An Accelerated Deep Learning Model Can Accurately Identify Clinically Important Humeral and Scapular Landmarks on Preoperative and Postoperative Anatomic Arthroplasty Plain Radiographs

William Lincoln Crutcher[1], Ishan Maneesh Dane, Anastasia Whitson[1], Frederick A Matsen[2], Jason Hsu
[1]University of Washington, [2]Univ of Washington Med Ctr

INTRODUCTION:

Introduction: Accurate identification of radiographic landmarks is fundamental to characterizing glenohumeral relationships before and sequentially after shoulder arthroplasty. Correlating radiographic measurements with clinical outcomes will guide efforts to identify the surgeon-controlled variables that determine the results experienced by the patient. Currently, landmark identification is laborious and lacks standardization that impacts interobserver variability. Additionally, scaling the process to large numbers of cases across different practices is a challenge. Artificial intelligence is a rapidly developing field and, specifically, computer vision and deep learning models (DLMs) are increasingly utilized for image recognition and classification in an efficient manner. However, it is unclear whether deep learning models can accurately identify reproducible landmarks on shoulder radiographs that contain overlap of multiple osseous structures. We report on the development and testing of an accelerated computer vision and DLM that recognizes standard anatomic landmarks on preoperative and postoperative shoulder radiographs then utilizes these landmarks to make measurements related to component positioning in anatomic shoulder arthroplasty:

1) How much deviation is there between computer vision / DLM-identified landmarks and surgeon-identified (SI) landmarks on preoperative and postoperative shoulder radiographs?

2) How much deviation is there between DLM and SI measurements of scapular, humeral, and glenohumeral relationships in shoulder arthroplasty?

METHODS:

Materials & Methods: To train a deep learning model, 120 preoperative and 120 postoperative true anteroposterior (Grashey) shoulder X-rays were annotated using 11 standard anatomic landmarks (surgeon-identified landmarks) (Figure 1A) – these were divided into 9 cortical landmarks and 2 non-cortical landmarks. The number of training images was artificially increased ten-fold by making various modifications to each base image resulting in a total of 2,640 images in the training model. The accuracy of DLM landmarks was compared to manually annotated radiographs using 60 radiographs not used in the training model. In addition, we also performed 14 different measurements of component positioning (Figure 1B) and compared these to measurements made based off DLM landmarks.

RESULTS:

Results:

1) *__Deviation of DLM vs. SI landmarks__*: The DLM was found to be accurate with respect to anatomic landmark identification when compared to surgeon-identified landmarks for all anatomic locations based off cortical bone. The mean deviation between DLM vs SI cortical landmarks was 1.9 ± 1.9mm. The least deviation was seen at the inferior glenoid rim (1.2 ± 0.7mm) and the tip of the acromion (1.4 ± 1.9mm) (Table 1). Non-cortical landmarks had much higher areas of deviation with metaphyseal and diaphyseal midpoint measurements of 8.1 ± 4.4mm and 10.3 ± 10.5mm of deviation, respectively.

2) *__Deviation of scapular, humeral, and glenohumeral measurements__*: The DLM was also found to be accurate with respect to these measurements. The mean deviation of all measurements was 2.9 ± 3.8mm. Measurement of humeral head radius to humeral head center of rotation (both native and prosthetic) based off a perfect circle had deviation of only 1.2 ± 1.1mm.

DISCUSSION AND CONCLUSION:

Discussion: While many artificial intelligence applications require thousands of inputs to develop accurate output generation, we were able to train a deep learning model using only 240 annotated images. This model was able to reliably identify cortical landmarks with low levels of deviation on preoperative and postoperative radiographs. The reliability and efficiency of these deep learning models represents a powerful tool to analyze preoperative and postoperative x-rays in the context of anatomic shoulder arthroplasty without assessment bias. Similar models for axillary radiographs, reverse shoulder arthroplasty designs, and other considerations could be developed using a similar strategy to improve the power and scope of radiographic studies.

Figure 1. Computer vision / deep learning model identification of (A) anatomic landmarks and (B) measures of scapular, humeral, and glenohumeral alignment.

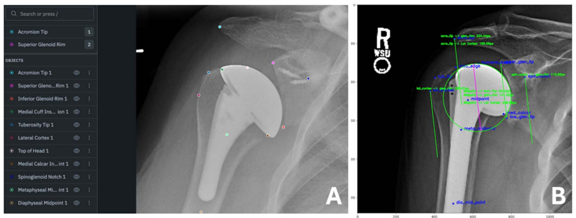

Table I. Anatomic Landmarks and Scapular, Humeral, and Glenohumeral Measurements

Anatomic Landmarks	**Deviation (mm)**
Cortical Landmarks	
Medial insertion of supraspinatus	2.5 ± 3.3
Superior aspect of humeral head	2.2 ± 1.4
Tip of acromion	1.4 ± 1.9
Tip of greater tuberosity	1.5 ± 1.1
Lateral cortex of greater tuberosity	1.9 ± 1.5
Medial calcar inflection point	2.5 ± 1.5
Superior glenoid rim	1.9 ± 1.7
Inferior glenoid rim	1.2 ± 0.7
Spinoglenoid notch	1.7 ± 2.2
Non-cortical Landmarks	
Metaphyseal midpoint	8.1 ± 4.4
Diaphyseal midpoint	10.3 ± 10.5
Scapular, Humeral, and Glenohumeral Measurements	**Deviation (mm)**
Medial humeral head (HH) – center of rotation (COR)	1.2 ± 1.1
Medial HH – humeral shaft axis	2.1 ± 2.3
Medial HH – lateral cortex	2.0 ± 1.4
COR – humeral shaft axis	1.7 ± 2.5
COR – lateral cortex (humeral plane)	2.4 ± 2.2
Humeral shaft axis – lateral cortex	2.6 ± 3.5
Humeral head height	2.9 ± 2.7
Glenoid rim – COR	3.0 ± 3.1
Glenoid rim – acromion	3.4 ± 4.8
Glenoid rim – lateral cortex	3.2 ± 4.8
Glenoid rim – spinoglenoid notch	2.4 ± 3.2
COR – acromion	8.9 ± 5.9
COR – lateral cortex (glenoid plane)	1.8 ± 1.7
Acromion – lateral cortex	3.5 ± 4.0