Appropriateness and Total Hip Arthroplasty: Determining the Structure of the American Academy of Orthopaedic Surgeons System of Classification

Daniel L. Riddle ⓘ and Robert A. Perera ⓘ

ABSTRACT. *Objective.* In late 2017, the American Academy of Orthopaedic Surgeons (AAOS) published an appropriateness classification system using the RAND/University of California, Los Angeles (UCLA) approach for patients with hip osteoarthritis (OA). We determined the contribution of predictor variables in the system to final classification, rated as "appropriate," "may be appropriate," or "rarely appropriate" for hip arthroplasty.

Methods. An AAOS-appointed expert panel developed 270 clinical vignettes incorporating all permutations of 5 evidence-driven indication variables associated with hip arthroplasty outcome or need. Indication variables were age, function-limiting pain severity, radiographic hip OA severity, hip motion, and presence of modifiable prognostic risk factors. Multinomial regression determined the relative contribution of each variable and a classification tree method determined variable combinations contributing to final classification.

Results. Patient age and hip OA severity were the dominant predictors of appropriateness classification in both statistical models. Function-limiting pain made a slight contribution relative to age and hip OA severity while hip motion and the presence of modifiable prognostic factors did not meaningfully contribute to final classification. The regression model explained about 99% of the variance and the classification tree had an accuracy of 87.8%.

Conclusion. Classification for hip arthroplasty appropriateness in the AAOS system is driven almost exclusively by age and OA severity. Function-limiting pain, a major reason patients seek surgery, contributes only slightly to the AAOS appropriateness criteria. The system relies heavily on traditional variables of patient age and radiographic hip OA severity. Future study of actual patient outcomes is needed to further test the validity of the AAOS system. (J Rheumatol First Release March 1 2019; doi:10.3899/jrheum.180911)

Key Indexing Terms:
HIP   ARTHROPLASTY   PAIN   SURGERY   CLASSIFICATION   RISK FACTORS

The American Academy of Orthopaedic Surgeons (AAOS) has invested substantial effort toward the development of appropriate use criteria targeting a variety of musculoskeletal conditions, including knee and hip osteoarthritis (OA). These criteria define patient-level characteristics that can be used to determine treatments. Judgments are categorized as "appropriate," "may be appropriate," or "rarely appropriate" for a given condition. The RAND/University of California, Los Angeles (UCLA) appropriateness method (also referred to as the RAND system) was used by the AAOS to develop appropriateness use criteria[1]. Briefly, the RAND system is a consensus-based method that relies on a comprehensive review of prognostic evidence, and multiple Delphi-type surveys among multidisciplinary panels of clinical experts. In the AAOS, an expert panel identified key prognostic/predictor indicator variables from a comprehensive literature review. A set of ordinal categories was then developed for each prognostic/predictor variable (e.g., range-of-motion limitation: minimal, moderate, severe) and brief clinical vignettes were written covering all permutations of the levels of the prognostic/predictor variables. A second independent expert panel then rated each clinical vignette as "appropriate," "may be appropriate," or "rarely appropriate" for a given treatment using defined methods. The end product is an algorithm that defines appropriateness ratings for all combinations of key prognostic/predictor variables. The AAOS appropriate use criteria are designed to serve as decision aids for informing clinicians and patients about the extent of appropriateness of various orthopedic interventions.

We recently published an analysis of AAOS Appropriate Use Criteria for knee arthroplasty[2] and found that the system relied heavily on traditional variables of age, knee OA

severity, and knee OA pattern for appropriateness classification. Function-limiting pain, the primary reason patients seek out knee arthroplasty[3], made a minimal contribution in the AAOS system. In our current study, we applied a parallel analytic approach to recently developed AAOS Appropriate Use Criteria for hip arthroplasty. Because the AAOS hip OA appropriateness system is available worldwide to clinicians and the public through a no-cost app (www.orthoguidelines.org/go/auc), it is important to study the system to determine the contributions of the prognostic/predictor variables to final classification.

