Machine-Learning Model for Quantification of Deltoid Characteristics

Hamidreza Rajabzadeh-Oghaz, Josie Elwell, Vikas Kumar, Lhoussein Mabrouk, Clement Daviller,
David Berry, Anshuman Singh, Sandrine Polakovic, Bradley Schoch, Chris Roche

Disclosures: Rajabzadeh-Oghaz, Elwell, Kumar, Mabrouk, Daviller: 3A-Exactech, Inc; Berry, Singh: Nothing to disclose;
Polakovic P: 3A-Exactech, Inc; Schoch: 3B-Exactech; Inc, Roche: 3A-Exactech, Inc. 4-Exactech Inc.

INTRODUCTION: Treatment planning of patients who need shoulder arthroplasty relies on various pre-operative factors, including integrity and condition of patient's joint, bone, and muscle quality. The deltoid is the primary elevator of the joint contributing to internal/external rotation and joint stability. Despite importance of the deltoid and its potential impact on shoulder arthroplasty outcomes, particularly reverse total shoulder arthroplasty, there have been only a limited number of studies that have attempted to establish a connection between deltoid characteristics and functional results post shoulder arthroplasty.[1], [2] This is likely because current clinical practice to evaluate soft tissue relies on visual inspection of medical images. The manual delineation of boundaries and characterization of muscle quality from image modalities such as CT scan is challenging and time-consuming leading to subjectivity in interpretations and inconsistencies. In this study, we aimed to develop and validate CT-based machine-learning frameworks to segment deltoid muscle and quantify various muscle characteristics, including basic and radiomic features.

METHODS: A total of ninety-eight pre-operative CT images from patients (71% female, average age of 70 ± 7, 62% with osteoarthritis, 26% rotator cuff tear, 31% rotator cuff arthroplasty) enrolled in a multi-center, IRB-approved database undergoing total shoulder arthroplasty with a single system (Exactech, Inc., Gainesville, FL) were identified to train and test the model. The patients included in the dataset were selected at random, ensuring a diverse representation of demographics, diagnoses, and various image attributes, such as kernel and manufacturer. Trained technicians, under supervision of an experienced orthopedic shoulder specialist, delineated deltoid boundaries. Seventy-eight labeled masks were utilized to fine-tune a pre-trained model [3], which was deployed and tested on the remaining 20 images.

Following segmentation of the 20 images in the test set, a quantification algorithm overlaid segmented masks onto the raw images and stratified voxels to fat and muscle groups based on their Hounsfield values. The quantification framework was used to measure several features, including deltoid volume, *normalized volume* $NV = V_{Deltoid}/V_{Scapula}$; *normalized atrophy* $Atrophy = NV/NV_{age,gender}$, where $NV_{age,gender}$ is the average of NV of patient within same gender and age group; and *fatty infiltration* $FI = V_{fat}/V_{fat} + V_{muscle}$. Beyond the basic features, additional radiomic features capturing voxel intensity and texture characteristics are extracted as well [4]. The segmentation and quantification frameworks outcomes from a sample CT scan is illustrated in Figure 1. This figure illustrates three key steps – 1) deltoid segmentation using ML, and 2) stratification of fat and muscle, and 3) the quantification of both basic and radiomic features.

Subsequent to segmentation and quantification of the 20 masks in the test set, a qualitative assessment was conducted categorizing the segmentations as requiring no correction, minor correction, or major correction. A comparison between the machine-learning-generated masks and the benchmark ("ground truth") masks was performed. Metrics included are dice coefficient, distance map (average of distance between benchmark and machine-generated masks, as shown in Figure 2), as well as the percentage error in atrophy and fatty infiltration.

