly demonstrated statistically significant higher ICCs in all angles for both intraobserver and interobserver reliability (>0.79), except for AP talar–first metatarsal angle (0.16-0.94) and lateral talocalcaneal angle (0.29-0.98), which included 0.7 for intraobserver ICC.

Two-way ANOVA with each angle as a dependent factor and with observer and time as independent factors was also used to calculate the mean value and standard deviation for each angle. Main effects tested include observer, time, and interaction between observer and time. All preoperative angles were once again found to be reliable across observers ($P>.05$). However, statistically significant ($P<.05$) differences were found between observers on AP talar–first metatarsal angle (1 SD of 3°) and AP talocalcaneal angle (1 SD of 5°) ($P<.05$). Review of the individual observer outcomes revealed that all these angles suffered from observer 6's results. Observer 6 had an AP talocalcaneal angle of 29°, as compared with the mean of 36° from other 5 observers (range, 36°-37°). Despite his extreme variability with regard to the mean values of others, his standard deviation remained extremely close and comparable. This implies that, though he was inaccurate, at least his values were consistent. It should be noted that the standard deviation derived in the 2-way ANOVA cannot be directly interpreted to equate a level or degree of expected variance between angles, as these are derived from the mean values of each angle in all patients with varying degrees of deformity. Rather, it represents a degree of reproducibility within each observer.

Discussion

Although teaching of "clubfoot-typical" radiographic parameters remains common in orthopedic literature, controversy persists regarding use of radiographs in the treatment of talipes equinovarus and other pediatric foot deformities.[4,5,9-15] Some investigators have attempted to compare functional outcomes and revision rates with radiographic findings.[13,16] However, very few studies have been conducted to determine whether the interpretation of radiographic angles in the pediatric foot offers any reproducibility at all, thus calling any conclusions made from these studies into question.

Cook and colleagues[10] studied interobserver and intraobserver variability in the radiographic assessment of patients with metatarsus adductus. With 4 observers, he evaluated 4 radiographic angles, including talus–first metatarsal angle, AP talocalcaneal angle, and lateral talocalcaneal angles. He found the angles to have a large degree of interobserver and intraobserver variability and concluded that treatment should be guided by clinical examination. However, patients in the study ranged in age from 9 weeks to 19 months, and no mean was defined. Porter and colleagues[14] studied the reproducibility of several radiographic measurements in clubfoot and found them to be more reliable than clinical examination. However, radiographic assessment was concentrated on patients ranging in age from 4 to 12 months. In our study, we included patients whose mean age at the preoperative and intraoperative assessments was 6 and 9 months, respectively. We did this because we believe that assessment and treatment can have their most significant impact during this age. Delaying definitive preoperative radiographic assessment to 6 months allows further definition of tarsal bones with increasing ossification, whereas including younger patients would likely substantially decrease measurement reproducibility and would aid little in the ultimate treatment, particularly surgery.

At our institution, radiographs are routinely used both before and during surgery to guide care and judge surgical release and correction. Other studies have shown that clinical examination alone is less reproducible than radiographs in terms of quantifying clubfoot severity and the need for surgery, particularly with borderline cases.[11,14,17] Our study was the most comprehensive and well-controlled study addressing the reliability of radiographs in patients with clubfeet. Its results show that 6 commonly used radiographic angles are

reproducible across both time and observers in the preoperative setting and that 4 of these angles are reproducible in intraoperative assessment (even these intraoperative parameters showed agreement, with the exception of 1 observer).

In general, preoperative angles were found to be more reliable than intraoperative angles, except for AP talocalcaneal angle with its wide range of 95% CIs (Table). This seems somewhat counterintuitive given the more immature status of tarsal ossification. A possible explanation could include presence of intraoperative hardware, which may have drawn some observers to design an axis line along the hardware instead of along the true longitudinal axis of the bone. However, closer scrutiny of the data revealed that AP talar–first metatarsal angle and lateral intraoperative talocalcaneal angle, which suffered on the ICC analysis (with a wide range of 95% CIs), and the 2 intraoperative angles found to be unreliable in the 2-way ANOVA were all statistically affected by the same observer. This illustrates another important point for interpretation. Although all observers were trained in radiographic interpretation of clubfoot angles, these observers were residents with a relatively low level of orthopedic training. One would expect the reliability of these measurements to improve when performed by qualified pediatric orthopedic surgeons who are routinely exposed to the disease entity. Furthermore, the reliability of the angles may have failed simply because of observer fatigue. Each observer measured a total of 720 angles, which, though needed to establish statistical power, is a fairly cumbersome undertaking.

Our study demonstrated that the value of the interobserver ICC is always higher than that of the intraobserver ICC in both preoperative and intraoperative measurements. This indicates that, though the radiographic measuring skill of 1 observer deviated from that of the other 5 observers, overall there were no significant variations between observers. In our model, interobserver reliability was directly dependent on intraobserver reliability. In other words, to establish a strong correlation across observers, one must first establish a strong correlation across time. Our radiographic measurements were reliable across age groups, as compared with studies conducted in adult patients with other foot deformities[18,19] showing intraobserver reliability higher than interobserver reliability for intermetatarsal distance, calcaneal inclination, and so forth.

