Original Clinical Article

JOURNAL OF CHILDREN'S ORTHOPAEDICS

Is the modified Gartland classification system important in deciding the need for operative management of supracondylar humerus fractures?

Tammie L. Teo[1]
Emily K. Schaeffer[1]
Eva Habib[1]
Ron El-Hawary[2]
Patricia Larouche[3]
Benjamin Shore[4]
Alexander Aarvold[5]
Sasha Carsen[6]
Christopher Reilly[1]
Kishore Mulpuri[1]

Abstract

Purpose: This study examined levels of agreement between paediatric orthopaedic surgeons in the need for operative management of extension-type supracondylar humerus fractures.

Methods: This was the second phase of a two-part study. De-identified baseline anteroposterior and lateral elbow radiographs from 60 paediatric patients with extension-type supracondylar humerus fractures were compiled. After classifying each fracture according to Gartland classification guidelines, radiographs were randomized, and surgeons indicated whether they would use operative or non-operative management to treat each fracture. Kappa statistics using pairwise comparisons were calculated to determine agreement levels.

Results: In total, 11 international surgeons participated, and 10/11 completed both survey rounds. The overall weighted interobserver agreement was moderate (0.530, 95%CI [0.215,0.854]) while overall weighted intraobserver agreement was substantial (0.740, 95%CI [0.513,0.963]). The largest variability in preferred treatment methods between surgeons was observed for type IIA fractures, with 6/11 preferring non-operative and 5/11 preferring operative management. The largest individual surgeon variability was observed for type IIA fractures, with 8/11 showing variability (defined by not having made the same decision for at least 90% of the cases) in choosing whether to operate.

Conclusions: Our findings suggest moderate interobserver, and substantial intraobserver agreement in treatment decision making. The largest disagreements between surgeons were observed for type IIA and IIB fractures and treatment decisions did not follow expected trends based on surgeons' preferred treatment methods for each fracture type. This suggests differences in treatment approaches between surgeons in the management of type IIA fractures and highlights the role of other variables that underlie differences between surgeons' treatment preferences.

Level of evidence: III

Cite this article: Teo TL, Schaeffer EK, Habib E, El-Hawary R, Larouche P, Shore B, Aarvold A, Carsen S, Reilly C, Mulpuri K. Is the modified Gartland classification system important in deciding the need for operative management of supracondylar humerus fractures? *J Child Orthop* 2020;14: 502-507. DOI: 10.1302/1863-2548.14.200093

Keywords: supracondylar; humerus; trauma; surgical decision making; Gartland classification system

[1] Department of Orthopaedic Surgery, British Columbia Children's Hospital, Vancouver, BC, Canada
[2] Division of Orthopaedics, Dalhousie University, Halifax, NS, Canada
[3] Département d'orthopédie, CHU de Québec-Université Laval, Québec, QC, Canada
[4] Department of Orthopaedic Surgery, Boston Children's Hospital and Harvard Medical School, Boston, USA
[5] Department of Paediatric Orthopaedics, Southampton Children's Hospital, University Hospital Southampton, Southampton, Hampshire, UK
[6] Division of Orthopedic Surgery, University of Ottawa, Ottawa, ON, Canada

Correspondence should be sent to Kishore Mulpuri, Department of Orthopaedic Surgery, BC Children's Hospital,1D.66-4480 Oak Street Vancouver, BC V6H 3V4, Canada.
E-mail: kmulpuri@cw.bc.ca

Introduction

The Gartland extension-type supracondylar humerus fracture is the most common fracture of the elbow in the paediatric population.[1] In general, the Wilkins[2] modification of the Gartland classification system is a useful guideline by which orthopaedic surgeons determine the management for extension-type supracondylar fractures. While type I fractures are usually treated non-operatively using tape, splint or cast immobilization, type IIB and type III

fractures are usually treated operatively using closed reduction and percutaneous pinning (CRPP) or open reduction and pinning.[3] However, there still exists a lack of consensus as to the appropriate management for type IIA fractures. While some surgeons use operative management as their default treatment of choice, other surgeons just as strongly believe that immobilization is sufficient. Unfortunately, the current evidence is not strong enough to support the collective adoption of one management strategy over the other,[4] despite recent publications suggesting specific treatment approaches based on Gartland fracture type.[5-7]

