

## Objective Structured Assessments of Technical Skills (OSATS) Does Not Assess the Quality of the Surgical Result Effectively

Donald D. Anderson PhD, Steven Long BSE, Geb W. Thomas PhD,  
Matthew D. Putnam MD, Joan E. Bechtold PhD, Matthew D. Karam MD

Published online: 26 October 2015  
© The Association of Bone and Joint Surgeons® 2015

### Abstract

**Background** Performance assessment in skills training is ideally based on objective, reliable, and clinically relevant indicators of success. The Objective Structured Assessment of Technical Skill (OSATS) is a reliable and valid tool that has been increasingly used in orthopaedic skills training. It uses a global rating approach to structure expert evaluation of technical skills with the experts working from a list of operative competencies that are each rated on a 5-point Likert scale anchored by behavioral descriptors. Given the observational nature of its scoring, the OSATS might not effectively assess the quality of surgical results.

**Questions/purposes** (1) Does OSATS scoring in an intraarticular fracture reduction training exercise correlate with the quality of the reduction? (2) Does OSATS scoring in a cadaveric extraarticular fracture fixation exercise correlate with the mechanical integrity of the fixation?

The institutions of the authors received funding from the National Institute of Arthritis and Musculoskeletal and Skin Diseases of the National Institutes of Health under award numbers P50AR48939 (DDA) and P50AR055533 (DDA, GWT), from the National Board of Medical Examiners® (NBME®) Edward J. Stemmler, MD Medical Education Research Fund (DDA, GWT, MDK), from Core Competency Innovation Grants provided by the OMeGA Medical Grants Association (DDA), and from the Orthopaedic Trauma Association (MDK, DDA).

All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

*Clinical Orthopaedics and Related Research*® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA-approval status, of any drug or device prior to clinical use.

Each author certifies that his or her institution approved the human protocol for this investigation, that all investigations were conducted in conformity with ethical principles of research, and that informed consent for participation in the study was obtained.

**Methods** Orthopaedic residents at the University of Iowa (six postgraduate year [PGY]-1s) and at the University of Minnesota (seven PGY-1s and eight PGY-2s) undertook a skills training exercise that involved reducing a simulated intraarticular fracture under fluoroscopic guidance. Iowa residents participated three times during 1 month, and Minnesota residents participated twice with 1 month between their two sessions. A fellowship-trained orthopaedic traumatologist rated each performance using a modified OSATS scoring scheme. The quality of the articular reduction obtained was then directly measured. Regression analysis was performed between OSATS scores and two metrics of articular reduction quality: articular surface deviation and estimated contact stress. Another skills training exercise involved fixing a simulated distal radius fracture in a cadaveric specimen. Thirty residents, distributed across four PGY classes (PGY-2 and PGY-3,

This work was performed at The University of Iowa, Iowa City, IA, USA, and the University of Minnesota, Minneapolis, MN, USA.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11999-015-4603-4) contains supplementary material, which is available to authorized users.

D. D. Anderson (✉), S. Long, G. W. Thomas, M. D. Karam  
Department of Orthopaedics and Rehabilitation, The University of Iowa, 2181 Westlawn Building, Iowa City, IA 52242, USA  
e-mail: don-anderson@uiowa.edu

D. D. Anderson, S. Long  
Department of Biomedical Engineering, The University of Iowa, Iowa City, IA, USA

D. D. Anderson, G. W. Thomas  
Department of Mechanical and Industrial Engineering, The University of Iowa, Iowa City, IA, USA

n = 8 each; PGY-4 and PGY-5, n = 7 each), simultaneously completed the exercise at individual stations. One of three faculty hand surgeons independently scored each performance using a validated OSATS scoring system. The mechanical integrity of each fixation construct was then assessed in a materials testing machine. Regression analysis was performed between OSATS scores and two metrics of fixation integrity: stiffness and failure load.

**Results** In the intraarticular fracture model, OSATS scores did not correlate with articular reduction quality (maximum surface deviations:  $R = 0.17$ ,  $p = 0.25$ ; maximum contact stress:  $R = 0.22$ ,  $p = 0.13$ ). Similarly in the cadaveric extraarticular fracture model, OSATS scores did not correlate with the integrity of the mechanical fixation (stiffness:  $R = 0.10$ ,  $p = 0.60$ ; failure load:  $R = 0.30$ ,  $p = 0.10$ ).

**Conclusions** OSATS scoring methods do not effectively assess the quality of the surgical result. Efforts must be made to incorporate assessment metrics that reflect the quality of the surgical result.

**Clinical Relevance** New objective, reliable, and clinically relevant measures of the quality of the surgical result obtained by a trainee are urgently needed. For intraarticular fracture reduction and extraarticular fracture fixation, direct physical measurement of reduction quality and of mechanical integrity of fixation, respectively, meet this need.

