Artificial Intelligence for Automated Identification of Total Shoulder Arthroplasty Implants on Plain Radiographs

Kyle N Kunze[1], Seong Jun Jang, Tim Yiran Li, Anthony Finocchiaro, Michael Fu, Scott Alan Rodeo[2], Frank A Cordasco[2], Theodore A Blaine[3], Riley Joseph Williams[4], David W Altchek[2], Stephen Fealy[3], Robert G Marx[2], Samuel Arthur Taylor[5], Joshua S Dines[3], David M Dines[3], Russell F Warren[2], Lawrence V Gulotta[6]
[1]Orthopaedic Surgery, Hospital for Special Surgery, [2]Hosp for Special Surgery, [3]Hospital For Special Surgery, [4]Hospital For Special Surgery - Weill Cornell Med, [5]Hosp for Special Surgery-Cornell, [6]Hosp for Special Surg-Cornell

INTRODUCTION: Accurate and rapid identification of implant manufacturer and model is critical in the evaluation and management of complications surrounding patients that have undergone total shoulder arthroplasty (TSA) and those necessitating revisions or reoperations. Failure to correctly identify implant designs in these circumstances may lead to delay in care, unexpected intraoperative challenges, increased morbidity, and excess healthcare costs. Deep learning permits automated image processing and holds the potential to mitigate such challenges while improving value of care rendered. The purpose of the current study was to develop an automated deep learning algorithm to accurately and rapidly identify the manufacturer and model of shoulder arthroplasty implants from plain radiographs.

METHODS: A Resnet-34 base architecture with transfer learning was trained and validated to classify reverse (rTSA) and anatomic (aTSA) implants as one of 20 implant models from eight different implant manufacturers used between 2011-2021 by 26 fellowship-trained surgeons at two independent tertiary academic hospital in the Pacific Northwest and Mid-Atlantic Northeast. From these radiographs, 1,008 implants were used for training and 252 for independent validation. Model performance was assessed using standardized metrics including area under the receiver-operator characteristic curve (AUC), sensitivity, specificity, as compared with a reference standard of implant model data from operative reports.

RESULTS: A total of 100 training epochs enhanced with data augmentation applied during each individual cycle was found to optimize model performance. The software classified implants at a mean speed of 0.079 seconds per image by identifying key differentiating characteristics across implant designs (**Figure 1**). After training, the deep learning model discriminated between the 20 implant models with a mean AUROC of 0.999 (95% confidence interval: 0.957-1.000, **Figure 2**), accuracy of 97.9%, sensitivity of 89.5%, and specificity of 98.8% on the independent testing set of 252 radiographs not used for training (**Figure 3**).

DISCUSSION AND CONCLUSION:

A deep learning model demonstrated excellent accuracy in identifying 20 unique TSA implants from eight manufacturers. This algorithm may provide a clinically meaningful adjunct in assisting with preoperative planning for the failed TSA and allows for scalable expansion with additional radiographic data.