

Accuracy and Reliability of Radiographic Estimation of Volar Lip Fragment Size in PIP Dorsal Fracture-Dislocations

HAND
2019, Vol. 14(6) 797–802
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1558944718777831
hand.sagepub.com

Daniel S. Donovan¹, Jeremy D. Podolnick¹, Wayne Reizner²,
O. Alton Barron³, Louis W. Catalano III³, and Steven Z. Glickel³

Abstract

Background: A cadaveric study was performed to evaluate the accuracy and reliability of radiographic estimation of the volar lip fragment size in proximal interphalangeal joint fracture-dislocations. **Methods:** Middle phalangeal base volar lip fractures of varying size and morphology were simulated in 18 digits. Radiographs and digital photographs of the middle phalangeal joint surface were obtained pre- and postinjury. Ten orthopedic surgeons of varying levels of training estimated the fracture size based on radiographs. The estimated joint involvement on radiograph was compared with the digitally measured joint involvement. **Results:** Radiographic estimation underestimated the volar lip fragment size by 9.02%. Estimations possessed high intraobserver (0.76-0.98) and interobserver (0.88-0.97) reliabilities. No differences were detected between levels of surgeon training. **Conclusions:** The significant underestimation of the volar lip fragment size demonstrates the lack of radiographic estimation accuracy and suggests that surgeons should be mindful of these results when making treatment plans.

Keywords: proximal interphalangeal joint, fracture-dislocation, volar lip, radiograph, diagnostic

Introduction

Fractures of the proximal interphalangeal joint (PIPJ) with or without dislocation are very common and can cause significant disability and morbidity. The decision of whether to treat a dorsal fracture-dislocation surgically depends largely on the stability of the joint after reduction. One of the metrics commonly used in this decision-making process is the percentage of articular surface which is fractured.³ The most commonly used modality for estimating this percentage of articular involvement is plain radiographs or fluoroscopy. However, the true size of the fracture fragment may be higher or lower than initially estimated on plain radiograph.

We performed a cadaveric study in which we created intra-articular volar shear fractures of the PIPJ of a variety of sizes and correlated these known fracture sizes to estimations of fracture size based on blinded readings of fluoroscopic images. The purpose of this study is to assess the accuracy of visual radiographic estimation of volar lip fracture size in the setting of fracture-dislocations of the PIPJ.

Materials and Methods

Cadaveric Specimen Preparation

The index, middle, and ring fingers of 6 cadaveric hands were used (18 specimens). Posteroanterior (PA) and lateral radiographs of the PIPJ of each digit were obtained prior to injury. The PIPJ of each specimen was then exposed through a volar approach. Using a sagittal saw with a cut thickness of 0.43 millimeters, a retrograde osteotomy was directed toward the base of the middle phalanx of each digit to mimic a volar lip fracture. To prevent removal of bone with the kerf of the blade, the osteotomy into the joint was completed with an osteotome and not the saw. Osteotomy sites

¹Mount Sinai St. Luke's-Roosevelt Hospital, New York, NY, USA

²Mount Sinai Hospital, New York, NY, USA

³The Roosevelt Hand to Shoulder Center, New York, NY, USA

Corresponding Author:

Daniel S. Donovan, Department of Orthopaedics, University of California, Irvine, 101 The City Drive, Pavilion 3, 2nd Floor, Orange, CA 92868, USA.