## Introduction

Performance assessment plays a vital role in surgical education and training. It is most valuable when based on objective, reliable, and clinically relevant indicators of success. The Objective Structured Assessment of Technical Skill (OSATS) is a reliable and valid tool for assessing technical skill that has been increasingly used in surgical skills training [3, 11, 15]. It relies on a global rating approach to structure expert evaluation of technical skills. Evaluators work from a list of operative competencies, each rated on a 5-point Likert scale and anchored by behavioral descriptors. Evaluating residents with this type of standardized, multiple-item global rating scale is reliable and has demonstrated construct validity in other surgical disciplines [12, 15].

M. D. Putnam, J. E. Bechtold  
Department of Orthopaedic Surgery, University of Minnesota,  
Minneapolis, MN, USA

J. E. Bechtold  
Departments of Mechanical Engineering and Biomedical  
Engineering, University of Minnesota, Minneapolis, MN, USA

OSATS-based assessments have also been used in orthopaedics [16–18]. However, not all factors that influence orthopaedic surgical outcomes are amenable to expert visual evaluation, and these factors may be equally or more important than directly observable technical expertise. For instance, in the case of intraarticular fracture reduction, performance evaluation should consider the precision with which the joint is restored, an important factor in determining the likelihood of posttraumatic osteoarthritis [2, 4]. Similarly, in the case of extraarticular fracture reduction, performance evaluation should consider the strength of the associated fixation construct, because the construct serves to maintain the reduction during the course of early recovery and its mechanical integrity substantially influences fracture healing [14]. Previous work in surgical skills assessment and training has rarely addressed these objective metrics of surgical performance.

The present study addresses the following questions: (1) Does OSATS scoring in an intraarticular fracture reduction training exercise correlate with the quality of the reduction? (2) Does OSATS scoring in a cadaveric extraarticular fracture fixation exercise correlate with the mechanical integrity of the fixation?

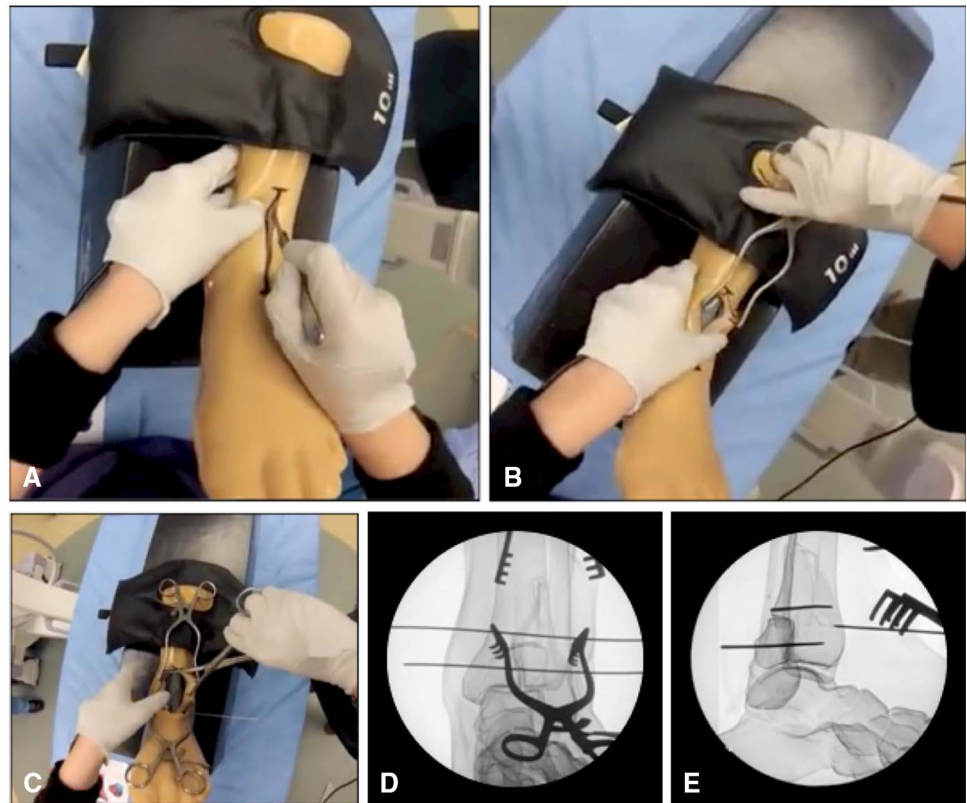
## Materials and Methods

The first study assessed reductions obtained with an intraarticular fracture reduction simulation focusing on restoration of articular congruity. The second study assessed fixations obtained with a cadaveric model of an extraarticular distal radius repaired with a plate, focusing on mechanical stability.

