Correlation of Radiographic Measurements With Patient-Centered Outcomes in Hallux Valgus Surgery

Michael Matthews, DPM, AACFAS1, Erin Klein, DPM, FACFAS1, Angie Youssef, BA1, Lowell Weil, Jr, DPM, MBA, FACFAS1, Matthew Sorensen, DPM, FACFAS1, Lowell S. Weil, Sr, DPM, FACFAS1, and Adam Fleischer, DPM, FACFAS1

Abstract

**Background**: Evaluation of patients undergoing hallux valgus surgery has historically emphasized radiographic angles and relationships. However, patient-reported outcomes are increasingly important as health care systems trend towards a “value-based” delivery approach.

**Methods**: We conducted a retrospective analysis of pre-existing data in our practice to examine whether patient-reported outcomes after bunion surgery, determined via Foot and Ankle Outcome Scores (FAOS), correlated with radiographic parameters commonly measured in hallux valgus deformity. Pearson correlation statistics and simple and multiple linear regression models were used to identify important radiographic predictors. There were 80 patients (80 feet) with mean follow-up of 59.3 ± 11.6 weeks (median 55, range 45.7-96.3 weeks) with complete data.

**Results**: No radiographic measurement/variable achieved anything more than a weak correlation with any of the FAOS subscale scores at final follow-up; the study's best was postoperative first-second intermetatarsal (IM) angle with sports and recreation scores \((r = -0.328, P = .005)\). There was no correlation found between change in hallux valgus angle, change in first-second IM angle, magnitude of preoperative hallux valgus angle or magnitude of preoperative first-second IM angle \((P > .05\) for all). Furthermore, none of the study’s final multivariable models achieved an \(R^2 > 0.24\), and nearly all fell between 0.10 and 0.17.

**Conclusion**: We conclude that radiographic angles were not well correlated with patient-centered outcomes in hallux valgus surgery. This study calls into question the current emphasis that is placed on x-ray values both pre- and postoperatively.

**Level of clinical evidence**: Level III, comparative study.

**Keywords**: hallux valgus, surgery, outcome, PROM, FAOS, radiographs

Introduction

Hallux valgus is a complicated pathology that has had more than a hundred different operative solutions advocated for its correction.4 A common theme among past and even many recent studies is the value that surgeons have placed on radiographic and other “physician-centered” clinical outcomes like postoperative range of motion, toe purchase, and toe alignment.3,11,16 Such an approach to hallux valgus surgery tends to de-emphasize what is a primary focus of any procedure: the patient and his or her experience. Recently, there has been a shift in the medial literature toward obtaining and reporting on patient-centered outcomes before and after encounters with the health care system. Although advancements are being made in this area regularly, there is still a sizeable void in the published literature evaluating hallux valgus correction from the patient’s point of view.

Most experienced foot and ankle surgeons have encountered the problem of “perfect” radiographic correction of

---

1Weil Foot and Ankle Institute, Des Plaines, IL, USA

**Corresponding Author:**
Michael Matthews, DPM, AACFAS, Weil Foot and Ankle Institute, 1455 East Golf Road, Des Plaines, IL 60016, USA.
Email: mmatthews@weil4feet.com
hallux valgus, only to have the patient displeased with the outcome. Conversely, many patients with less than perfect correction often express satisfaction. This not so uncommon observation has led some investigators to call into question the value of radiographic measurements in hallux valgus assessment and surgery. Thordarson and colleagues first examined this phenomenon in 2001 reporting on a series of 196 patients who underwent hallux valgus surgery, with a follow-up study in 2005. Their group found no relation between the amount of deformity correction, amount of postoperative residual deformity, or magnitude of correction and outcome with regard to various physician- and patient-reported metrics (ie, SF-36, AAOS Lower Limb Outcomes Data Collection Questionnaire, and AOFAS forefoot score). To our knowledge, no group has yet attempted to replicate the findings of Thordarson and colleagues. Furthermore, although clearly insightful, their original work did not make use of a patient-centered outcome measure validated for use in hallux valgus surgery. The authors also did not examine the effects of sesamoid position on outcomes—a relationship that has become more popular in recent years as it can serve as a proxy for frontal plane rotation of the first metatarsal head. Therefore, the purpose of this study was to compare patient-reported outcomes, using a validated outcome measure, to radiographic measurements before and after hallux valgus correction in a group of patients treated at a single urban-based foot and ankle specialty clinic.