Email: danielsonovan86@gmail.com

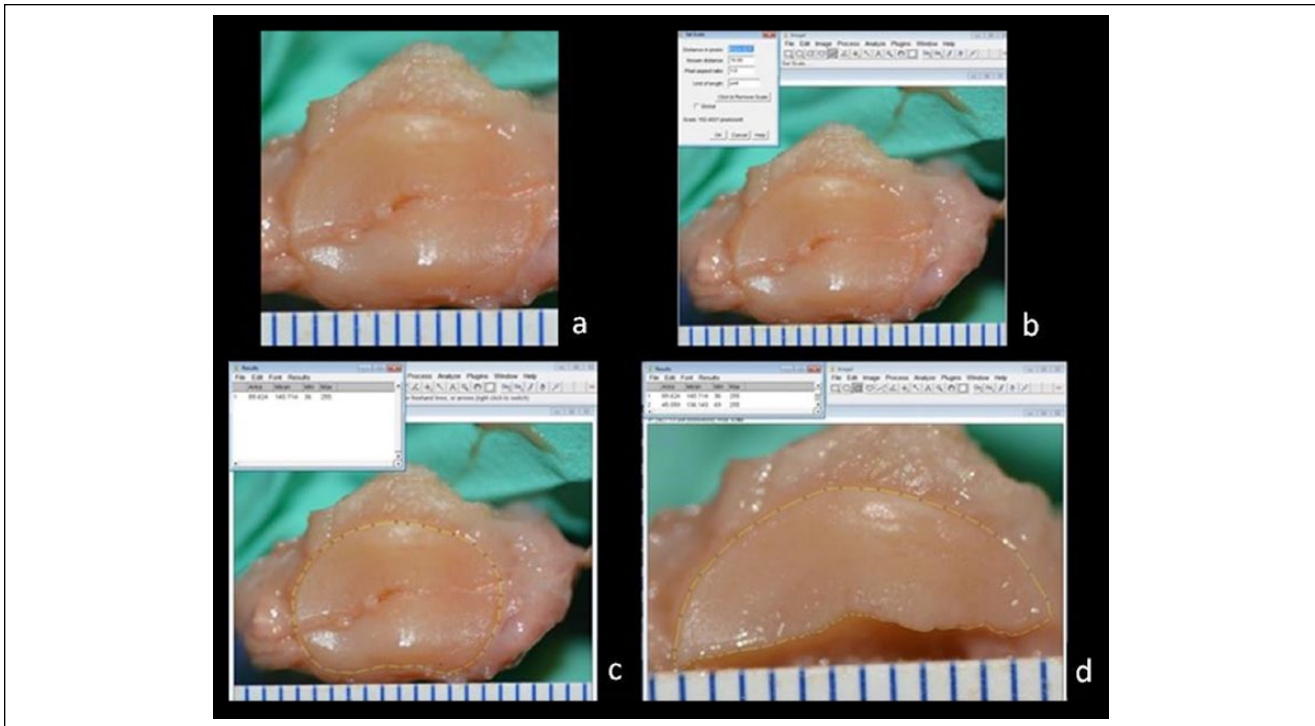


Figure 1. (a) The shotgunned joint with fracture joint surface reduced. The volar fragment is adjacent to the ruler. (b) The known value from the ruler at the joint surface is used to calibrate the software. (c and d) Measuring total area of the joint and the dorsal area (two of the measurements used to determine the total area fractured).

were placed at varying distances from the volar articular surface to create fracture fragments of varying size (range 10% to 90% of articular surface). Postinjury PA and lateral radiographs were then obtained.

Digital Measurement

A digital photograph of the articular surface of each middle phalanx base was performed using a Nikon D7100 digital camera. Each photograph was taken perpendicular (bird's-eye view) to the articular surface with a ruler held at the level of the articular surface to provide scale. Photographs were taken of the articular surface both before and after the osteotomy was created. Using ImageJ software (National Institutes of Health), the total surface area of the articular surface and the surface area of the volar lip fragment were calculated by 2 of the investigators. Each investigator performed these measurements twice. The software converted pixels in a demarcated area to millimeters squared through calibration to the photographed ruler with each photograph. The digital measurements were used to define the true percent of articular surface involvement by the osteotomy (Figures 1a-1d).

Radiographic Assessment

Ten orthopedic physicians served as assessors: 5 orthopedic residents (1 PGY1, 1 PGY2, 1 PGY3, 1 PGY4, and 1



Figure 2. Posteroanterior and lateral radiographs of a fractured digit blinded for review.

PGY5), 2 orthopedic hand fellows, and 3 attending orthopedic physicians with formal hand fellowship training. Four blinded PowerPoint (Microsoft, Redmond, Washington) presentations were prepared. Two presentations included only postinjury PA and lateral radiographs, and 2 presentations included both preinjury and postinjury radiographs (Figures 2 and 3). In practice, a radiograph of the