### Study 1: Reduction Quality in an Intraarticular Fracture Reduction Simulation

The intraarticular fracture reduction simulation used a three-segment, radioopaque polyurethane foam distal tibia surrogate housed in a synthetic soft tissue sleeve (Sawbones, Inc, Vashon Island, WA, USA) [18]. The trainee's task was to reduce (Fig. 1A–C) and fix with Kirschner wires (Fig. 1D–E) this simulated tibial plafond fracture operating through a limited direct anterior exposure with the aid of a C-arm fluoroscope and standard surgical instruments. Participants were given 15 minutes to complete the exercise. Performance assessments for the simulation included: the number of fluoroscopic images obtained, the task duration, and OSATS score [18]. A fellowship-trained orthopaedic traumatologist (MDK), blinded to the residents' experience level, rated each participant using a modified OSATS

**Fig. 1A–E** (A–C) The articular fracture reduction model is shown at various stages in the exercise, revealing the radioopaque surrogate bone specimen with soft tissue sleeve. (D) AP and (E) lateral fluoroscopic images of the intraarticular distal tibia fracture are shown, nearing final reduction with Kirschner wires having been placed by the resident.



scoring scheme (Supplemental Table 1 [Supplemental materials are available with the online version of *CORR*®.] [18] either by directly observing or viewing a video of the exercise. No intraobserver reliability testing was done as part of this study.

Six University of Iowa postgraduate year (PGY)-1 orthopaedic residents, seven University of Minnesota PGY-1 orthopaedic residents, and eight University of Minnesota PGY-2 orthopaedic residents participated. The University of Iowa residents performed the simulation three times during the month of January 2013. The University of Minnesota residents performed the simulation twice, first during March 2013 and again in April 2013. This resulted in 48 different sessions with the simulator. Iowa residents' first two trials occurred on the same day with approximately 3 hours between trials. Dedicated training between the first and second trials involved one-on-one coaching and didactic instruction. A third trial was done 2 weeks later to assess skills retention. The Minnesota residents were randomly assigned to two cohorts: (1) an intervention group of four PGY-1 and four PGY-2 residents received video coaching between their first and second trial; and (2) a control group of three PGY-1 and four PGY-2 residents received training after their second trial. We previously reported that the OSATS scores improved in the video-coached group compared with their control cohort [6].

Rather than adopting the traditional metric of “stepoff” magnitude in a radiographic view for assessing the quality of the fracture reductions, a method previously shown to be unreliable [10], we developed a full-surface deviation analysis. This describes the full three-dimensional (3-D) exterior surface of the bone, not just specific points selected by the orthopaedic surgeon. The reduction was first directly assessed by measuring the 3-D deviations of the reduced articular surfaces from their intact positions. To understand the impact of these surface deviations on the ankle contact mechanics, a second analysis method assessed contact stresses predicted in the fracture-reduced configuration using previously established computational stress analysis methods [7, 9].

The assessments of reduction quality were based on the reduced articular surfaces, which were described by 3-D digital models of the final reductions, obtained either with a NextEngine 3-D laser scanner (NextEngine, Inc, Santa Monica, CA, USA) or from segmented CT data. Although the laser scans and the CT segmentations might be expected to produce slightly different surface representations, we did not study this nor would we expect the differences to be large enough to appreciably alter the metrics of reduction quality that were used. Both acquisition methods produce generally noisy surface models with occasional missing and/or aberrant point locations that

were filtered, smoothed, and cleaned up (holes filled and noise reduced) with the built-in capabilities of the Geomagic Studio (Geomagic, Inc, Research Triangle Park, NC, USA) software.

The reduced surface models were compared with an ideal, intact digital model of the distal tibia provided by Sawbones. First, the reduced and intact models were aligned relative to the proximal (intact) segment of the tibia using an iterative closest point algorithm. Then the other fragments were positioned appropriately relative to this proximal base. Next the ideal, intact surface was manually partitioned into individual fragments to match the fragments in the reduced model. Then each fragment was registered to its corresponding ideal counterpart using the Geomagic iterative closest point algorithm, yielding the complete spatial transformation between the two surfaces.

A MATLAB (Version 2014a; MathWorks, Inc, Natick, MA, USA) script was used to perform a surface deviation mapping using a point-to-point method, computing the Euclidean distances between 3-D locations of points on the reduced test surface and corresponding points on the ideal intact surface (Fig. 2A). The mean and maximum of the spatial distribution of differences were both examined as potentially useful metrics of articular surface restoration.

The second assessment of surface incongruity used an expedited computational stress analysis method to estimate the resulting contact stresses associated with a given