This study has a few identifiable shortcomings. First, and probably most important, is whether the standardized technique of taking radiographs is in itself reproducible. Unfortunately, to determine this, one would need to take multiple radiographs in the same patient over short periods to negate the effects of growth. This may be neither feasible nor ethical in the setting of the pediatric patient with clubfoot. However, there is reasonable evidence to suggest that implementing standardized techniques when taking radiographs would produce a reliable result.[14,20] Furthermore, variance produced by minor malpositioning errors is likely relatively small.[20] We believe it is important to emphasize that each radiograph must be scrutinized for adequate technique before placing significance on any measurements. Simons[20] identified several "clues" that can aid in pinpointing "unacceptable" radiographs. Lateral radiographic signs include extreme posterior positioning of the fibula with respect to the tibia, loss of metatarsal overlap, and lack of ankle dorsiflexion. Similarly, AP signs include visualization of the shafts of the tibia and fibula, recognition of the anterior ends of the talus and calcaneus at different levels, and significant metatarsal overlap. Another possible shortcoming of this study was the relatively short period between each observer's 2 readings. We believe that this is less important given that each observer measured such a large quantity of angles that the likelihood of recalling a previous measurement was low.

Conclusions

Radiographic angles measured after taking standardized radiographs are reliable both from interobserver and intraobserver standpoints. Radiographic evaluation is sufficiently reproducible to be used as an aid in assessment and treatment of children with clubfeet. The next step is to study whether radiographic parameters correlate with functional outcomes for clubfoot.

Authors' Disclosure Statement

The authors report no actual or potential conflict of interest in relation to this article.

References

1. Barwell R. On various forms of talipes as depicted by x-rays. *Lancet.* 1896;148(3803):160-162.
2. Davis LA, Hatt WS. Congenital abnormalities of the feet. *Radiology.* 1955;64(6):818-825.
3. Heywood AWB. The mechanics of the hindfoot in clubfoot as demonstrated radiographically. *J Bone Joint Surg Br.* 1964;46:102-107.
4. Wisbrun W. Neue gesichtspunkte zum redressement des angeborenen klumpfusses und daraus sich ergebende schlussfolgerungen bezuglich der atiologie. *Arch Orthop Unfall-Chir.* 1932;31(1):451-464.
5. Thometz JG, Simons GW. Deformity of the calcaneocuboid joint in patients who have talipes equinovarus. *J Bone Joint Surg Am.* 1993;75(2):190-195.
6. Simons GW. Analytical radiography of club feet. *J Bone Joint Surg Br.* 1977;59(4):485-489.
7. Howard CB, Benson MK. The ossific nuclei and the cartilage anlage of the talus and calcaneum. *J Bone Joint Surg Br.* 1992;74(4):620-623.
8. Armitage P. Correlation. In: Armitage P, Colton T, eds. *Encyclopedia of Biostatistics*. Vol 1. New York, NY: Wiley; 1998:971-976.
9. Bansal VP, Daniel J, Rai J. Radiological score in the assessment of clubfoot. *Int Orthop.* 1988;12(3):181-185.
10. Cook DA, Breed AL, Cook T, DeSmet AD, Muehle CM. Observer variability in the radiographic measurement and classification of metatarsus adductus. *J Pediatr Orthop.* 1992;12(1):86-89.
11. Cummings RJ, Davidson RS, Armstrong PF, Lehman WB. Congenital clubfoot. *J Bone Joint Surg Am.* 2002;84(2):290-308.
12. Johnston CE, Hobatho MC, Baker KJ, Baunin C. Three-dimensional analysis of clubfoot deformity by computed tomography. *J Pediatr Orthop.* 1995;4(1):39-48.
13. Magone JB, Torch MA, Clark RN, Kean J. Comparative review of surgical treatment of the idiopathic clubfoot by three different procedures at Columbus Children's Hospital. *J Pediatr Orthop.* 1989;9(1):49-58.
14. Porter RW, Roy A, Rippstein J. Assessment in congenital talipes equinovarus. *Foot Ankle.* 1990;11(1):16-21.
15. Herbsthofer B, Eckardt A, Rompe JD, Kullmer K. Significance of radiographic angle measurements in evaluation of congenital clubfoot. *Arch Orthop Trauma Surg.* 1998;117(6-7):324-329.
16. Moses W, Allen B, Linda I, Stasikelis PJ. Predictive value of intraoperative clubfoot radiographs on revision rates. *J Pediatr Orthop.* 2000;20(4):529-532.
17. Bensahal H, Catterall A, Dimeglio A. Practical applications in idiopathic clubfoot: a retrospective multicentric study in EPOS. *J Pediatr Orthop.* 1990;10(2):186-188.
18. Saro C, Johnson DN, Maritinez DE, Argon J, Lindgren U, Fellandertsai L. Reliability of radiological and cosmetic measurements in hallux valgus. *Acta Radiol*. 2005;46(8):843-851.
19. Taranto MJ, Taranto J, Bryant A, Singer KP. Radiographic investigation of angular and linear measurements including first metatarsophalangeal joint dorsiflexion and rear foot to forefoot axis angle. *J Foot Ankle Surg*. 2005;44(3):190-199.
20. Simons GW. A standardized method for the radiographic evaluation of clubfeet. *Clin Orthop*. 1978;(135):107-118.