We previously sought to investigate whether differences between surgeons in treatment preferences for type II fractures are due to surgeons classifying the same types of fractures differently.[8] Substantial overall levels of agreement were found between surgeon respondents in the classification of fractures using the modified Gartland criteria, suggesting that variability in classification may not be a significant contributor to the current treatment controversy for type IIA fractures. If the controversy over treatment cannot be explained by variability in classification, then there is a need to identify where the differences in practice patterns lie and the factors underlying these differences, with particular regard to the role of the modified Gartland classification system in the treatment decision-making process. Once a better understanding of these differences is achieved, outcomes between operative and non-operative management can then be more appropriately compared, with the ultimate goal of standardizing patient care for these fractures.

The primary objective of this study was to examine levels of agreement between surgeons from around the world in the decision to operate on extension-type supracondylar humerus fractures. The secondary objective was to explore how the modified Gartland classification system impacts individual surgeon decision making.

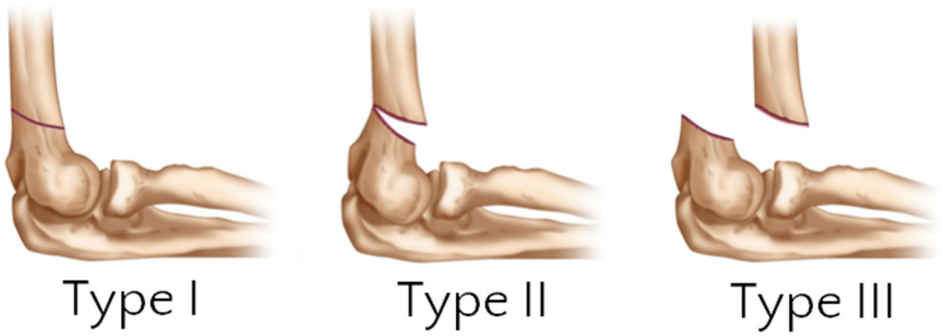
Materials and methods

This is the second phase of the study previously described.[8] The first phase investigated levels of agreement between surgeon respondents in fracture type classification using the modified Gartland classification system. After Institutional Research Ethics Board approval was obtained, we retrospectively reviewed the radiographs and charts of all patients from two years to 12 years of age who had been treated for an extension-type supracondylar humerus fracture at our institution from 1st January 2005 to 31st December 2016. Each patient had to have adequate pre-reduction radiographs available, defined by a true anteroposterior (AP) view with orthogonal visualization of the distal humerus and clear delineation of the hourglass sign and capitellum on lateral radiographs.

Following radiographic and chart review, we selected 60 patients for inclusion, all of whom had been diagnosed by one of seven surgeons at our institution with a Gartland type I (n = 10), type II (n = 25) or type III (n = 25) fracture. Baseline AP and lateral plain elbow radiographs for each patient were de-identified and compiled into surveys administered using Research Electronic Data Capture (REDCap) software (Vanderbilt University, Nashville, Tennessee, USA), which we used for both phases of the study.

Invitations to participate in the survey were sent out to fellowship-trained paediatric orthopaedic surgeons practising in tertiary care hospitals around the world. These surgeons were known to the study's senior authors as previous colleagues, mentees or research collaborators and were included in the study if they consented to participate. All invited surgeons practised in treatment centres that had sufficient radiological resources to perform closed reduction with percutaneous fixation of supracondylar fractures, and none were from countries with limitations in radiological resources that would incline them to use non-fixation methods over fixation in cases where they thought fixation would be warranted.

This study was conducted in two phases. For the first phase,[8] surgeons were provided with a brief pictorial and table summary of the Wilkins-modified Gartland classification system (Fig. 1), along with the compiled radiographs, and asked to classify the radiographs according to the classification system. After the two surveys in the first phase ('classification reliability survey') of the study were concluded, invitations to participate in the second phase ('treatment variability survey' – the present study) were sent out to all paediatric orthopaedic surgeons who had already consented to participate. Similar to the classification reliability survey, the treatment variability survey was conducted in two rounds. For the first round, surgeons were asked about their most preferred treatment method for each of the modified Gartland supracondylar fracture classifications. They were then provided with the same radiographs from the classification reliability survey and asked to indicate whether they would elect to treat each fracture non-operatively or operatively. If they had chosen non-operative management, they were asked to choose between splinting, taping or casting as immobilization options. If they had chosen operative management, they were asked to choose between closed reduction under general anaesthesia in the operating room, CRPP, or open reduction and internal fixation. Surgeons were blinded to the original Gartland classifications they had chosen for each radiograph in the classification reliability survey. For each set of radiographs provided, surgeons were also asked to indicate other factors that would influence their treatment decisions; options included degree of displacement, mechanism of injury, swelling and soft tissue status, neurovascular compromise and history of previous injury