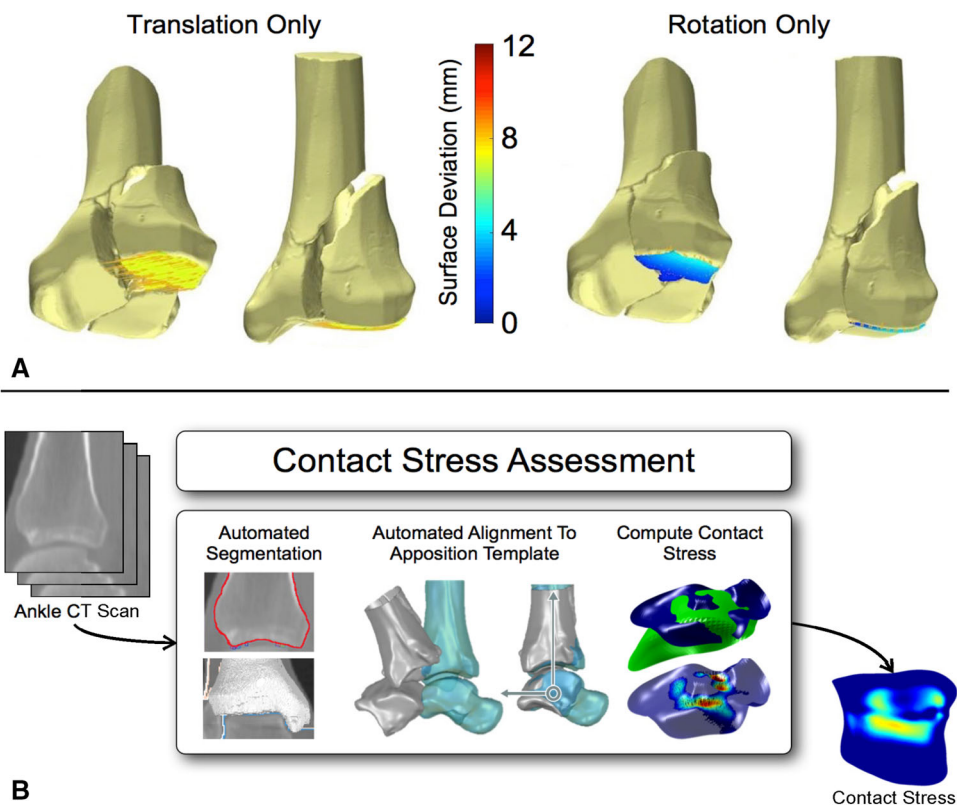
reduction (Fig. 2B). The method was validated in previous work examining the relationship between contact stress and eventual posttraumatic osteoarthritis development in tibial plafond fractures [1, 7, 9]. Each of the 3-D models from the fracture reduction exercise was run through a 13-pose flexion-extension arc representing the stance phase of gait. The contact stress distributions from each loaded pose were queried to compute the maximum contact stress for each ankle model, which was previously shown to be highly predictive of 2-year postoperative posttraumatic osteoarthritis development [1].

Study 2: Mechanical Construct Stability in an Extraarticular Fracture Fixation Simulation

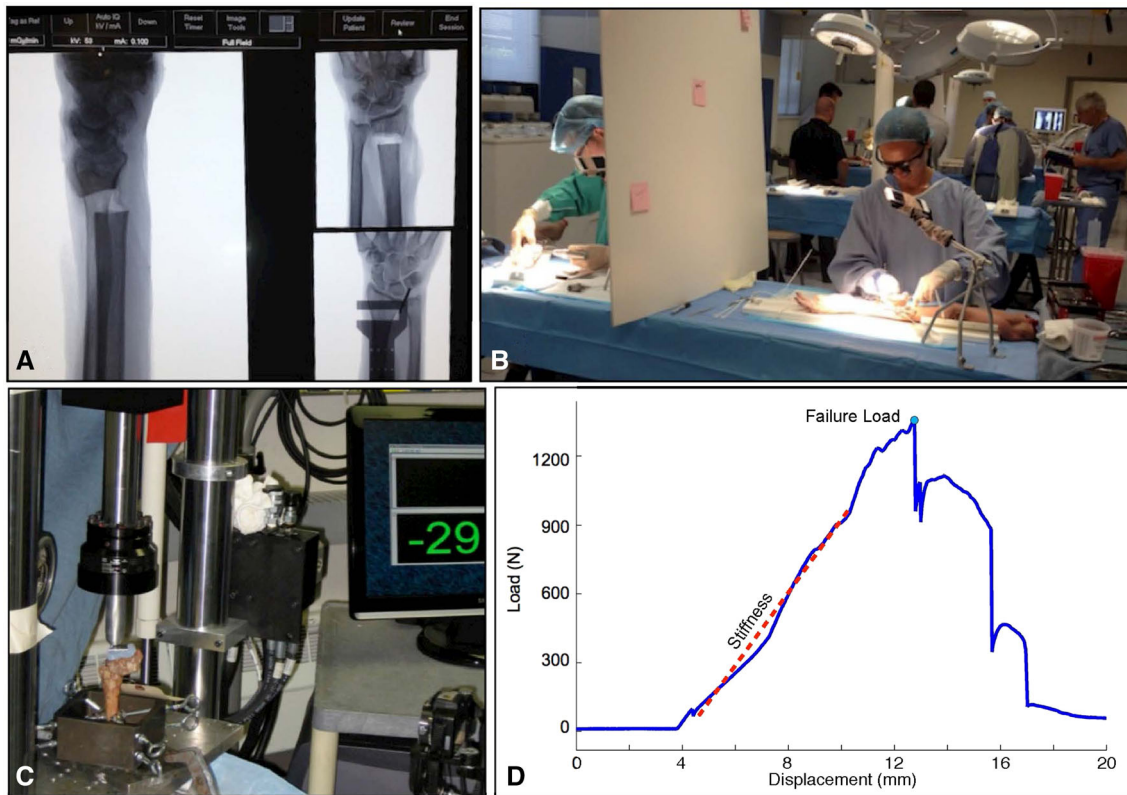
The second simulation involved fixing a simulated extraarticular distal radius fracture in a cadaveric specimen containing a hand, wrist, forearm, and elbow [13]. Identical extraarticular osteotomies of the distal radii were performed using a jig designed to remove 1 cm of metaphyseal bone immediately proximal to the proximal margin of the distal radioulnar joint (Fig. 3A). The osteotomies were performed through a dorsal approach to leave the volar wrist and forearm soft tissues intact.