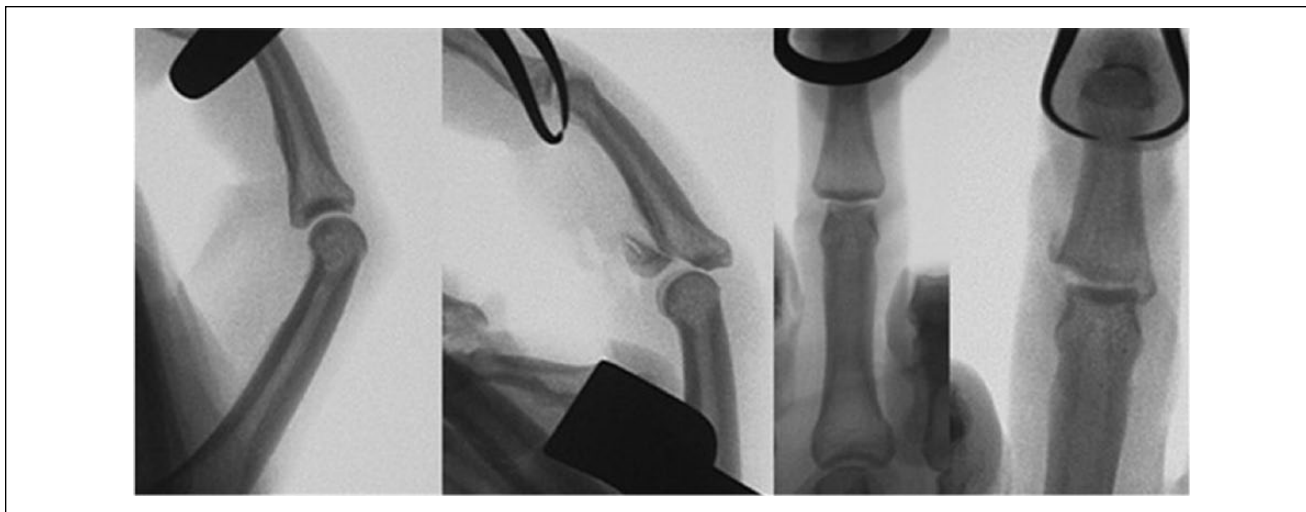


Figure 3. Posteroanterior and lateral radiographs of a fractured digit including prefracture image for comparison—both blinded for review.

contralateral uninjured digit may be obtained for comparison. To simulate this, a preinjury radiograph was obtained and used as a proxy for obtaining the contralateral uninjured digit. The order of the radiographs was randomized in each presentation. Slideshows were presented to the assessors at 3-week intervals. Each assessor estimated the percentage of the articular surface represented by the volar fragment. No time limit was set for this task, but all estimates for each round were done in the same session.

Statistical Analysis

Intraclass correlation coefficients (ICC) and 95% confidence intervals (CI) were calculated to assess intrarater reliability and interrater reliability for radiographic measurements among the 10 radiographic raters. Repeated-measures analysis models were used to evaluate the contribution of rater, training level, and presence of comparison preinjury radiographs to the overall variability in the radiographic measurements.

For each sample ($n = 18$), we calculated the mean and 95% confidence interval for radiographic measurements and digital measurements. We conducted paired t tests to determine whether the mean differed by measurement method, for each sample, and overall. Throughout all analyses, significance was assessed at the 5% level.

Results

Overall, the volar lip fragment size was routinely underestimated based on the measurements from radiographs. Table 1 and Figure 4 present the characteristics of the specimens as well as the mean measurements by radiographic and digital measurement. The “radiographic volar fragment size” repre-

sents the mean size (%) of each specimen as calculated using the 10 assessors’ estimates. The ranges of the estimates are included. The mean difference in measurement was -9% (95% CI -13.7 to -4.3), with a range from -35% to 7% . The radiographic measurements underestimated the fracture size on average by 9% ($P < .001$). When analyzed by sample, there were several significant differences between digital readings and radiographic readings (Table 1, Figure 3).

The ICC was calculated to determine intraobserver reliability. It ranged from 0.76 to 0.98, averaging 0.93, suggesting the agreement among a single rater’s 4 readings was high (Table 2, Figure 5).

The interclass correlation coefficient was calculated to determine interrater reliability. It was 0.93; 95% confidence limit was 0.88 to 0.97, demonstrating high levels of concordance between raters.

Raters were stratified by training level and their results compared. There was no difference detected ($P = .17$) between the groups.

Raters were provided only the fractured images (Rounds 1 and 3) or the fractured and prefractured images for comparison (Rounds 2 and 4). There was no difference between the samples measured with the prefractured image for comparison and those with only the fractured images ($P = .36$).