Type I	Non-displaced
Type II A	Intact posterior cortex, hinged in extension. No rotation or translation.
Type II B	Intact posterior cortex, hinged in extension, with some degree of rotational displacement or translation.
Type III	Complete displacement.

Fig. 1 Summary of Wilkins-modified Gartland classification system[2] provided to survey respondents. Image sourced from Alton et al.[1]

to the same area. Following a minimum of two weeks, the same radiographs were shuffled and provided to each surgeon for another round of treatment decision making.

We calculated Kappa (κ) statistics using pairwise comparisons and averaged the κ values to determine overall interobserver and intraobserver levels of agreement. The Landis and Koch[9] guidelines were used to interpret the κ values as follows: values less than 0.00 indicate poor agreement, 0.00–0.20 indicates slight agreement, 0.21–0.40 indicates fair agreement, 0.41–0.60 indicates moderate agreement, 0.61–0.80 indicates substantial agreement, and 0.81–1.00 indicates excellent or almost perfect agreement.

In order to determine the proportion of cases surgeons had decided to operate on for the respective categories in the modified Gartland criteria, we matched each surgeon's treatment decisions from both rounds of the treatment variability survey with the classifications they had assigned the fracture in the classification reliability survey. The grouping of operative decisions by fracture classification for each surgeon was thus reliant on the classifications that they had previously assigned to the fractures. Since there were two rounds of the classification reliability survey and two rounds of the treatment variability survey, data for four matches were compiled and the total number of operative decisions categorized by fracture type was summed. For these groupings, surgeons who made the same decision for more than 90% of the radiographs were defined as being consistent, while surgeons who made the same decision in less than 90% of their treatment decisions in each fracture type category were defined as showing variability in their treatment decision-making

Results

Participant demographics

Of the 21 surgeons who had participated in the classification reliability survey, 11 surgeons (six from Canada, three from the United States, one from the United Kingdom, one from India), representing ten tertiary care hospitals, consented to participate in the treatment variability survey. Across all respondents, the mean length of practice was 8.4 years (range 1 years to 20 years), with each treating between 20 to 70 supracondylar humerus fractures annually. Following the two-week period, ten of the original survey respondents completed the second round of surveys. One of the original respondents could not respond to the survey in time due to work commitments and thus had to be excluded from the second round.

Interobserver level of agreement

The overall weighted interobserver agreement for operative decision making was moderate (Table 1).

Table 1 Interobserver and intraobserver κ values

No. of patient cases	Combined interobserver κ [95% CI] (n =11)	Combined intraobserver κ [95% CI] (n = 10)
60	0.530 [0.215, 0.854]	0.740 [0.513, 0.963]

JOURNAL OF CHILDREN'S ORTHOPAEDICS

Intraobserver level of agreement

The overall weighted intraobserver agreement for treatment decision making was substantial (Table 1). Individual intraobserver agreements for each of the ten respondents who completed both surveys ranged from 0.474 (moderate) to 0.879 (excellent or almost perfect).

Table 2 details each surgeon's most preferred treatment method for the various modified Gartland fracture classifications. No variability between surgeon preferences was observed for type I and type III fractures, with all respondents preferring non-operative and operative management respectively. There was also close to no variability observed for type IIB fractures, with only 1/11 surgeon preferring non-operative management and the remaining 10/11 preferring operative management. The largest variability in preferences was observed for type IIA fractures, with 6/11 preferring non-operative management and 5/11 preferring operative.