On the examination day, 30 residents simultaneously completed the exercise at individual stations separated by

**Fig. 2A–B** This graphic depicts the basis of measurements to assess the final articular fracture reduction. (A) The measure of 3-D surface deviations directly reflect the degree of imprecision in the reduction, here shown illustratively for a fragment simply translated a fixed amount or rotated a fixed amount. (B) This graphic shows the basis for computationally estimating contact stress distributions, which indicate the influence of the surface deviations on the ankle contact mechanics.







**Fig. 3A–D** This montage depicts the articular fracture simulation involving the fixing of a simulated extraarticular fracture of the distal radius in an upper extremity cadaver specimen. **(A)** The radiographic images in the upper left show the osteotomy with the cutting jig immediately proximal to the distal radioulnar joint (DRUJ) shown in the lower right. **(B)** Three residents are taking the examination at the same

time here. A faculty member is grading one resident in the background. The faculty grading the residents in the foreground have stepped out of the picture, and a C-arm available to the residents is seen in the background. **(C)** This image shows the loading of the fracture fixation construct. **(D)** This illustrative displacement versus force data tracing shows load uptake and the basis for assessing stiffness and failure load.

vertical screens (Fig. 3B). The residents were evenly distributed across the PGY classes (PGY-2 and PGY-3,  $n = 8$  each; PGY-4 and PGY-5,  $n = 7$  each). Each station included an instrument tray, a fixed-angle distal radius volar plate, a drill/Kirschner wire driver, and a cadaver specimen attached to a forearm-holding device that stabilized the arm/fracture and enabled traction to be applied to the simulated fracture through the fingers. Appropriate tools were available to apply the plate to the bone, including a C-arm fluoroscope for visualizing the reduction and plate position. Participants were allowed 60 minutes and up to six screws, including at most three locking screws, to fixate the fracture with the provided plate. One of three faculty hand surgeons (MDP, TV, CMW) independently scored each participant's performance using three subjective tools: a validated OSATS scoring system (Supplemental Table 2 [Supplemental materials are available with the online version of *CORR*®]) [16], a checklist, and a direct pass/fail determination. No intraobserver reliability testing was done as part of this study.

After completing the exercise, the distal radius was harvested, and the proximal end (proximal to the insertion

of the pronator teres tendon) was embedded in bone cement for mounting before mechanical testing. Compression testing was performed using an MTS 858 Mini Bionix II (MTS Systems Corporation, Eden Prairie, MN, USA) with a 25-mm ball on the distal radius (Fig. 3C). A compressive force was applied at 10 mm/s until construct failure with load and displacement data collected throughout the test (Fig. 3D). The construct was deemed to pass the test if its failure load was greater than 400 N and its stiffness was greater than 80 N/mm. These cutoff values were chosen to be twice the loads applied to the wrist during active digital flexion (an activity promoted immediately postoperatively in clinical care to reduce swelling, pain, and finger stiffness) but less than those achievable by experts using modern plate fixation [5, 8, 14].

#### Data Analysis

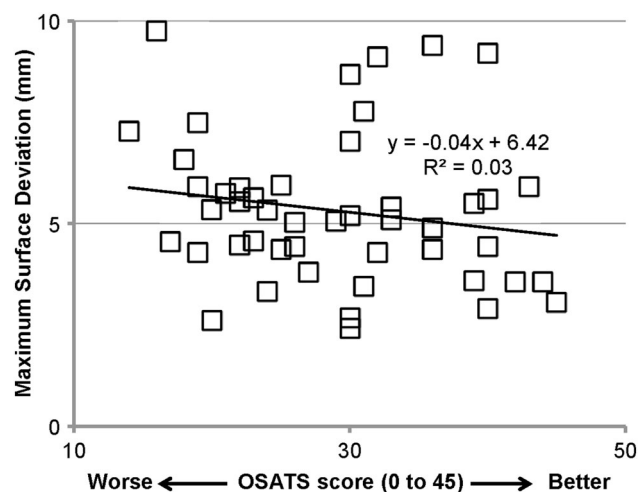
Relationships between OSATS scores and the different metrics of residual articular incongruity (the maximum 3-D surface deviation and the maximum contact stress) and of

fixation stability (the construct stiffness and the failure load) were investigated using simple linear regression techniques with the coefficient of determination ( $R$ ) used as a measure of the goodness of fit. Each performance of an exercise was treated independently for the purposes of the regressions, although individual residents performed multiple trials. This was done because most of the repeat performances were on different days and there was no reason to expect that the relationship between OSATS scores and the different metrics would vary by resident. For the intraarticular fracture reduction model, the 3-D surface deviations measured indicated that a single stepoff measure could not adequately represent the fracture reductions that were obtained. The surface deviations were multidirectional, including both translations and rotations.