Methods

We conducted a retrospective analysis of patients who underwent a scarf bunionectomy in our practice during the past 3 years. All data were already in existence at the time of the study. Exempt determination and HIPAA waiver was obtained from our local Institutional Review Board. Included patients underwent a scarf osteotomy and intra-articular release of the capsule as previously described. All patients were minimally weight bearing in a postoperative shoe after surgery and transitioned into running shoes starting at 1 week postoperatively. Eligible subjects had to have Foot and Ankle Outcome Scores (FAOS) completed preoperatively with respect to their operative foot and again at least 10 months postoperatively. In our practice, FAOS data are obtained on many of our operative patients to help with clinical decision making and to monitor postoperative progress. The FAOS is a validated patient-based tool for evaluating postoperative hallux valgus operative outcomes that has demonstrated internal consistency, convergent validity, and structural validity. The survey consists of 42 items and 5 subscales: pain, symptoms, function/activities of daily living, function/sport and recreation, and foot and ankle–related quality of life. Subscale scores are reported on a scale that ranges from 0 to 100, with higher values indicating better scores (eg, less pain, fewer symptoms, higher functioning).

Included subjects also had to have preoperative weight-bearing foot radiographs, and postoperative weight bearing foot radiographs at 6 weeks and final follow-up available for review. Furthermore, included subjects could not have appreciable first MTP joint arthritis, defined as Coughlin and Shurnas grade 2 or greater, where arthrodesis might be considered more appropriate. We included patients who had additional minor forefoot surgeries (eg, hammer toe repair, neurectomy) during their bunionectomy surgery, but excluded those who underwent concomitant lesser metatarsal osteotomy or midfoot/hindfoot osteotomy or arthrodesis.

All radiographs were evaluated by a single rater (M.M.) using commercially available software. The following parameters were evaluated: the first-second intermetatarsal (IM) angle, the hallux valgus angle (HVA), metatarsus adductus angle via the Engles angle, metatarsal protrusion distance using Nilsonne’s method and tibial sesamoid position using the method of Hardy and Clapham. Engle’s method for measuring the metatarsus angle described the angle between the longitudinal bisection of the second metatarsal and the longitudinal bisection of the second cuneiform (as opposed to the using the longitudinal axis of the lesser tarsus). Nilsonne’s method of examining metatarsal protrusion distance involved measuring the distance in millimeters between a perpendicular line to the longitudinal axis of the second metatarsal at the most distal point of the second metatarsal head to a parallel line that passed through the distal most point of the first metatarsal head.

There were 80 patients (80 feet, 75 female) with a mean age of 52.3 ± 13.3 years and mean follow-up of 59.3 ± 11.6 weeks (median 55, range 45.7–96.3 weeks). Most patients underwent a scarf/akin bunionectomy (63/80, 79%), whereas the remainder underwent scarf only to address the hallux valgus. Fourteen percent (11/80) underwent concomitant additional minor forefoot surgeries that consisted mostly of hammertoe correction and tailor’s bunionectomies without osteotomy. There were 7 radiographic hallux varus at final follow-up (HVA ranging from −1.6 to −13.7). Nearly all surgeries (78/80) were performed by one of the 2 surgeons (L.S.W. or L.W.J.).

In building the statistical models for analysis, radiographic covariates were grouped into continuous and dichotomous variables. Continuous variables included magnitude of preoperative HVA, magnitude of preoperative first-second IM angle, magnitude of preoperative metatarsal protrusion distance, magnitude of preoperative metatarsus adductus angle, magnitude of initial change in HVA (HVA_B – HVA_6wk), magnitude of initial change in first-second IM angle (IM_B – IM_6wk), magnitude of initial change in metatarsal protrusion distance (MPD_B – MPD_6wk), magnitude of overall change in HVA (if HVA ≥ 0, HVA_B – HVA_59wk), magnitude of overall change in first-second
**Table 1.** Change in FAOS Subscale Scores and Radiographic Parameters After Surgery (n = 80).a

<table>
<thead>
<tr>
<th>FAOS subscale</th>
<th>Preoperative (Average 59 wk)</th>
<th>Change</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain</td>
<td>71.5 ± 18.5</td>
<td>15.5 ± 22.4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Symptoms</td>
<td>82.5 ± 15.3</td>
<td>4.9 ± 17.5</td>
<td>.0150</td>
</tr>
<tr>
<td>Function, daily living</td>
<td>81.8 ± 19.3</td>
<td>11.8 ± 22.8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Function, sports and recreation</td>
<td>68.8 ± 24.3</td>
<td>20.1 ± 27.3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Quality of life</td>
<td>51.1 ± 18.4</td>
<td>26.4 ± 27.9</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Radiographic parameter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hallux valgus angle (degrees)</td>
<td>27.2 ± 9.5</td>
<td>−19.3 ± 8.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>First/second IM angle (degrees)</td>
<td>12.2 ± 3.8</td>
<td>−7.3 ± 3.8</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Tibial sesamoid position</td>
<td>4 (2-6)</td>
<td>−3 (0-5)</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Metatarsus adductus angle (degrees)</td>
<td>25.0 ± 5.8</td>
<td>−2.3 ± 2.5</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Metatarsal protrusion distance (mm)b</td>
<td>−3.1 ± 2.9</td>
<td>−5.4 ± 3.1</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: FAOS, Foot and Ankle Outcome Score; IM, intermetatarsal; NS, non-significant test result.