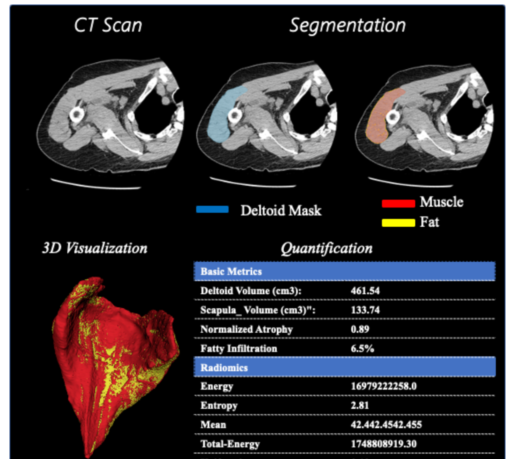

Figure 1. Workflow for segmentation, quantification, and visualization of deltoid muscle.

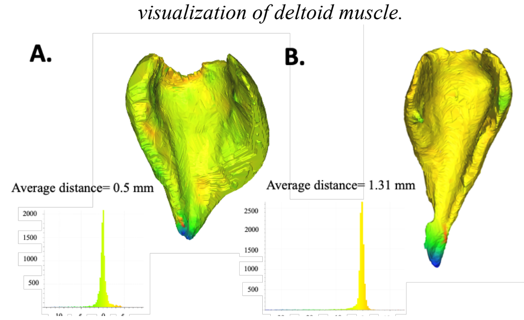

Figure 2. Distance map between the ML-generated mask and benchmark mask for two examples.

RESULTS SECTION: For the qualitative evaluation, the 20 ML-generated masks were rated as follows: 9 masks required no correction, 5 masks required minor correction, and 6 masks required major correction. The quantified validation metrics are presented in Table 1 according to qualitative rating. Across all masks, the average dice-coefficient was 0.93 ± 0.03. The average distance for all cases was 1.04 ± 0.98mm. The percentage error for atrophy and fatty infiltration was 6.29% ± 5.69% and 13.73% ± 16.87%, respectively. Although the dice-coefficient exhibited similar values for cases categorized based on the need for correction, distance and percentage error in atrophy and fatty infiltration were higher in cases requiring major correction compared to those needing no or minor correction.

Table 1. Comparison between ML-Generated masks and Benchmark masks

	Dice Coefficient	Distance Map (mm)	PE in Muscle Atrophy	MPE in Fatty Infiltration [Median]
Total (20)	0.93±0.03	1.04±0.98	6.26±5.69	13.73 [5.68]
no correction needed (9)	0.95±0.02	0.52±0.26	3.31±1.89	11.43 [4.58]
minor correction needed (5)	0.91±0.03	0.87±0.54	8.28±9.60	14.26 [3.02]
major correction needed (6)	0.92±0.01	1.96±1.34	9.01±3.83	16.74 [9.70]

PE: Percentage Error, MPE: Mean Percentage Error

DISCUSSION: In this paper, we developed and validated a CT-based ML models to delineate the boundary of deltoid muscles. We conducted an evaluation of the machine-learning-generated masks through both qualitative and quantitative assessments. We also demonstrated that ML-generated deltoid masks could be used to quantify both basic and radiomic features. This study had several limitations, including a small sample size, especially for our test cohorts, and a lack of evaluation by multiple users.

SIGNIFICANCE/CLINICAL RELEVANCE: CT-based ML models can be a reliable tool for automated and efficient quantification of muscle characteristics. Correlation of muscle characteristics to clinical outcomes after shoulder arthroplasty should be the subject of future work, which may aid surgeons in treatment planning to optimize range of motion and joint stability.

REFERENCES:

[1] J.-D. Werthel *et al.*, "Three-dimensional muscle loss assessment: A novel computed tomography–based quantitative method to evaluate rotator cuff muscle fatty infiltration," *J Shoulder Elbow Surg*, vol. 31, no. 1, pp. 165–174, 2022.

[2] D. Berry *et al.*, "Assessment of Deltoid Muscle Volume and Fatty Infiltration on Clinical Outcomes After Reverse Total Shoulder Arthroplasty," in *ORS*, 2023.

[3] Y. Tang *et al.*, "Self-supervised pre-training of swin transformers for 3d medical image analysis," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 20730–20740.

[4] A. Fedorov *et al.*, "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magn Reson Imaging*, vol. 30, no. 9, pp. 1323–1341, 2012, doi: https://doi.org/10.1016/j.mri.2012.05.001.