## Results

OSATS scores did not correlate well with the quality of the articular reduction, measured either by the maximum surface deviations ( $R = 0.17$ ,  $p = 0.25$ ; Fig. 4) or by the maximum contact stress values ( $R = 0.22$ ,  $p = 0.13$ ; Fig. 5). This lack of agreement could be easily appreciated by inspecting the results in individual case scenarios, where improved OSATS scores were in contrast associated with an elevation of contact stress in the final reduced configuration when going from a baseline (Fig. 6A) to a followup exercise (Fig. 6B) for a given trainee.

Similarly, OSATS scores did not correlate with mechanical integrity of the fixation construct, measured either by the stiffness ( $R = 0.10$ ,  $p = 0.60$ ; Fig. 7) or by the failure load values ( $R = 0.30$ ,  $p = 0.10$ ; Fig. 8)



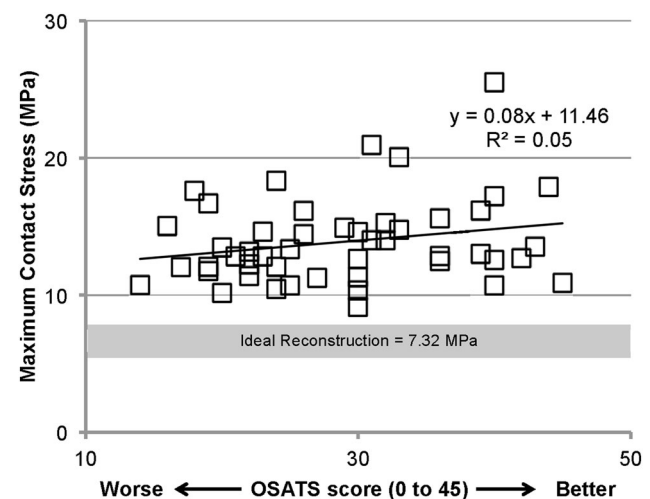
**Fig. 4** The maximum surface deviation was very weakly correlated with the OSATS score, suggesting that the two parameters measure different facets of performance in articular fracture reduction surgery.

Interestingly, it was clear that the OSATS scores rose with increasing experience level with greater certainty than did the mechanical integrity of the fixation construct. For the extraarticular fracture fixation model, OSATS scores were higher ( $p < 0.002$ ) for more senior residents (PGY-4 and 5 mean  $\pm$  SD =  $24.71 \pm 3.67$ ) relative to more junior residents (PGY-2 and 3 =  $20.44 \pm 3.18$ ). The distal radius fracture fixation model could be stabilized to the standards set for failure load and stiffness by the majority of residents. Of the five residents who did not pass for failure load, four were PGY-2 and one was a PGY-5. Only one resident did not exceed the stiffness goal, and it was the same PGY-5 who failed to meet the failure load goal.

## Discussion

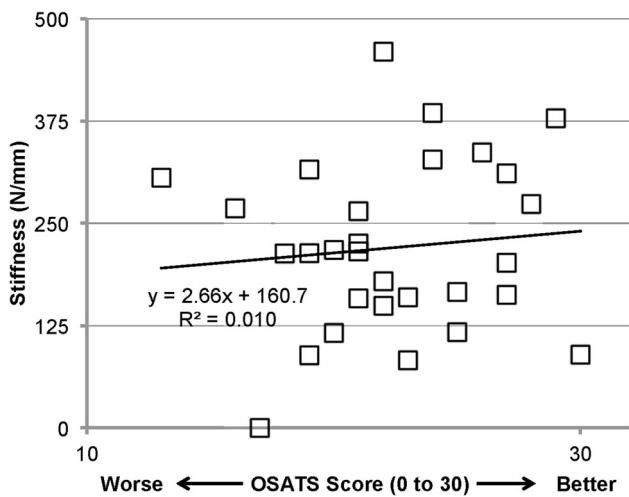
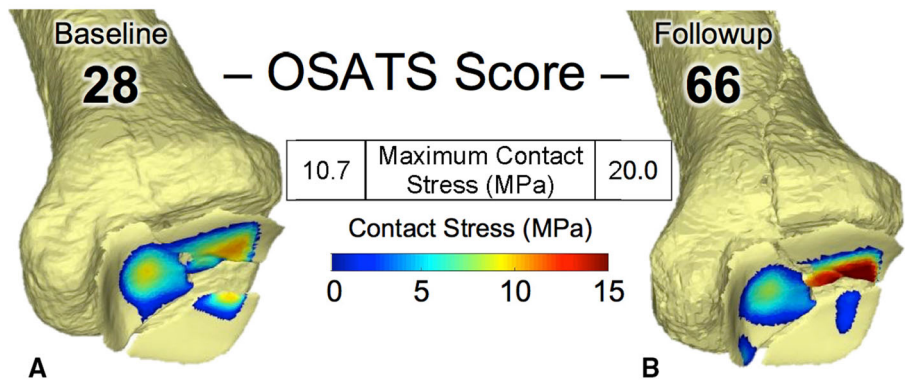
Although prior studies have established progression in the success with which surgical technique is mastered (eg, using OSATS as a measure), few have documented how well trainees restore anatomy and/or achieve adequate fixation. Because expert evaluations such as those done using OSATS scoring depend on external observation of competencies, they risk missing the evaluation of critical elements in the surgical result. In the present study we found that important surgical outcomes associated with precisely reducing intraarticular fractures and solidly fixing extraarticular fractures did not correlate with OSATS scores.