The overall observed trends in operative decisions when separated by surgeon-rated fracture classification (Table 3) generally reflect the practice patterns predicted by each surgeon's indicated preferred treatment method (Table 2), with the decision to operate making up 8.1% of treatment decisions for surgeon-rated type I fractures (compared to the predicted 0%), 50.8% for surgeon-rated type IIA fractures (compared to 45.5%), 82.9% for surgeon-rated type IIB fractures (compared to 90.9%) and 99.8% for surgeon-rated type III fractures (compared to 100%). However, individual decisions to operate did not follow the predicted patterns as closely. The largest amount of variability was observed for the type IIA fracture category, with 8/11 surgeons (D, E, F, J, L, M, O, Q) showing variability (as defined by less than 90% of their decisions being the same) in whether they chose to operate on fractures they had previously judged as type IIA. This is followed by the type IIB and type I categories, with 4/11 surgeons (A, B, F, J) and 1/11 surgeon (J) respectively exhibiting variability in their decision making. Within each fracture type, variations between surgeons in the decision to operate was largest for type IIA fractures, ranging from 0% (B) to 90.2% (H), followed by type IIB fractures, which ranged from 28% (B) to 100% (H, Q).

Table 2 Individual surgeon treatment preferences, grouped by fracture classification. Shaded boxes indicate a preference for operative management, non-shaded boxes indicate a preference for non-operative management.

Surgeon / Fracture Type	A	B	D	E	F	H	J	L	M	O	Q	Total n/N*
I												0/11
IIA												5/11
IIB												10/11
III												11/11

*n/N: Number of times the preference to operate was indicated over the total number of surgeons who responded.

Discussion

The findings of this study demonstrate only a moderate level of interobserver agreement between surgeons in operative decision making for extension-type supracondylar fractures at baseline. Overall intraobserver agreement was substantial, although individual agreements demonstrated a range between moderate to excellent or almost perfect. The greatest disagreement between surgeons was observed for fractures that surgeons had previously rated as type IIA and, to a lesser extent, those rated as type IIB, highlighting clearly that the practice pattern variability lies within these fractures types. This suggests the existence of differences in treatment approaches between surgeons and a lack of consensus specific to the appropriate management of type IIA fractures. Interestingly, surgeons' operative decision making for type IIA fractures and type IIB fractures did not completely follow the patterns expected by their preferred treatment methods for each fracture type. This suggests that the modified Gartland classification system plays a smaller role than thought in operative decision-making for type II fractures. Decision

Table 3 Percentage of decisions to operate based on respondents' pooled fracture classifications. Highlighted in bold are the surgeons observed to be variable in their operative decision-making.

Surgeon	A	B	D	E	F	H	J	L	M	O**	Q	
Fracture Type	(%)											Total n/N* (%)
I	0	0	0	0	0	9.5	**33.3**	7.5	5.3	0	5.3	31/385 (8.1)
IIA	8.2	0	**64.2**	**24.1**	**23.9**	90.2	**37.8**	**73**	**71.4**	**22.2**	**73.3**	509/1091 (50.8)
IIB	**61.8**	**28**	91.2	92.6	**73.2**	100	**83.3**	95.5	95.2	90.2	100	452/545 (82.9)
III	100	100	100	100	98.4	100	100	100	100	100	100	498/499 (99.8)
Total n/N* (%)	102/240 (42.5)	70/240 (29.2)	164/240 (68.3)	120/240 (50)	126/240 (52.5)	192/240 (80)	144/240 (60)	174/240 (72.5)	174/240 (72.5)	60/120 (50)	164/240 (68.3)	

*n/N: Number of decisions to operate over the total number of fractures of each type rated by each surgeon. For ease of reference: for each fracture type, surgeons who had previously indicated a preference for operative management have been shaded grey.
**Data based on only one round of the treatment variability survey as respondent did not fill out second round.

making may be more subject to surgeon bias and take into greater account other factors, including neurovascular injury, swelling, soft tissue status, mechanism of injury, degree of varus or valgus and patient age.

Leung et al.[10] previously explored this issue. They asked three fellowship-trained paediatric orthopaedic surgeons and three orthopaedic residents to review 200 radiographs of supracondylar humerus fractures and classify them according to the modified Gartland classification system, and then indicate whether they would treat the fracture non-operatively or operatively. Their analysis found a substantial level of interobserver agreement (κ = 0.691–0.784) between attending-level surgeons and fair to moderate agreement (κ = 0.397–0.569) between residents. Intraobserver agreement was substantial with a κ value of 0.760. Additionally, they demonstrated close to unanimous agreement between surgeons on whether to operate on type I (3%), IIB (99%) and III (100%) fractures, but fractures classified as type IIA showed greater variability, with decisions to operate being made only 27% of the time. The study authors suggested that the modified Gartland classification system may have limited utility particularly in the management of type IIA and IIB fractures, and that other qualitative characteristics of severity or displacement that are not well quantified by the classification system may be used by surgeons to determine treatment. Ernat et al.[11] similarly argued that classification plays a relatively small role in patient management for these fractures and is perhaps only useful as a surrogate for the extent of soft tissue injury to identify potential neurovascular compromise.

