Reliability and Accuracy of Templating Humeral and Ulnar Components for Total Elbow Arthroplasty

Nick D. Pappas, MD, Jeffry T. Watson, MD, John M. Erickson, MD, Keith D. Baldwin, MD, MSPT, MPH, and Donald H. Lee, MD

Abstract

We conducted a study to examine intraobserver reliability, interobserver reliability, and accuracy of preoperative templating in approximating humeral and ulnar component sizes in total elbow arthroplasty (TEA).

Twenty-two patients underwent cemented TEA with 1 of 2 commonly used implants. Four independent reviewers performed templating in 2 separate sessions spaced a minimum of 2 weeks apart. Reviewers were blinded to patient information and used appropriately magnified templates provided by the implant manufacturer. Preoperative and postoperative films were assessed for humeral and ulnar stem width and length.

For both implants combined, there was substantial ($\kappa > 0.7$) intraobserver reliability for humeral width, humeral length, and ulnar length. Interobserver reliability was fair for humeral width ($\kappa = 0.28$), substantial for humeral length ($\kappa = 0.64$), and moderate for both ulnar width ($\kappa = 0.44$) and ulnar length ($\kappa = 0.49$). Preoperative templating accurately predicted exact stem size 72.7% of the time and within 1 size variation 96.9% of the time. Attending surgeons were slightly more accurate than fellows (75.5% vs 71.5%) in predicting stem sizes.

Preoperative templating is moderately reliable and largely accurate in planning TEA.

Total elbow arthroplasty (TEA) is used to treat severe degeneration of the ulnohumeral and radiocapitellar joints. It provides reliable pain relief and restores elbow motion in the vast majority of patients.[1] However, because there are significantly fewer cases of advanced arthritis of the elbow than advanced arthritis of other major joints (eg, hip, shoulder, knee), the overall incidence of elbow arthroplasty is relatively low.[2] Even orthopedic surgeons with busy upper extremity practices may perform only a handful of TEAs per year.

Before any surgery is performed, it is important to engage in planning to help ensure a successful outcome. In joint arthroplasty, preoperative templating is often beneficial. Templating can provide additional information that may be useful to the surgeon, such as the size of the implant that may be needed and where bone cuts should be made. Templating has been efficacious in other arthroplasties—namely, hip and shoulder.[3] To our knowledge, however, no study has examined the effectiveness of preoperative templating in TEA. Therefore, we conducted a study to measure the reliability and accuracy of preoperative templating in TEA and to determine its usefulness for approximating humeral and ulnar stem sizes.

Materials and Methods

Institutional review board approval was obtained for this study. Twenty-two patients underwent cemented TEA. Of these patients, 13 received the Discovery Elbow System (Biomet, Warsaw, Indiana); the other 9 received the Coonrad/Morrey Total Elbow (Zimmer, Warsaw, Indiana). Four reviewers (2 experienced hand and upper extremity attending surgeons, and 2 hand and upper extremity fellows) analyzed the patients' preoperative and postoperative standard anteroposterior and lateral radiographs at 2 separate review sessions a minimum of 2 weeks apart. At each session, the reviewers were blinded to patient information and used appropriately magnified templates provided by the implant manufacturers. Preoperative radiographs were assessed for humeral and ulnar stem width and length, and postoperative radiographs for appropriate length and width of the implanted stem.

We used free-marginal κ analysis to calculate the reliability of intraobserver and interobserver agreement for preoperative templating.[4-6] Kappa values were interpreted according to a commonly used scale (**Table I**). The accuracy of templating was calculated as a percentage agreement between template size and implanted size for each patient. Postoperative analysis of

Authors' Disclosure Statement: Dr. Lee is a consultant to Biomet; has received payment for lectures, royalties, and development from Biomet and royalties from Elsevier. Dr. Pappas, Dr. Watson, Dr. Erickson, and Dr. Baldwin report no actual or potential conflict of interest in relation to this article.

Table I. Interpretation of κ Values

κ	Agreement
<0.0	Less than chance
0.0–0.20	Slight
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Substantial
0.81–0.99	Almost perfect

Table II. Intraobserver Agreement for Preoperative Templating

		κ	Agreement	% Agreement
Humeral	Width	0.659	Substantial	77.3
	Length	0.795	Substantial	86.4
Ulnar	Width	0.523	Moderate	68.2
	Length	0.761	Substantial	84.1

Table III. Interobserver Agreement for Preoperative Templating

		κ	Agreement	% Agreement
Humeral	Width	0.284	Fair	52.3
	Length	0.636	Substantial	75.8
Ulnar	Width	0.443	Moderate	62.3
	Length	0.489	Moderate	65.9

Table IV. Accuracy of Templating for Total Elbow Arthroplasty

		Accurate Prediction, %	1 Size Variation, %	>1 Size Variation, %
Humeral	Width	68.2	96.6	13.4
	Length	81.8	98.9	1.1
Ulnar	Width	64.7	97.7	2.3
	Length	76.1	94.3	5.7
Average		**72.7**	**96.9**	**5.6**

the subjective grading of implant fit was based on percentage agreement between raters.

