Predicting Shoulder Arthroplasty Outcomes: Powered with CT-based Musculoskeletal Radiomics

Gregory Spangenberg[1], Caleb Bowyer[1], Josie Elwell[1], Christine Allen[1], Francois Boux de Casson[1], William R. Aibinder[2], Bradley Schoch[3], Bruno Gobbato[4], Scott W. Trenhaile[5], [6]Anshuman Singh, [7]John Erickson, Hamidreza Rajabzadeh-Oghaz[1], Christopher P. Roche[1]

[1] AdvitaOrtho., Gainesville, FL, USA, [2]University of Michigan, Ann Arbor, MI, USA [3]Mayo Clinic, Jacksonville, FL, US, [4]R. José Emmendoerfer, Jaraguá do Sul, SC, Brazil, [5]OrthoIllinois, Rockford, IL, USA, [6]University of California, San Diego, CA USA. [7]Atlantic Medical Group, New Providence, NJ, USA

Disclosures: G. Spangenberg: 3A; AdvitaOrtho, **C. Bowyer:** 3A; AdvitaOrtho, **J. Elwell:** 3A; AdvitaOrtho, **Christine Allen:** 3A; AdvitaOrtho, **Francois Boux de:** 3A; AdvitaOrtho, **W. Aibinder:** 3B; AdvitaOrtho, **B.Schoch:** 3B; AdvitaOrtho, **B.Gobbato:** 3B; AdvitaOrtho. **S.Trenhaile:** 3B; AdvitaOrtho. **A. Singh:** 3B; AdvitaOrtho. **J. Erick son:** 3B; AdvitaOrtho, **H. Rajabzadeh-Oghaz:** 3A; AdvitaOrtho. **C. Roche:** 3A; AdvitaOrtho.

INTRODUCTION: Previous studies have demonstrated that ML-based clinical decision support tools (CDSTs) can predict postoperative outcomes, including range of motion (ROM), patient reported outcome measures (PROMs), and pain following anatomic and reverse total shoulder arthroplasty (aTSA and rTSA) using a minimal, manually entered input feature set comprising patient demographics, diagnosis, comorbidities, and preoperative clinical data including ROM, pain, and function . However, obtaining preoperative clinical data, may require in person physical assessment, can be time-consuming and subject to variability [1]. Previous studies have shown that muscle radiomics can be extracted from CT images and integrated with a minimal feature set of patient data to improve prediction accuracy [2]. The objective of this study is to evaluate 1) whether excluding pre-operative clinical data from the manual minimal feature set (e.g. ROM and pain) significantly affects model accuracy in predicting post-operative outcomes and 2) whether automatically quantified, image-based radiomic data for shoulder muscles can serve as substitutes for these manual inputs to regain potential performance loss.

METHODS: Preoperative CT images and clinical data from primary shoulder arthroplasty patients treated with a single-platform shoulder prosthesis (Equinoxe; Advita, Inc., Gainesville, FL, USA) were analyzed. All patients were prospectively enrolled and data was collected via standardized, institutional review board approved forms. Using an internally developed multiclass segmentation model, muscle boundaries of the rotator cuff and deltoid were automatically delineated the. Figure 1 demonstrates an example of the segmented images. All cases were subsequently processed to exclude those with incomplete CT scans with incomplete acquisition of the deltoid or those with metal artifacts. Six radiomic features were then extracted for each muscle: flatness, sphericity, elongation, volume, surface-to-volume ratio, and mean Hounsfield units.

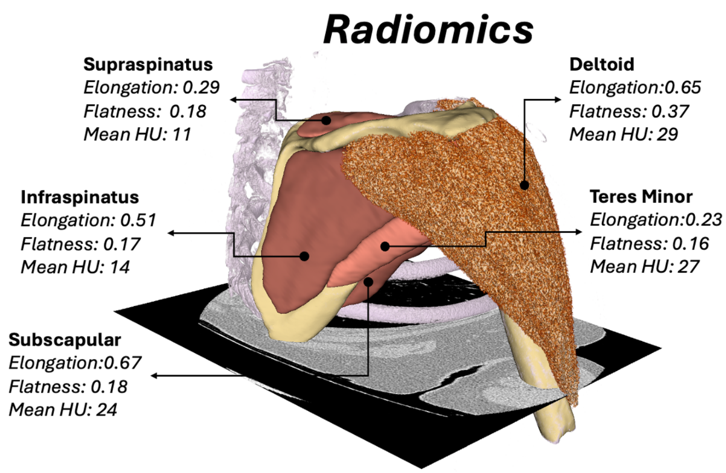

Figure 1: segmentation and radiomic extraction of shoulder muscles, deltoids (muscle property color) and rotator cuffs (solid colors).

Three sets of XGBoost regression prediction models were trained to predict nine outcomes (VAS pain, global shoulder function, SAS score, active abduction, forward elevation, external rotation, IR score, ASES score and Constant score) at two postoperative timepoints (1 and 2 year) after aTSA and rTSA, resulting in 18 models for each procedure. Base models (Base) were trained using a set [1] of 15 inputs consisting of patient demographics (6), diagnosis (1), comorbidities (1), and preoperative clinical data (7). A second set of models (Base-light) was developed using only patient demographics, diagnosis, and comorbidities. A third set of models (Radiomics-light) was developed with features used by the second model (patient demographics, diagnosis, and comorbidities) and the aforementioned radiomics of deltoid and rotator cuff muscles. The models were trained using an 80:20 train-test split ratio, where the test data set was held out completely from model training. Performance of these three models was compared via MAE.

