ORIGINAL REPORTS

Examination of Skill Acquisition and Grader Bias in a Distal Radius Fracture Fixation Model

Matthew D. Putnam, MD,[*] Julie E. Adams, MD,[†] Paul Lender, BS,[*] Ann E. Van Heest, MD,[*] Janet R. Shanedling, PhD,[‡] David J. Nuckley, PhD,[§] and Joan E. Bechtold, PhD[*,‖]

[*]Department of Orthopaedic Surgery, University of Minnesota, Minneapolis, Minnesota; [†]Mayo Clinic, Rochester, Minnesota; [‡]Clinical and Translational Science, University of Minnesota, Minneapolis, Minnesota; [§]Spine Division, Stryker Corporation, Allendale, New Jersey; and [‖]Minneapolis Medical Research Foundation and Excelen Center for Bone and Joint Education and Research, Minneapolis, Minnesota

OBJECTIVES: *Primary*: Assess the ability of faculty graders to predict the objectively measured strength of distal radius fracture fixation. *Secondary:* Compare resident skill variation and retention related to other knowable training data.

DESIGN: Residents were allowed 60 minutes to stabilize a standardized distal radius fracture using an assigned fixed-angle volar plate. Faculty observed and subjectively graded the residents without providing real-time feedback. Objective biomechanical evaluation (construct strength and stiffness) was compared to subjective grades. Resident-specific characteristics (sex, PGY, and ACGME case log) were also used to compare the objective data.

SETTING: A simulated operating room in our laboratory.

PARTICIPANTS: Post-graduate year 2, 3, 4, and 5 orthopedic residents.

RESULTS: *Primary:* Faculty were not successful at predicting objectively measured fixation, and their subjective scoring suggests confirmation bias as PGY increased. *Secondary:* Resident year-in-training alone did not predict objective measures ($p = 0.53$), but was predictive of subjective scores ($p < 0.001$). Skills learned were not always retained, as 29% of residents objectively failed subsequent to passing. Notably, resident-reported case-specific experience alone was inversely correlated with objective fixation strength.

CONCLUSIONS: This testing model enabled the collection of objective and subjective resident skill scores. Faculty graders did not routinely predict objective measures, and their subjective assessment appears biased related to PGY. Also, in vivo case volume alone does not predict objective results. Familiar faculty teaching consistency, and resident grading by external faculty unfamiliar with tested residents, might alter these results. (J Surg Ed ▮:▮▮▮-▮▮▮. © 2018 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: fracture, competence, resident, training, bias, skill assessment

COMPETENCIES: Patient Care, Practice-Based Learning and Improvement, Medical Knowledge

INTRODUCTION

Swensen and Cortese[1] have called for transparency in medicine, and identified Ernst Codman as a leader in transparency. Codman[2] details his efforts, beginning in 1905, to standardize surgical reporting which he called "The End Result Idea" in his 1934 book *The Shoulder.* Specifically, in an address to the Philadelphia County Medical Society in 1913, Codman[3] stated that the results of Medical Education directly depended on the "quality" of the medical/surgical work performed. Warshaw,[4] in his

Level of Evidence: Prognostic—Level 2. The residents are assumed to be subjects; they are studied prospectively with IRB approval as exempt. A blinded (biomechanical) standard is applied.

Sources of Funding: This study was directly funded by a Residency/Core Competency Innovation OMeGA grant (xyhi9237XY). In-kind material support (plates, screws, and kits) was provided by Acumed, Arthrex, Biomet (now Zimmer Biomet), DePuy (now Depuy Synthes (a Johnson & Johnson company)), Stryker, Synthes (now Depuy Synthes [a Johnson & Johnson company]), TriMed, and Zimmer (now Zimmer Biomet). The funding sources had no involvement in this study's design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

Correspondence: Inquiries to Matthew D. Putnam, MD, c/o: Joan E. Bechtold, PhD, Department of Orthopaedic Surgery, University of Minnesota, 2450 Riverside Avenue South, Suite R200, Minneapolis, MN 55454; fax: (612) 273-7959; *e-mail addresses:* matthew.d.putnam@gmail.com, bechtold@umn.edu

2014 American College of Surgeons (ACS) Presidential address, detailed Codman's pursuit of excellence noting that Codman struggled to move his "End Result Idea" forward and that transparency has yet to become the standard in surgical activities. In our time of complex surgical options and increasing variation in resident training due to work hours, transparent proof of skill acquisition seems logical.[5]

To this end, we began a program of yearly testing of resident skills. From review of the results of these annual exercises, we documented a utility for Objective Structures Assessment Tools (OSATs) and discovered a disconnect between yearly OITE testing and objective biomechanical measurement of resident skill development.[6-8]

To evaluate skill in fracture care, we developed a distal radius fracture model to mimic a clinical condition. These fractures are common throughout adulthood, affect men and women, can lead to disability, and are increasing in frequency.[9,10] Since this fracture is considered unstable, and cast support alone is insufficient to maintain reduction,[11-13] operative fixation is needed, providing us an opportunistic model to assess operative skill.