Discussion

Fracture-dislocations of the PIPJ pose both diagnostic and therapeutic challenges. Variable fracture morphology, delayed patient presentation, different treatment modalities, and the propensity for PIPJ stiffness after injury make achieving excellent results challenging.⁴ The classification scheme of Kiefhaber and Stern consider fractures less than 30% of the volar articular surface of the middle phalanx

Table I. Characteristics of the Specimens and Average Radiographic and Digital Measurements.

Sample	Digit	Age/sex of specimen	Radiographic volar fragment size (%)	Digitally measured volar fragment size (%)	Average difference in measurement (%)	P value for measurement—radiograph vs digital software	P value for difference between exposure to prefractured image vs only fractured image
1	2	61 M	23.28 [10-45]	57.85	-34.575	<.0001	.978
2	3	61 M	32.95 [5-50]	30.35	2.6	.3281	.4328
3	4	61 M	32.33 [20-41]	36.95	-4.625	.0571	.169
4	2	55 M	33.20 [25-40]	37.70	-4.5	.1225	.9275
5	3	55 M	54.70 [15-65]	61.25	-6.55	.0526	.1529
6	4	55 M	10.30 [0-20]	20.90	-10.6	.004	.4106
7	2	57 F	14.20 [5-25]	30.40	-16.2	.0002	.4599
8	3	57 F	45.85 [30-60]	63.25	-17.4	<.0001	.6671
9	4	57 F	45.35 [15-55]	58.55	-13.2	.0095	.5634
10	2	51 M	75.60 [70-80]	80.90	-5.3	.0261	.7927
11	3	51 M	39.15 [25-50]	53.50	-14.35	.0006	.6021
12	4	51 M	26.23 [10-40]	43.25	-17.025	.0018	.7659
13	2	44 M	26.50 [15-40]	38.40	-11.9	.0004	.5673
14	3	44 M	27.88 [15-40]	26.50	1.375	.7345	.8423
15	4	44 M	38.65 [15-50]	46.40	-7.75	.0002	.931
16	2	51 M	63.13 [25-85]	65.90	-2.775	.8085	.6434
17	3	51 M	89.73 [70-100]	82.35	7.375	.2062	.5831
18	4	51 M	47.85 [35-66]	54.85	-7	.0951	.0456
					Overall average difference in measurements	Overall average P value for measurement—radiograph vs digital software	
					-9.02	.0008	

Note. P values < .05 are indicated in bold.