Results

Among the 13 patients who received the Discovery implant, intraobserver agreement was substantial for humeral width (κ = 0.65), almost perfect for ulnar width (κ = 0.89), and almost perfect for humeral and ulnar length (κ > 0.80). Interobserver agreement was fair for humeral width (κ = 0.37), moderate for ulnar width (κ = 0.48), and substantial for humeral length (κ = 0.65) and ulnar length (κ = 0.64).

In the 9 patients who received the Coonrad/Morrey implant, intraobserver agreement was substantial for humeral width (κ = 0.75) and length (κ = 0.67), and moderate for ulnar width (κ = 0.54) and length (κ = 0.58). Interobserver agreement was slight for humeral width (κ = 0.17) and ulnar length (κ = 0.19), substantial for humeral length (κ = 0.61), and fair for ulnar width (κ = 0.40).

For the 2 implant types combined, intraobserver reliability was substantial for humeral width and length and ulnar length (κ > 0.7), and moderate for ulnar width (κ = 0.52) (**Table II**). Interobserver reliability was fair for humeral width (κ = 0.28), substantial for humeral length (κ = 0.64), and moderate for ulnar width (κ = 0.44) and length (κ = 0.49) (**Table III**).

Interobserver agreement regarding appropriate stem size on postoperative radiographs was 83.7%. Agreement was almost perfect for humeral length and ulnar width and length (κ > 0.80), and moderate for humeral width (κ = 0.47). Overall, stem size was deemed appropriate 89.2% of the time, with the humeral stem deemed too narrow 25% of the time and too long 2.3% of the time; the ulnar stem was deemed too narrow 6.8% of the time and too short 6.8% of the time.

Among all reviewers, preoperative templating accurately predicted exact stem size 72.7% of the time and within 1 stem size 96.9% of the time (**Table IV**). Reviewers were slightly more accurate in predicting exact Discovery stem size (75.7%) than exact Coonrad/Morrey stem size (68.0%). They were also slightly more accurate in predicting exact humeral stem size (75.0%) than exact ulnar stem size (70.4%). Regarding level of experience, attending surgeons were slightly more accurate (75.5%) than fellows (71.5%) in predicting exact stem size.

Discussion

Selecting an appropriately sized implant is key in successful joint arthroplasty. Preoperative templating can assist in this endeavor. Studies have shown templating to be efficacious in both hip and shoulder arthroplasty.[3,7]

The study reported here is the first to examine the reliability and accuracy of templating in TEA. On average, our results showed substantial intraobserver reliability for humeral width and length and ulnar length (κ > 0.7), and moderate reliability for ulnar width (κ = 0.52) (**Table II**). When the implants were analyzed in combination, however, interobserver agreement was only moderate (Discovery, κ = 0.53) to fair (Coonrad/Morrey, κ = 0.34). Although we do not know the exact reason for the slight difference between the implants, we speculate that since Coonrad/Morrey has an additional length option for its humeral and ul-

nar stems compared to Discovery, reviewers had more length options to choose from, and thus a greater chance of picking a stem different from the one implanted. Nevertheless, the relatively low overall interobserver agreement for both implants is still an issue. Why is agreement only moderate at best for predicting stem sizes?

First, there are no real guidelines regarding the lengths required for ulnar and humeral stems in TEA. For example, there are no data regarding when to use a 4-inch-long humeral stem and when to use a 6-inch-long humeral stem in primary TEA. Although a recent article indicated that longer humeral stems may cause elbow joint alignment issues, it did not elucidate the ideal stem lengths in different scenarios.[8] Generally, surgeons choose longer stems in trauma cases and shorter stems for arthritis. However, the literature does not provide any evidence supporting this practice in the elbow, though common sense suggests such an approach.

Whereas interobserver agreement was moderate for predicting stem sizes, it was excellent for judging the appropriateness of the implanted stems on postoperative radiographs. As a group, the reviewers subjectively graded the postoperative radiographs as having appropriately sized stems almost 90% of the time, which is a high percentage.

In this study, the overall accuracy of preoperative templating in predicting exact stem size was good (72.7%), and the accuracy was excellent in predicting within 1 stem size (96.9%). The general goal of templating stem sizes is to give the surgeon an approximate idea of the size of the implants that he or she will need during surgery. If a surgeon can predict within 1 stem size what will be needed almost 97% of the time, then templating is clearly worth the effort.