There are limitations to this work that warrant mention. First, neither of the models presented have been tested using practicing orthopaedic surgeons, so the observations should be considered to pertain only to residents in



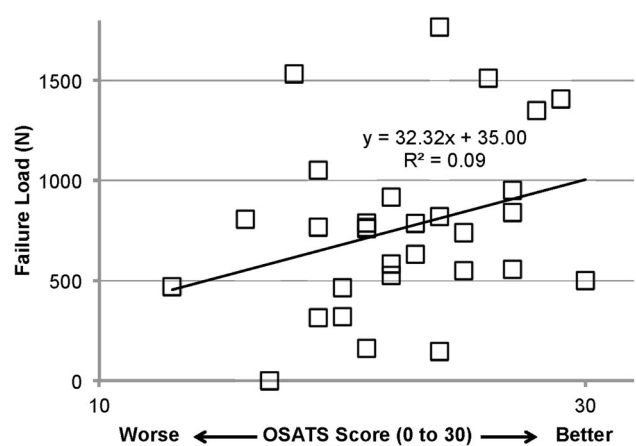
**Fig. 5** The maximum contact stress was very weakly correlated with the OSATS score, suggesting that the two parameters measure different facets of performance in articular fracture reduction surgery.

**Fig. 6A–B** The results from these two trials with the articular fracture reduction exercise illustrate differences in the restoration of the articular surface and the degree of contact stress elevation. These trials were chosen because they showed a representative improvement in OSATS score going from (A) a baseline to (B) a followup exercise, which did not correlate with improvements in contact stress distributions.



**Fig. 7** The stiffness of the fracture fixation construct that was achieved did not correlate well with the OSATS scores.

training. Second, the validity of the data reported depends on equal motivation being applied by all residents who were assessed. This seems a reasonable assumption given the relatively high-stakes circumstances (watched and scored by their instructors) under which they performed. Third, the metrics of reduction quality are not used clinically. We would argue that there is value in a reliable metric of the surgical result, even if it is not readily available in the clinical setting. Fourth, although the fracture fixation model presented has been used in multiple biomechanical testing centers to test the type of fixation examined, the model is a fairly simple version of the fracture often treated. Fifth, although a trainee repeating a task increases the number of observations for the purposes of the regressions, these repeats are not truly independent observations, and they can falsely increase observed effect size and therefore depress measures of statistical significance. However, there was no reason to expect that the relationship between OSATS scores and the different metrics would vary by resident or by performance. If repeat performances were not strictly independent, this would



**Fig. 8** The failure load of the fracture fixation construct that was achieved did not correlate well with the OSATS scores.

tend to falsely suggest significance where there was not any, an outcome that we did not see. Sixth, we did not include inter- or intraobserver reliability testing of OSATS scoring as part of our experimental design. Prior work has established the general reliability of the OSATS approach to the rating of competencies [3, 11, 15], and our prior experience with the specific OSATS scoring schemes used [6, 13, 16, 18] has shown consistency across the relatively few expert raters who have participated.

The results of our first study indicate that OSATS scoring of surgical competency among orthopaedic residents does not correlate with success in restoring articular congruity. To the authors' knowledge, this is the first time that this subject has been specifically studied, which is somewhat alarming. Although beyond the scope of this study, it is possible that a more extensile approach, or a series of limited approaches based on fracture morphology, combined with use of a distraction frame, would yield a better reduction in the hands of trainees. However, regardless of the approach taken, the acquisition of technical competency in performing a surgery is an important first step toward its successful mastery. In the case of

intraarticular fracture reduction, the surgical result importantly influences the clinical outcome.