RESULTS AND DISCUSSION: In total, data from 2426 patients and 6902 follow-up visits were included in the current study (1298 F/1128 M; age 70.0 ± 8.3 years; 521 aTSA/1,905 rTSA). Table 1 presents the mean absolute error (MAE) results on the holdout test datasets. In the Base-light models, performance decreased for 16 out of 18 models for both aTSA and rTSA compared to the Base models. However, when radiomic features were added, the Radiomics-light models showed better performance in 13 out of 18 models for both procedures compared to the Base-light models. Radiomics-light models also achieved the same or better accuracy than Base models in 9 and 3 models, for aTSA and rTSA, respectively. These results highlight the utility of radiomic features as potentially effective replacements for manually assessed pre-operative ROM and pain metrics, enabling a more automated, objective, and efficient predictive modeling pipeline. One limitation of this study is that radiomic features from the scapula and humerus were not included. Another limitation is that we did not include all possible radiomic features but rather focused on those identified as robust and less sensitive to imaging parameters.

SIGNIFICANCE/CLINICAL RELEVANCE: Incorporating radiomic features derived from preoperative standard of care medical CT scans can reduce reliance on time-intensive and subjective clinical assessments requiring manual input entry, facilitating automated and more standardized predictive modeling to support preoperative decision-making in shoulder arthroplasty.

REFERENCES:

[1] Rajabzadeh-Oghaz, H., et al. (2025). Radiomic analysis of the deltoid and scapula: Identification of CT image–based measurements predictive of pain, motion, and function before and after shoulder arthroplasty. JSES International.

[2] Kumar, V. (2021). Using machine learning to predict clinical outcomes after shoulder arthroplasty with a minimal feature set. J. Shoulder Elb.

Table 1: Comparison of Test MAE associated with three different ML models to predict clinical outcomes, 1 year and 2 year for aTSA and rTSA patients

	aTSA						**rTSA**					
	1-year			**2-year**			**1-year**			**2-year**		
Model	**Base**	**Base-Light**	**Radiomics-Light**	**Base**	**Base-Light**	**Radiomics-Light**	**Base**	**Base-Light**	**Radiomics-Light**	**Base**	**Base-Light**	**Radiomics-Light**
Abduction	22.14 ± 1.89	22.93 ± 2.22	23.47 ± 2.14	17.17 ± 2.29	23.03 ± 3.24	19.63 ± 2.57	23.83 ± 1.22	27.6 ± 1.39	24.39 ± 1.25	23.84 ± 1.46	26.42 ± 1.54	23.84 ± 1.39
External Rotation	13.88 ± 1.55	17.32 ± 1.72	15.14 ± 1.51	15.5 ± 1.6	18.01 ± 1.76	13.36 ± 1.56	13.23 ± 0.84	17.47 ± 1.01	14.33 ± 0.84	12.76 ± 0.82	17.34 ± 1.23	13.91 ± 0.94
Forward Elevation	18.02 ± 2.0	19.99 ± 2.27	17.22 ± 1.93	19.97 ± 2.7	23.37 ± 3.76	17.84 ± 2.97	18.14 ± 1.11	21.45 ± 1.38	18.66 ± 1.25	15.09 ± 1.0	18.59 ± 1.24	15.84 ± 0.92
ASES	11.17 ± 1.22	10.18 ± 1.15	10.64 ± 1.11	11.58 ± 1.53	12.26 ± 1.36	10.97 ± 1.45	11.94 ± 0.68	12.61 ± 0.7	12.62 ± 0.65	12.2 ± 0.77	12.11 ± 0.79	12.2 ± 0.72
Constant	8.83 ± 0.96	8.9 ± 0.96	8.83 ± 1.06	7.12 ± 1.1	7.2 ± 1.09	8.0 ± 1.1	8.08 ± 0.65	8.18 ± 0.65	9.33 ± 0.63	8.84 ± 0.54	9.75 ± 0.76	9.64 ± 0.61
Internal Rotation	1.08 ± 0.11	1.27 ± 0.11	1.17 ± 0.11	1.03 ± 0.11	1.14 ± 0.13	1.15 ± 0.14	1.34 ± 0.07	1.58 ± 0.08	1.41 ± 0.07	1.37 ± 0.08	1.39 ± 0.1	1.42 ± 0.08
VAS pain	1.21 ± 0.15	1.4 ± 0.17	1.15 ± 0.12	1.18 ± 0.16	1.51 ± 0.19	1.43 ± 0.16	1.33 ± 0.08	1.47 ± 0.1	1.44 ± 0.07	1.34 ± 0.08	1.29 ± 0.09	1.44 ± 0.09
SAS	7.71 ± 0.88	9.12 ± 1.06	6.35 ± 0.69	7.61 ± 1.04	8.41 ± 1.16	7.69 ± 1.07	7.79 ± 0.44	9.23 ± 0.56	8.06 ± 0.46	7.16 ± 0.43	7.6 ± 0.55	6.72 ± 0.46
Shoulder Function	1.37 ± 0.15	1.49 ± 0.17	1.32 ± 0.15	1.59 ± 0.2	1.47 ± 0.21	1.75 ± 0.22	1.46 ± 0.09	1.61 ± 0.1	1.46 ± 0.09	1.37 ± 0.1	1.58 ± 0.11	1.4 ± 0.09