Do the AAOS Appropriate Use Criteria Match Treatment Recommendations from Arthroplasty Surgeons?

Adam I. Edelstein M.D., T. Jacob Selph Jr. B.S., Manasa Pagadala B.S., Rachel Bergman M.D., Abbie Bennett M.S., Patricia Franklin M.D. M.P.H. M.B.A., Linda I. Suleiman M.D.,
Department of Orthopaedic Surgery, Northwestern University, Chicago, IL

P136 AAOS Annual Meeting 2024

Introduction

The purpose of this study was to evaluate the accuracy of the American Academy of Orthopedic Surgeons (AAOS) Appropriate Use Criteria (AUC) *Osteoarthritis of the Knee: Surgical Management* and *Osteoarthritis of the Hip: Management* by examining the frequency with which AUC total knee arthroplasty (TKA) and total hip arthroplasty (THA) appropriateness ratings align with fellowship-trained arthroplasty surgeon recommendations.

Methods

AUC TKA and THA appropriateness classifications generated for 558 osteoarthritis patients (397 knee, 161 hip) presenting to our clinic were compared with surgeon total joint arthroplasty recommendation. SAS software (Cary, NC) was used for analysis. Strength of association between AUC and surgeon recommendation was assessed via Spearman's ρ. Level of AUC and surgeon recommendation agreement was quantified with Cohen's weighted Kappa (KW) using Fleiss-Cohen weighting. Values assigned to variables in calculations included in table below.

Assigned Value	AUC Classification	Surgery Recommendation
1	Rarely Appropriate	No
2	May Be Appropriate	-
3	Appropriate	Yes

Results

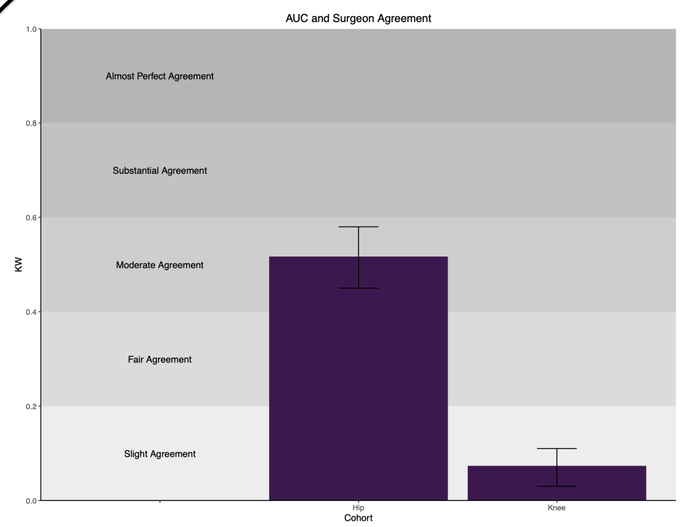

Figure 1: Weighted kappa plot demonstrating inter-rater agreement between AUC and surgeon for hip and knee cohorts. Hip AUC had moderate agreement (KW = 0.5) with surgeons and Knee AUC had slight agreement (KW = 0.07).

Cohort	ρ	95% Confidence Interval
Hip	0.61861	0.515 to 0.722
Knee	0.22641	0.15 to 0.302

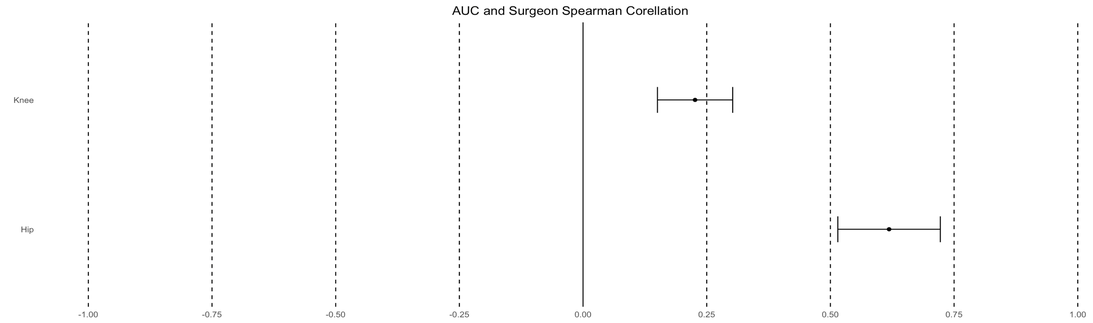

Figure 2: Forest plot demonstrating Spearman correlation of AUC appropriateness with surgeon recommendation for surgery. Hip AUC had moderate positive correlation and knee AUC had negligible positive correlation with surgeon recommendation for surgery. Both reported correlations had p<0.0001. Correlation was rated on a scale of 0 to ±0.3 (Negligible correlation), 0.3 to ±0.5 (Weak correlation), 0.5 to ±0.7 (Moderate correlation), 0.7 to ±0.9 (High correlation), 0.9 to ±1.0 (Very high correlation).