The results of our second study indicate that OSATS scoring of surgical competency among orthopaedic residents does not correlate with success in achieving mechanically competent extraarticular fixation. Putnam et al. [13] previously reported on the lack of correlation between traditional assessments of orthopaedic resident knowledge (Orthopaedics In-Training Exam [OITE] overall score, OITE trauma score, and an Objective Structured Assessment of Technical Skill [OSATS] score) and the structural integrity of fixation in the extraarticular fracture model that was used in the current study. That prior study tested only a limited number of residents ( $n = 15$ ), and it included only PGY-3 ( $n = 8$ ) and PGY-4 ( $n = 7$ ) residents. The data reported here confirm those earlier findings while adding information regarding the variation over the course of training during orthopaedic residency.

Our results indicate that fracture reduction and fixation skills are often overestimated by OSATS scoring. New objective, reliable, and clinically relevant measures of the quality of the surgical result obtained by a trainee are urgently needed to improve resident assessment. For intraarticular fracture reduction and extraarticular fracture fixation, direct physical measurement of reduction quality and of mechanical integrity of fixation, respectively, meet this need.

**Acknowledgments** We thank Mr Gary Ohrt and Mr Andrew Kern for helping with the work in Iowa, and we gratefully acknowledge Mr Paul Lender for his help with data management and analyses in Minnesota. We also thank Thomas Varecka MD, and Christina M. Ward MD, for their assistance in OSATS scoring of performance on the extraarticular fracture fixation exercise.

## References

1. Anderson DD, Van Hofwegen CJ, Marsh JL, Brown TD. Is elevated contact stress predictive of post-traumatic osteoarthritis for imprecisely reduced tibial plafond fractures? *J Orthop Res.* 2011;29:33–39.
2. Buckwalter JA, Brown TD. Joint injury, repair, and remodeling: roles in post-traumatic osteoarthritis. *Clin Orthop Relat Res.* 2004;423:7–16.
3. Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med.* 1996;71:1363–1365.
4. Furman BD, Olson SA, Guilak F. The development of posttraumatic arthritis after articular fracture. *J Orthop Trauma.* 2006;20:719–725.
5. Gesensway D, Putnam MD, Mente PL, Lewis JL. Design and biomechanics of a plate for the distal radius. *J Hand Surg [Am].* 1995;20:1021–1027.
6. Karam MD, Thomas GW, Koehler DM, Westerlind BO, Lafferty PM, Ohrt GT, Marsh JL, Van Heest A, Anderson DD. Surgical coaching from head-mounted video in the training of fluoroscopically guided articular fracture surgery. *J Bone Joint Surg Am.* 2015;97:1031–1039.
7. Kern AM, Anderson DD. Expedited patient-specific assessment of contact stress exposure in the ankle joint following definitive articular fracture reduction. *J Biomech.* 2015;48:3427–3432.
8. Koh S, Morris RP, Patterson RM, Kearney JP, Buford WL Jr, Viegas SF. Volar fixation for dorsally angulated extra-articular fractures of the distal radius: a biomechanical study. *J Hand Surg [Am].* 2006;31:771–779.
9. Li W, Anderson DD, Goldsworthy JK, Marsh JL, Brown TD. Patient-specific finite element analysis of chronic contact stress exposure after intraarticular fracture of the tibial plafond. *J Orthop Res.* 2008;26:1039–1045.
10. Martin J, Marsh JL, Nepola JV, Dirschl DR, Hurwitz S, DeCoster TA. Radiographic fracture assessments: which ones can we reliably make? *J Orthop Trauma.* 2000;14:379–385.
11. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg.* 1997;84:273–278.
12. Moulton CA, Dubrowski A, Macrae H, Graham B, Grober E, Reznick R. Teaching surgical skills: what kind of practice makes perfect? A randomized, controlled trial. *Ann Surg.* 2006;244:400–409.
13. Putnam MD, Kinnucan E, Adams JE, Van Heest AE, Nuckley DJ, Shanedling J. On orthopedic surgical skill prediction—the limited value of traditional testing. *J Surg Educ.* 2015;72:458–470.
14. Putnam MD, Meyer NJ, Nelson EW, Gesensway D, Lewis JL. Distal radial metaphyseal forces in an extrinsic grip model: implications for postfracture rehabilitation. *J Hand Surg [Am].* 2000;25:469–475.
15. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative ‘bench station’ examination. *Am J Surg.* 1997;173: 226–230.
16. Van Heest A, Kuzel B, Agel J, Putnam M, Kalliainen L, Fletcher J. Objective structured assessment of technical skill in upper extremity surgery. *J Hand Surg [Am].* 2012;37:332–337.
17. Van Heest A, Putnam M, Agel J, Shanedling J, McPherson S, Schmitz C. Assessment of technical skills of orthopaedic surgery residents performing open carpal tunnel release surgery. *J Bone Joint Surg Am.* 2009;91:2811–2817.
18. Yehyawi TM, Thomas TP, Ohrt GT, Marsh JL, Karam MD, Brown TD, Anderson DD. A simulation trainer for complex articular fracture surgery. *J Bone Joint Surg Am.* 2013;95:1–8.