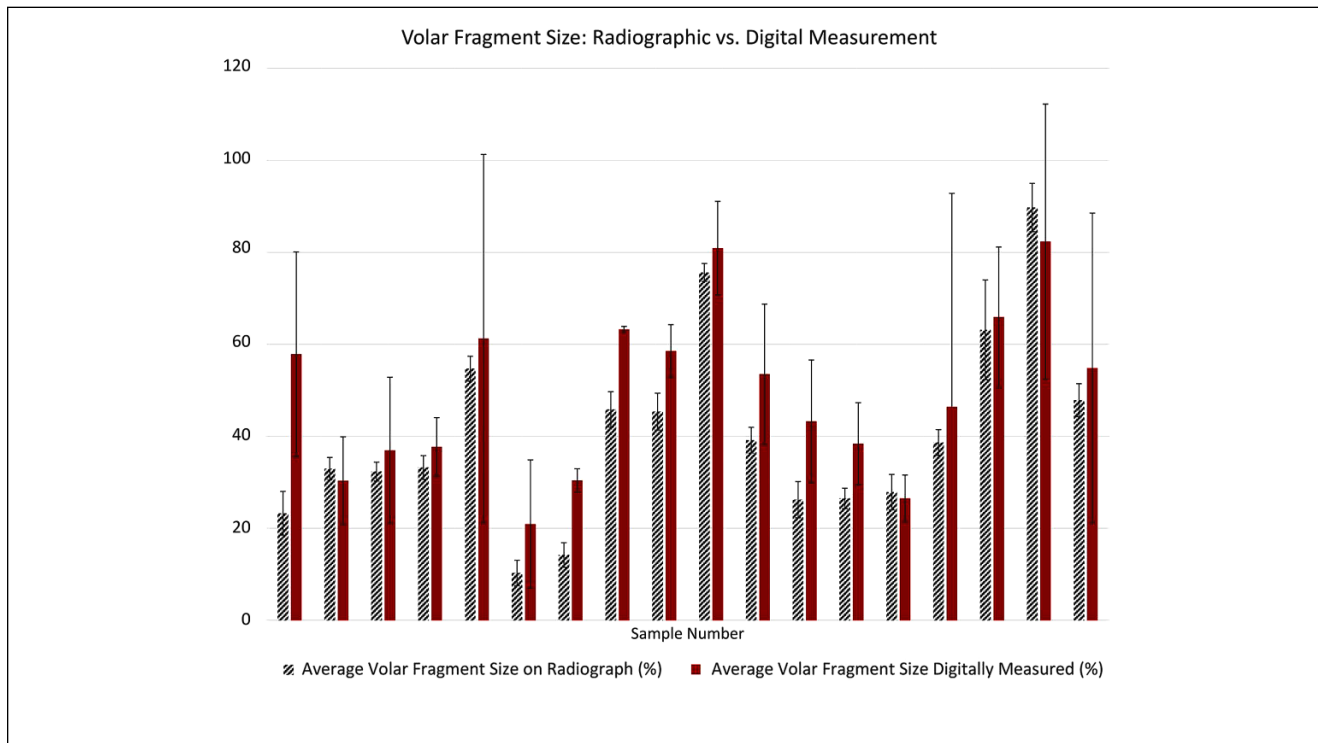


Figure 4. Average percentage of volar fragment size with error bars as 95% confidence intervals. Measurements made from radiographs to the left in blue, and digitally measured in red on the right.

Table 2. Intrarater Reliability: The Intraclass Correlation Coefficient (ICC) for the Agreement Among Each Single Rater’s Four Readings.

Rater	ICC	Lower	Upper
Attending 1	0.96	0.92	0.98
Attending 2	0.97	0.94	0.99
Attending 3	0.95	0.91	0.98
Fellow 1	0.96	0.92	0.98
Fellow 2	0.95	0.9	0.98
PGY 1	0.94	0.89	0.97
PGY 2	0.91	0.83	0.96
PGY 3	0.76	0.6	0.88
PGY 4	0.98	0.95	0.99
PGY 5	0.92	0.85	0.96

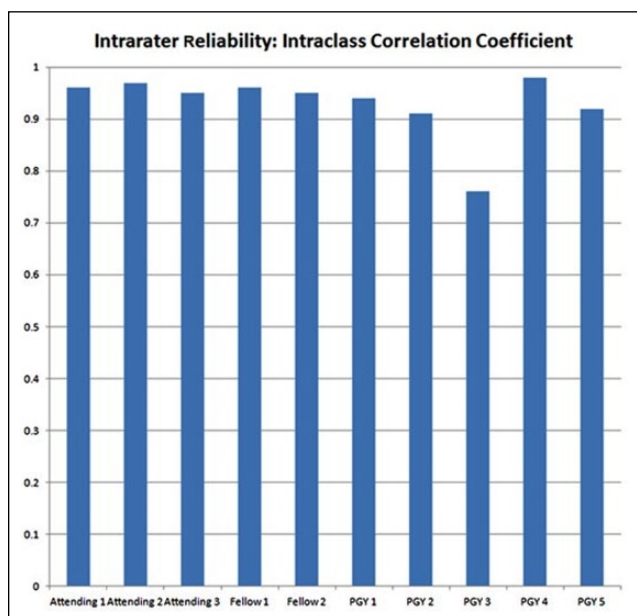


Figure 5. The intraclass correlation coefficient for the agreement among each single rater’s 4 readings.

stable, 30% to 50% as tenuous, and greater than 50% as unstable.⁴ This classification scheme can guide treatment algorithms. Generally, the severity of injury and loss of stability increase with amount of intra-articular volar fracture surface.² As the most widely used treatment algorithms center around the amount of articular involvement,¹ the ability to accurately measure the fracture size and estimate articular involvement with radiographs is paramount. Unfortunately, accurate estimates may be difficult to achieve, and clinicians have acknowledged that the severity of the fracture and degree of involvement of the middle phalanx may be greater than they appear in radiographs.⁷

The accuracy of using radiographs to measure volar lip fracture size in PIPJ fracture-dislocations had previously

not been evaluated. However, prior studies have examined the accuracy of visual estimation of fracture size on radiographs compared with digital measurement techniques among other orthopedic injuries. For instance, a radiographic study measuring distal radius fractures demonstrated poor reliability in all parameters measured, which similarly did not improve with level of training.^{5,6}

Our study aimed to evaluate the accuracy of orthopedic physicians in estimating volar lip fragment size based on radiographs of a fracture of the base of the middle phalanx simulating a PIP fracture-dislocation injury pattern. Our null hypothesis was that there would be no difference between the radiographically estimated fracture size and the digitally measured fracture size. This hypothesis was rejected. There was an average underestimation of fracture size by radiographs of 9%, with estimates ranging from -35% to 7% ($P < .001$, CI -13.7 to -4.3). Clinically, this underestimation could impact the physician’s treatment plan based on the available algorithms. For example, the treatment for a fracture involving 25% of the joint surface is dramatically different from one involving 60%. These findings should be used to caution providers to consider that a volar lip fracture is larger than what is estimated on radiograph.