Currently, no single tool has proven superior at stabilizing distal radius fractures.[14,15] We chose to use volar plate fixation since this technique is commonly performed by recent graduates and has demonstrated a large increase in use in recent years.[16] Besides achieving adequate reduction, the fracture construct has to support loads during rehabilitation.[17] This is particularly important since current American Academy of Orthopedic Surgery (AAOS) Clinical Practice Guidelines (CPGs) related to distal radius fracture management recommends: "begin post-treatment therapy and active finger motion as soon as possible."[18] In order to withstand the forces experienced during therapy and active digital motion, a certain threshold strength of construct must be obtained.[17] We know from our biomechanical tests of correctly applied plate constructs that they can resist 1.5 to $\geq$ 2 times rehabilitation loads.[19,20] However, inaccurate reductions can magnify non-axial loads and lead to recurring deformity.[21] Our model allows objective assessment of the resident's ability to adequately create the biomechanical stability needed to allow rehabilitation without fracture deformation in addition to subjective evaluation by experienced and trained faculty graders.

Transparent and unbiased measurement of resident skill fulfills an obligation of training programs. To pursue this goal, we measured our residents over 5 years using objective biomechanics as our control and asked if resident fixation skill: (1) was predicted by faculty graders; (2) varied and/or was retained related to other knowable training data.

MATERIALS AND METHODS

Our Institutional Review Board (IRB) ruled this study exempt from HHS policies for human research subjects protections under 45 CFR 46.101(b) as Category 2: Research involving educational tests. Testing occurred in May of each of 5 consecutive years (2010-2014). Participation of PGY 2 to 5 residents was required (excepting for special circumstances). A mandatory instruction meeting occurred 2 to 4 weeks prior to testing. The meeting reviewed grading methods, test structure, and available equipment including fracture fixation devices.

Our institution's Anatomy Bequest Program was the cadaver source. Specimens were matched pairs with no prior injury or wrist disease, and were uniquely numbered for each resident. Residents were assigned personal work stations; arm holders with traction; appropriate surgical instruments; FDA-approved fixed-angle plating systems; access to "mini C-arm"; and all necessary personal protection. The test duration was a firm 60 minutes, and residents received no feedback during the test.

Faculty-completed specimen preparation the week of the exam. Specimens were then refrigerated until test day. Preparation employed a jig designed to ensure creation of repairable extra-articular fractures[6] (Fig. 1B). No specimens were assigned to residents for fixation wherein the bone shape/size was inadequate for fixation.

Graders were faculty members, all expert in fixed-angle distal radius plating. For years 2010 and 2011, one faculty member was not acquainted with the tested residents. The purpose of this grader was to provide information about potential bias, as this grader had not worked with the residents previously and was not aware of their year-in-training status.

Figure 1 details the entire process of this exam (link to video). Figure 1G is a specific picture of the testing set-up. A PowerPoint™ of the process is available (link). Faculty graders and manufacturer's representatives strictly adhered to the policy of providing no advice/assistance to residents. Faculty were present for the entire exam. For each observed resident, faculty completed a Global Rating Scale (GRS), a Checklist, and a Subjective pass/fail rating. After the completion of plate placement, residents dissected the plated radius free from soft tissues leaving the plated radius exposed for visual inspection. Faculty graders visually assessed the specimens specifically to identify any screws penetrating the radiocarpal joint or distal radioulnar joint cartilage or any screws of excessive length, as well plate position and gross restoration of anatomy (i.e., Was fracture reduction achieved?). The specimens were frozen until biomechanics testing (within 2 weeks).

Before testing, radii osteotomized at a level 2 cm proximal to the pronator teres insertion and were potted in acrylic cement. An electromechanical loading machine (MTS Inc., Eden Prairie, MN) axially loaded the distal radius using a pre-formed extension centered on the lunate fossae[19] (link to video). Load and deformation were measured, and stiffness (slope of load vs. deformation) and yield (initiation of failure) were derived for each specimen (Fig. 1I).