*All values are presented as mean ± SD, except for tibial sesamoid position, which is given as median (range). A paired t test was used for comparisons.

IM angle (if $HVA \geq 0,$ $IM_B - IM_{59wk}$), magnitude of overall change in metatarsal protrusion distance ($MPD_B - MPD_{59wk}$), and magnitude of residual HA deformity (if $HVA \leq 15,$ $HA_{59wk} - 15$). Dichotomous variables included preoperative sesamoid position = severe (5 or higher), suboptimal initial reduction of HVA (where optimal = 0 ≤ HA ≤ 15 at 6 weeks), suboptimal initial reduction of first-second IM angle (where optimal = 0 ≤ IM ≤ 10 at 6 weeks), suboptimal initial sesamoid reduction (4 or higher at 6 weeks), suboptimal final reduction of HVA (where optimal = 0 ≤ HA ≤ 15 at final follow-up), suboptimal final reduction of first-second IM angle (where optimal = 0 ≤ IM ≤ 10 at final follow-up), suboptimal final sesamoid position (4 or higher at final follow-up).

Multivariable models to predict FAOS scores were built using the following approach: Bivariate associations among the continuous radiographic and other predictor variables were initially examined using Pearson correlation coefficient ($r$) with each of the 5 FAOS subscales at final follow-up. Dichotomous predictors were examined in unadjusted linear regression models with each of the subscale scores at final follow-up. Any predictors with $P < .20$ in the bivariate/unadjusted analysis were then tested in multivariable models that included all potential predictor variables and adjusted for the respective baseline FAOS score. Patients with radiographic hallux varus ($n = 73$) or first-second IM angle were excluded when continuous forms of hallux valgus or first-second IM angle were being tested in models, but these patients were retained when dichotomous forms of these variables (ie, favorable/unfavorable postoperative radiographic correction) were being tested in the multivariable models. We excluded patients with negative HVA because while we would expect to find that smaller HVA angles would be associated with higher FAOS scores in the linear modeling, this would only be true to a point (up to 0 degrees). Malpositioned toes in varus with HVAs smaller than 0 would not be expected to continue to follow the linear trend. Kolmogorov-Smirnov goodness-of-fit test and normal probability plots of jackknife residuals were used to assess normality of the data. Plots of jackknife residuals against predicted values were also used to check the model assumptions of linearity, homoscedasticity, and independence. Two final models (1 excluding postoperative hallux varus patients [$n = 73$], and the other including hallux varus patients [$n = 80$]) were generated for each of the 5 FAOS subscales at final follow-up.

Descriptive statistics were generated for the study population. Pre- and postoperative FAOS scores and pre- and postoperative radiographic parameters were examined using paired $t$ test. $P$ values $<.05$ were considered statistically significant. All analyses were carried out using SAS version 9.4.

**Results**

Neither follow-up time nor surgeon was associated with any of the FAOS outcomes at final follow-up ($P > .05$). Statistically significant improvements for the cohort were seen postoperatively in all 5 FAOS domains and in all measured radiographic parameters (Table 1). Tests of the model assumptions of normality, linearity, homoscedasticity, and independence did not reveal any frank violations.

**Bivariate Analysis**

Results of the Pearson correlation analysis showed that no radiographic variable achieved anything more than a
weak correlation with any of the FAOS subscale scores at final follow-up. The variables that demonstrated the strongest linear correlation were first-second IM angle at final follow-up with sports and recreation subscale scores ($r = -0.328, P = .005$) and metatarsal protrusion distance at final follow-up with function/daily living scores ($r = 0.326, P = .005$). All other covariates demonstrated a negligible Pearson correlation coefficient ($r \leq |0.29|$) for each of the 5 subscale scores at final follow-up. There was no correlation found between change in hallux valgus angle (initial or overall change), change in first-second IM angle (initial or overall change), magnitude of preoperative hallux valgus angle or magnitude of preoperative first-second IM angle ($P > .05$ for all). Unadjusted linear regression models demonstrated several dichotomous variables with $P$ values $< .20$. These were tested in multivariable models, which is described below. Suboptimal sesamoid position at final follow-up was tested in the multivariable model for sports and recreation only, but otherwise sesamoid position was not found to be correlated with any of the other 4 subscales in the bivariate analysis ($P > .20$ for all).