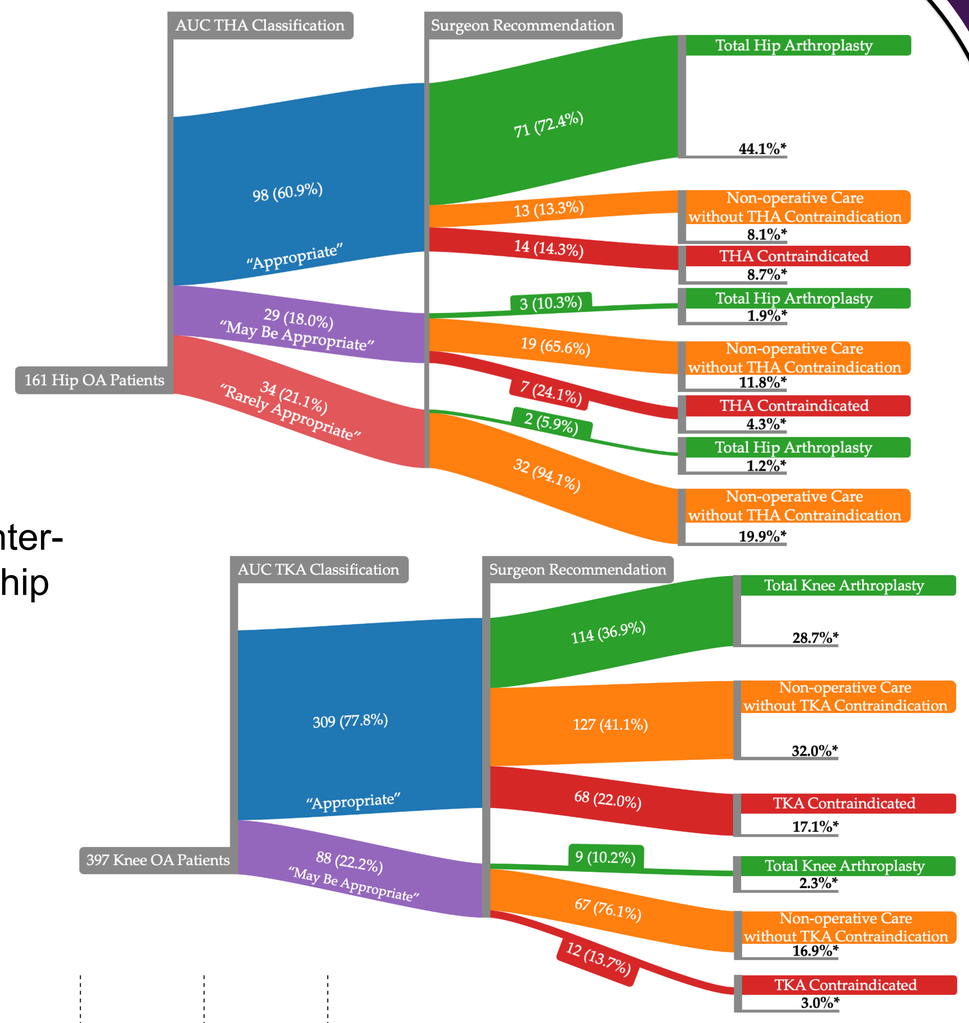

Figure 3: Sankey diagrams demonstrating the proportions of patients in each AUC appropriateness category falling into each treatment category. % is percent of patients originating from chart node. %* is percent of total knee/hip cohort.

Conclusion

In this patient population, AAOS AUC guidelines recommended surgery much more often than fellowship-trained arthroplasty surgeons. TKA AUC rating had negligible positive correlation and THA AUC had moderate positive correlation with surgeon decision to operate. AUC agreed only slightly with surgeons for TKA and moderately for THA. Providers making treatment recommendations should not rely on AUC appropriateness classification alone.

NORTHWESTERN
DEPARTMENT OF ORTHOPAEDIC SURGERY