We did not detect a difference in accuracy of measurement based on level of orthopedic training, though this may have been due to the relatively small number of participants. We are unable to state that higher levels of orthopedic training confer greater accuracy in judging the volar lip fragment of a PIP fracture on radiographs.

Interestingly, we did not find a benefit of measurements made with the inclusion of the preinjured film for comparison ($P = .36$). This finding suggests imaging the contralateral, uninjured finger for comparison would not yield superior accuracy.

The strengths of this article include our use of only the second, third, and fourth fingers to remove the confounding variable of different appearance of the thumb interphalangeal joint and small finger PIPJ, per Tyser et al.⁸ Radiographs were reviewed at 3-week intervals to prevent bias from prior measurements. The weaknesses of this study are chiefly those inherent in a cadaveric model. The injury mechanisms employed were not natural injuries to the PIPJ but rather iatrogenic. Another weakness is that only 18 digits were used and only 10 observers participated.

We used the percentage of articular surface involved in the fracture as measured digitally as the standard for comparison in this study. This method has both strengths and weaknesses. It allowed us to fully visualize the articular fracture pattern which radiographs fail to do. One possible reason for the underestimation of fracture size is that some fractures entered the joint at varying obliquities which is difficult to assess with radiographs. However, by using the digital pictures of only the joint surface as a standard, we were unable to see the amount of bone that was attached to

the articular piece. On a lateral radiograph, this is more identifiable. The amount of bone involvement of the fracture pattern may play a role when deciding on treatment course. While not investigated in this study, oblique radiographs or advanced imaging may provide insight into the spatial relationship of the fragment and offer clues as to the integrity of the collateral ligaments.

Based on these results, we recommend that the practitioner proceed with caution in their treatment algorithm based on radiographs alone, reaffirming the value of a sound clinical exam. This confirms previous suspicions that the magnitude of a fracture of the PIP joint is often greater than estimated on injury radiographs.

Ethical Approval

This study was approved by our institutional review board.

Statement of Human and Animal Rights

This article describes a cadaveric study.

Statement of Informed Consent

This article describes a cadaveric study.

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.

Funding

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

References

1. Glickel SZ, Barron OA. Proximal interphalangeal joint fracture dislocations. *Hand Clin.* 2000;16:333-344.
2. Khouri JS, Bloom JM, Hammert WC. Current trends in the management of proximal interphalangeal joint injuries of the hand. *Plast Reconstr Surg.* 2013;132:1192-1204.
3. Kiefhaber TR. Phalangeal dislocations/periarticular trauma. In: Premier CA, ed. *Surgery of the Hand and Upper Extremity.* New York, NY: McGraw-Hill; 1996:939-972.
4. Kiefhaber TR, Stern PS. Fracture dislocations of the proximal interphalangeal joint. *J Hand Surg Am.* 1998;23:368-380.
5. O'Malley MP, Rodner C, Ritting A, et al. Radiographic interpretation of distal radius fractures: visual estimations versus digital measuring techniques. *Hand.* 2014;9:488-493.
6. Robertson GAJ, Robertson BRM, Thomas B, et al. Assessing angulation on digital images of radiographs of fractures of the distal radius: visual estimation versus computer software measurement. *J Hand Surg Eur Vol.* 2011;36:230-235.
7. Ting BL, Blazar PE. Volar plate arthroplasty. In: Wiesel SW, ed. *Operative Techniques in Orthopaedic Surgery.* Philadelphia, PA: Lippincott & Williams; 2011:2431-2435.
8. Tyser AR, Tsai MA, Parks BG, et al. Biomechanical characteristics of hemi-hamate reconstruction versus volar plate arthroplasty in the treatment of dorsal fracture dislocations of the proximal interphalangeal joint. *J Hand Surg Am.* 2015;40:329-332.