**Multivariable Models**

**FAOS pain.** Two models were developed to predict pain at final follow-up. The first one used the continuous variables ($n = 73$) of first-second IM angle at final follow-up, metatarsal protrusion distance at final follow-up, and HVA at final follow-up (each with $P < .20$ in the bivariate analysis) while adjusting for preoperative FAOS pain subscale scores. The $R^2$ for this model was 0.109, and none of the variables were statistically significant in the final model. The second model ($n = 80$) used the same variables but tested the dichotomous handling of HVA and first-second IM angle (ie, suboptimal final reduction of HVA, and suboptimal final reduction of first-second IM angle) rather than the continuous handling. The $R^2$ for this model was 0.173, and the suboptimal final reduction of the first-second IM angle ($P = 0.034$) was the only statistically significant variable in the final model explaining 8% of the variance in FAOS pain at final follow-up (Table 2).

**FAOS Symptoms.** Two models were developed to predict FAOS Symptoms subscale scores postoperatively at 59 weeks. The first one ($n = 73$) tested the continuous variables of preoperative first-second IM angle, first-second IM angle postoperatively at final follow-up, metatarsal protrusion distance at final follow-up, and HVA at final follow-up (each with $P < .20$ in the bivariate analysis) while adjusting for preoperative Symptoms FAOS subscale scores. The $R^2$ for this model was 0.138, and the only significant variable was preoperative Symptoms subscale score (partial $R^2 = 0.071, P = .022$). The second model ($n = 80$) used the same
variables but tested suboptimal final reduction of HVA rather than the continuous handling of this variable. The $R^2$ for this model was 0.107, and again the preoperative Symptoms subscale score (partial $R^2=0.063, P=.024$) was the only statistically significant variable in the final model.

**FAOS function, sports and recreation.** Again 2 models were developed to predict sports and recreation scores at final follow-up. The first one used the continuous variables (n = 73) of age, preoperative first-second IM angle, first-second IM angle postoperatively at final follow-up, metatarsal protrusion distance at final follow-up, and hallux valgus angle at final follow-up (again all with $P < .20$ in the bivariate analysis), while adjusting for preoperative the FAOS sports and recreation subscale scores. The $R^2$ for this model was 0.233, and the first-second IM angle at final follow-up, hallux valgus angle at final follow-up, and age were each statistically significant in the final model (Table 3). The second model (n = 80) used the same variables but tested suboptimal final reduction of HVA rather than the continuous handling of this variable. The $R^2$ for this model was 0.131, and again metatarsal protrusion distance at final follow-up (partial $R^2=0.131, P = .001$) was the only statistically significant variable in the final model.

**FAOS quality of life.** Two models were developed to predict FAOS quality of life subscale scores at final follow-up. The first one (n = 73) tested the continuous variables of metatarsal protrusion distance at final follow-up, and the hallux valgus angle at final follow-up (both with $P < .20$ in the bivariate analysis), while adjusting for preoperative quality of life subscale scores. The $R^2$ for this model was 0.133, and the only significant variable was metatarsal protrusion distance at final follow-up (partial $R^2=0.059, P = .039$). The second model (n = 80) used the same variables but tested suboptimal final reduction of HVA rather than the continuous handling of this variable. The $R^2$ for this model was 0.138, and metatarsal protrusion distance at final follow-up (partial $R^2=0.068, P = .019$) was the only statistically significant variable in the final model.

**Discussion**

The study served to address one of the evolving “macro” trends in medicine, that of patient-centric, value-based
Our results indicate that there was nothing more than a minimal correlation between pre- and postoperative radiographic parameters and patient-reported outcomes following hallux valgus surgery. The metatarsal protrusion distance was statistically significant in 2 of the final models (quality of life and function/daily living), with shorter first metatarsals always being associated with lower FAOS scores. It was most closely associated with outcomes in the daily living subscale, but it still only explained 13% of the variance in this model. Similarly, postoperative first-second IM angle was retained in 2 of the study’s final predictor models (pain and function/sports and recreation), but again only explained anywhere from 8% to 10% of the variance in final follow-up FAOS scores in each of these respective domains. Furthermore, none of the study’s final predictive models achieved an $R^2 > 0.24$, and nearly all fell between 0.10 and 0.17. This suggests that radiographic variables alone do not do a very good job of explaining who is going to report favorable, or unfavorable, outcomes after hallux valgus surgery. This is surprising because so much emphasis is placed on radiographic angles and parameters by surgeons both preoperatively and postoperatively. What is more is that our group is not the first to report on this phenomenon. We are simply replicating something that was previously described (albeit using a separate outcome measure and different analytic approach) more than 10 years ago by Thordarson and colleagues.14,15