Our study supports the findings of Leung et al.[10] in that there is almost full agreement between surgeons for the non-operative and operative management of type I and III fractures respectively. However, they had concluded that there is agreement that operative management should be used for type IIB fractures and that non-operative management should be used for the majority of type IIA fractures. While we observed similar trends, the split in opinion was much larger, with 50.8% of decisions being to operate on type IIA fractures and only 82.9% of decisions being to operate on type IIB fractures. The strength of our study is that when surgeon respondents were asked to make treatment decisions, they were blinded to the classifications they had assigned each fracture in the previous phase of the study, and thus their decisions might not have been as significantly influenced by fracture classification as in Leung et al.[10] In addition, the respondents were comprised of 11 surgeons in total – the largest number of surgeon respondents in such a study to date – who represent multiple countries and hence may more accurately reflect practice variability across the countries represented.

A potential criticism of this study is that the disagreement between stated preferences and actual decisions could be reflective of either individual variability in fracture classification or actual deviance from preferences. Indeed, surgeon J, the surgeon who demonstrated variability in their treatment decisions for the fractures they had classified as types I, IIA and IIB, had had the lowest intraobserver agreement (0.554, moderate) in the first phase of the study where they classified fractures according to the modified Gartland classification. This suggests that there might be some variability in surgeons' understanding of the classification system. However, the rest of the surgeon respondents had intraobserver agreements ranging from 0.699 (substantial) to 0.898 (excellent or almost perfect) for their classification reliability surveys, suggesting that there should not be a substantial amount of variation in classification and that it would not be a major factor explaining the variability in treatment decisions within rated fracture types. Another limitation of this study is that one surgeon did not respond to the second round of the treatment decision-making surveys, so our conclusions may be based on a slightly smaller sample size and technically incomplete data. Nevertheless, as the remaining 10 (90.9%) surgeons did complete their surveys, the differences in the data should not have significantly altered our findings. A further limitation is that the grouping of operative decisions by fracture type was done based on each surgeon's assigned fracture classifications from the previous study, which is not a gold standard for classification. In addition, there were a significantly larger proportion of radiographs classified by surgeons as type IIA over types IIB, III and I, which could be a contributing factor to our finding that type IIA fractures saw the greatest variability in treatment decision making. However, even though we did not have an even distribution of fracture types, there was still a good distribution with enough numbers of each fracture type to allow for confidence in hypothesis testing. Finally, while we used the first edition of the Wilkins[2] modification of the Gartland classification in this study – where the type IIA fracture is described as having an intact posterior cortex, hinged in extension, with no rotation or translation, and the type IIB fracture is described as having some degree of rotational displacement or translation – there have been variations to the classification of these fractures over the years. This includes the classification of significant rotational deformities as type III instead of type IIB in the eighth edition of the same textbook. This may indicate some variability in the literature of surgeons' understanding of the classification system. However, we attempted to overcome this by including a standard table of the Wilkins-modified criteria with each survey in order to ensure that each respondent was classifying fractures using the same classification system.

In conclusion, our findings suggest an overall moderate level of interobserver agreement and overall substantial level of intraobserver agreement in operative

decision-making for Gartland extension-type supracondylar humerus fractures. The largest source of disagreement between surgeons was observed for fractures that they had previously rated as type IIA, and treatment decisions did not completely follow the patterns that would be expected going by surgeons' preferred treatment methods for each fracture type. This suggests that the modified Gartland classification may not be as useful in predicting the need for internal fixation in type II supracondylar fractures. Factors such as fracture stability, patient age and degree of swelling seem to be important determinants that would help make this decision. Future high-quality, prospective comparative trials are needed to identify important factors in the evaluation and treatment of supracondylar humerus fractures, adjust the classification system used if needed, compare patient outcomes between non-operative and operative management and, finally, establish consensus on the necessity of operative management for type II supracondylar fractures.