Our reviewers were slightly more accurate in predicting Discovery stem sizes (75.7%) than Coonrad/Morrey stem sizes (68.0%). Again, we hypothesize that the difference between these implants derives from the extra length option for the Coonrad/Morrey stems. Regarding level of templating experience, reviewers with more experience had a slightly higher accuracy (75.5%) than reviewers with less experience (71.5%), suggesting a learning curve for TEA templating.

Our study had several strengths and weaknesses. To our knowledge, this is the first study to analyze the reliability and accuracy of TEA templating. In addition, it used 4 separate reviewers, each of whom had been blinded to patient information and had reviewed the radiographs on 2 different occasions. We believe that these measures helped reduce bias in our data.

A potential weakness of this study is radiographic magnification. All films were printed radiographs, reportedly taken at the standard 40-inch source-to-image distance. Depending on patient size and amount of soft tissue around the arm, magnification is typically about 3% or higher. Our templates take into account approximately 3% to 5% of magnification. Magnifying the films more than 5% would have resulted in overestimating the required sizes of our components and thus would have skewed our data. One might argue that this issue could be solved with digital templating. Digital templating for TEA exists. However, it is costly and is not found in most institutions, including ours. The most practical templating method for most hand surgeons—certainly those in smaller community settings—still involves printed films and templating sheets, which is why we conducted our study the way we did.

Another potential weakness is our assumption that the size of the actual implant is the correct size for our assessment of accuracy. If the actual implant were undersized or oversized, it would limit our ability to truly measure accuracy. To control for this scenario, we had reviewers assess postoperative films for appropriate stem sizes. Given that the reviewers deemed stem sizes appropriate about 90% of the time, we believe that this potential source of error was minimized.

Conclusion

Preoperative planning is key in maximizing outcomes in joint arthroplasty. Specifically, templating for implant sizes before surgery can provide the surgeon with important information that can aid in intraoperative decision making. Our study showed substantial intraobserver agreement, fair-to-moderate interobserver agreement, and good accuracy for predicting exact stem sizes in TEA. Compared with less-experienced reviewers, experienced reviewers were slightly more accurate in predicting stem sizes. Overall, we believe that our results show that preoperative templating can be of significant value in performing TEA.

Dr. Pappas is Hand Surgeon at Greenville Health System/Steadman Hawkins Clinic of the Carolinas, Greenville, South Carolina. Dr. Watson is Assistant Professor of Orthopaedic Surgery, Vanderbilt Hand and Upper Extremity Center, Vanderbilt Orthopaedic Institute, Nashville, Tennessee. Dr. Erickson is Orthopaedic Surgeon, Raleigh Hand Center, North Carolina. Dr. Baldwin is Orthopaedic Surgeon, Children's Hospital of Philadelphia, Pennsylvania. Dr. Lee is Professor of Orthopaedic Surgery, Vanderbilt Hand and Upper Extremity Center, Vanderbilt Orthopaedic Institute, Nashville, Tennessee.

Acknowledgements: The authors thank Julie Daniels and Anita Pai for their assistance in data collection.

Address correspondence to: Donald H. Lee, MD, Suite 3200, Medical Center, South Tower, Vanderbilt Orthopaedic Institute, Nashville, TN 37232-8828 (tel, 615-322-4506; fax, 615-936-8116; e-mail, donald.h.lee@vanderbilt.edu).

Am J Orthop. 2013;42(7):321-323.

References

1. Kamineni S, Morrey BF. Distal humeral fractures treated with noncustom total elbow replacement. *J Bone Joint Surg Am.* 2004;86(5):940-947.
2. Stanley D. Prevalence and etiology of symptomatic elbow osteoarthritis. *J Shoulder Elbow Surg.* 1994;3(6):386-389.
3. Buzzell JE, Lutton DM, Shyr Y, Neviaser RJ, Lee DH. Reliability and accuracy of templating the proximal humeral component for shoulder arthroplasty. *J Shoulder Elbow Surg.* 2009;18(5):728-733.
4. Brennan RL, Prediger DJ. Coefficient kappa: some uses, misuses, and alternatives. *Educ Psychol Meas.* 1981;41(3):687-699.
5. Randolph JJ. Free-marginal multirater kappa (multirater κ_{free}): an alternative to Fleiss' fixed-marginal multirater kappa. Paper presented at: Joensuu University Learning and Instruction Symposium; October 2005; Vol Joensuu, Finland.
6. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
7. Unnanuntana A, Wagner D, Goodman SB. The accuracy of preoperative templating in cementless total hip arthroplasty. *J Arthroplasty.* 2009;24(2):180-186.
8. McDonald CP, Peters TM, Johnson JA, King GJ. Stem abutment affects alignment of the humeral component in computer-assisted elbow arthroplasty. *J Shoulder Elbow Surg.* 2011;20(6):891-898.