Perhaps most surprising was that postoperative HVA (what many surgeons believe to be the most essential element in hallux valgus repair) was not at all associated with patient-reported outcomes in our population in 4 of the 5 FAOS subscales, and it was only very weakly associated with quality of life (partial $R^2=0.06$). These findings suggest that the emphasis in prior studies on HVA and radiographic hallux valgus recurrence as the primary outcome may be misguided, especially if the overarching goal is to improve patient satisfaction with hallux valgus surgery. It also suggests that straight toes are not necessarily what is needed to get pain relief or to begin functioning again at a higher level postoperatively. Toe alignment does, however, appear to play a very modest role (explaining 6% of the variance) in quality of life and with patient confidence postoperatively.

We did not find an association between sesamoid position (preoperative or postoperative) and outcomes in this analysis. However, many of our patients achieved complete or near complete reduction of the sesamoid apparatus after distal osteotomy. In fact, 71 (89%) had a sesamoid position of 3 or less at final follow-up, leaving us with relatively small numbers to try and detect an association if there was one. The lack of variation in postoperative sesamoid position is a notable weakness in this work.

Further research is required to better understand patient-perceived outcomes in hallux valgus surgery. The proposed study evaluated patient experience with regard to a well-established validated scale, the FAOS scale. However, we recognize that FAOS is not necessarily the “gold standard” of hallux valgus surgery outcome measures. It is important to understand that as the larger trends in medicine turn to focus on value-based care, the systems used to evaluate the relationship between surgery and perceived value continue to evolve. Research using newer scales such as PROMIS and Patient Reported Outcome Measurement Information System Computerized Adaptive Test (PROMCAT) may be a more appropriate focus of future research as they have been shown to be superior and more widely applicable than traditional FAOS scoring.9 However, what our work has demonstrated is that accurate prediction as to which patients will report higher levels of improvement following hallux valgus surgery is a complex issue that goes far beyond radiographic appearance. Future work should focus on expanding the list of demographic predictor variables and examine other predictor variables that may have merit, such as patient biopsychosocial profiles, personality traits, and patient expectations, and even consider the role of pain exaggeration.

Limitations are inherent to all studies. Perhaps most importantly, our results may suffer some from selection bias, as we only included those patients who had completed FAOS surveys as part of their clinical care at our institution. It could be argued that patients completing surveys are perhaps more willing or eager to please their surgeon, so our population may not be entirely representative of all patients who underwent hallux valgus surgery. Also, the patient population analyzed in this work was fairly homogenous, with most patients experiencing fairly satisfactory reduction in their radiographic measurements postoperatively and only a minority of the group lying outside of our normally accepted radiographic ranges. This lack of variability among our independent variables may have made it difficult to detect certain correlations, even if they existed. We also restricted our analysis to exploring only linear associations among radiographic predictors and FAOS scores. It may well be that the “best fit” and most functional forms for our radiographic predictors were actually nonlinear (eg, quadratic, cubic) terms. However, the net effect of including nonlinear terms would have likely been very small as our scatter plots failed to suggest any obvious nonlinear associations among the variables examined in this analysis, and subsequent explorations into the study’s bivariate nonlinear associations using Spearman $\rho$ failed to show any meaningful departures from the Pearson $r$ values used throughout the article. Finally, FAOS is but one patient-reported measure commonly used to evaluate foot/ankle outcomes, and there are several new measures currently in use that may have been more appropriate.9

In conclusion, we found that there was minimal correlation between pre- and postoperative radiographic parameters and patient-reported outcomes following hallux valgus

Foot & Ankle International 00(0)
surgery. These findings further solidify previous findings on the subject by Thordarson and colleagues, and give further credence to the observation that many patients frequently improve with operative correction regardless of radiographic outcome and vice versa.\textsuperscript{14,15} Our findings also suggest that perhaps too much emphasis is currently being placed on radiographic angles and parameters by surgeons both preoperatively and postoperatively. Future work may consider what could add any additional insight into this subject, and further explore what aspects, beyond traditional radiographic measurements, are actually contributing to our best and worst patient-reported outcomes after hallux valgus surgery.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. ICMJE forms for all authors are available online.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References