Received 30 April 2020; accepted after revision 14 September 2020

COMPLIANCE WITH ETHICAL STANDARDS

FUNDING STATEMENT

This work was partially funded by I'm a HIPpy Foundation, BC Children's Hospital Research Institute, and the Pediatric Orthopaedic Society of North America (POSNA).

ETHICAL STATEMENT

Ethical approval: The study has been approved by an institutional research ethics board. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: This research involves human participants who have all given informed consent.

ICMJE CONFLICT OF INTEREST STATEMENT

RE reports grants and personal fees from Medtronic Canada and Depuy Synthes Spine, personal fees from Apifix Ltd., Globus Medical, Wishbone and Joint Solutions Alliance, outside the submitted work.

SC reports grants from ConMED Linvatec and Zimmer Biomet, outside the submitted work.

KM reports grants from International Hip Dysplasia Institute, during the conduct of the study; grants from Depuy, Johnson & Johnson, Pega Medical, Allergan, I'm a HIPpy Foundation, and a patent pending from Pega Medical, outside the submitted work.

The other authors declare no conflict of interest relevant to this work.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Jeffrey Bone for his valuable assistance with the data analysis for this study. We also wish to thank Drs. Alaric Aroojis, Mark Camp, Wudbhav Sankar and Vidyadhar Upasani for their substantial contribution of time as surgeon respondents for this study.

AUTHOR CONTRIBUTIONS

TLT: Developed the study, analysed the data and wrote the manuscript.
EKS: Contributed to the study development and edited the manuscript.
EH: Made intellectual contributions and edited the manuscript.
RH: Made intellectual contributions and edited the manuscript.
PL: Made intellectual contributions and edited the manuscript.
BS: Made intellectual contributions and edited the manuscript.
AA: Made intellectual contributions and edited the manuscript.
SC: Made intellectual contributions and edited the manuscript.
CR: Oversaw study completion and data interpretation and advised on the writing of the manuscript.
KM: Oversaw study completion and data interpretation and advised on the writing of the manuscript.

REFERENCES

1. **Alton TB, Werner SE, Gee AO.** Classifications in brief: the Gartland classification of supracondylar humerus fractures. *Clin Orthop Relat Res* 2015;473:738-741.

2. **Wilkins KE.** Fractures and dislocations of the elbow region. In: Rockwood CA, Wilkins KE, King R, eds. *Fractures in Children*. Philadelphia: Lippincott; 1984:363-575.

3. **Mallo G, Stanat SJC, Gaffney J.** Use of the Gartland classification system for treatment of pediatric supracondylar humerus fractures. *Orthopedics* 2010;33:19-22.

4. **Howard A, Mulpuri K, Abel MF, et al; American Academy of Orthopaedic Surgeons.** The treatment of pediatric supracondylar humerus fractures. *J Am Acad Orthop Surg* 2012;20:320-327.

5. **Ariyawatkul T, Eamsobhana P, Kaewpornsawan K.** The necessity of fixation in Gartland type 2 supracondylar fracture of the distal humerus in children (modified Gartland type 2A and 2B). *J Pediatr Orthop B* 2016;25:159-164.

6. **Ladenhauf HN, Schaffert M, Bauer J.** The displaced supracondylar humerus fracture: indications for surgery and surgical options: a 2014 update. *Curr Opin Pediatr* 2014;26:64-69.

7. **Kish AJ, Hennrikus WL.** Fixation of type 2a supracondylar humerus fractures in children with a single pin. *J Pediatr Orthop* 2014;34:e54-e57.

8. **Teo TL, Schaeffer EK, Habib E, et al.** Assessing the reliability of the modified Gartland classification system for extension-type supracondylar humerus fractures. *J Child Orthop* 2019;13:569-574.

9. **Landis JR, Koch GG.** The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-174.

10. **Leung S, Paryavi E, Herman MJ, Sponseller PD, Abzug JM.** Does the modified Gartland classification clarify decision making? *J Pediatr Orthop* 2018;38:22-26.

11. **Ernat J, Ho C, Wimberly RL, Jo C, Riccio AI.** Fracture classification does not predict functional outcomes in supracondylar humerus fractures: a prospective study. *J Pediatr Orthop* 2017;37:e233-e237.