The purpose of our study was to determine the contribution of the 5 predictor variables (i.e., age, function-limiting pain, hip radiographic evaluation, range-of-motion limitation, presence or absence of modifiable risk factors) in predicting hip arthroplasty appropriateness classifications made by the AAOS voting panel. Based on our prior work on the AAOS knee arthroplasty appropriate use criteria, we hypothesized that hip arthroplasty classification would be highly reliant on historically traditional variables of age and OA severity and that function-limiting pain would contribute in only a minor or inconsequential way. Pain relief following hip arthroplasty is rated by patients as the most important reason[3] for seeking the procedure. Additionally, function-limiting pain ranks as the most important predictor of appropriateness in a commonly reported hip arthroplasty RAND-based appropriateness system[4].

MATERIALS AND METHODS

We obtained the full report entitled "Appropriate Use Criteria for the Management of Osteoarthritis of the Hip" from the AAOS Website (www.aaos.org). The report provided complete versions of all vignettes (n = 270) rated by an expert voting panel of 16 experts (13 orthopedic surgeons, 1 physical therapist, 1 radiologist, and 1 rheumatologist). We did not have direct interactions with any member of the voting panel. Rather, we relied on the full AAOS report to extract all data. Specifically, the investigators extracted the appropriateness ratings for each of the 270 vignettes scored in the final voting as "appropriate," "may be appropriate," or "rarely appropriate." The median rating for each vignette also was recorded and was rated on a scale from 1 (rarely appropriate) to 9 (appropriate). A median score of 5 was predesignated as a score that indicated disagreement among the expert panel, defined as ≥ 4 members' ratings falling between 1–3, and ≥ 4 members' ratings falling between 7–9 for a given vignette. A total of n = 33 vignettes were coded as "disagreement" among the expert voting panel. We also recorded whether the expert voting panel agreed as a group (≤ 3 voting panelists rated outside of the 3-point range containing the median score for a given vignette). A total of n = 95 vignettes were scored as agreement among the expert voting panel. All other vignettes (n = 142) were considered as falling in the middle between disagreement and agreement among the expert voting panel members.

We included data from all 270 vignettes. Scores for each of the 5 prognostic/predictor variables for each vignette were analyzed (Table 1). Three of the prognostic variables (age, hip motion, function-limiting pain) had trichotomous responses, one (risk of negative outcome) had dichotomous responses, and one (hip OA radiographic evaluation) had 5 response options. The variables were combined by AAOS using a factorial approach creating vignettes covering all permutations of the 5 prognostic variables [(51 × 33 × 21) = 270]. Importantly, the AAOS system used all 270 vignettes to generate appropriateness ratings for clinical application.

Data analysis. We used multinomial regression to determine the contribution of the 5 predictor variables to appropriateness ratings. We studied the entire population of vignettes and all predictors were categorical. As a result, p values and CI were not needed because both are typically used to make inferences to the population from estimates obtained from a sample. Because we included the entire population of vignettes, our results reflect the entire population and not just an estimate, precluding p values and CI. Coefficients from the regression were used to assess the importance of each predictor variable in determining appropriateness classification. Coefficients are directly comparable because all predictors were categorical. Additionally, collinearity cannot have an effect when the entire population is represented. Nagelkerke r^2 was used to estimate explained variance.

We also used a classification tree approach [exhaustive Chi-Square Automatic Interaction Detection (CHAID)] to determine the optimal combination of prognostic variables for predicting each of the appropriateness ratings. We used exhaustive CHAID to construct the tree because it allows for the examination of all possible splits of polytomous predictor variables (e.g., age, scored as young, middle-aged, and elderly), unlike Classification and Regression Trees, which require dichotomous splitting for each predictor. Settings allowed for up to 5 levels of branching with a minimum of 25 vignettes in a parent node and a minimum of 15 subjects in a terminal node. This nonparametric approach systematically tests each of the 5 predictor variables to determine which variable most strongly associates with appropriateness classification. Once the variable with the highest chi-square is found, the tree branch is split and the process is repeated. The goal was to find the purest terminal nodes for each branch of the tree while also considering parsimony. The strength of this approach is that it identifies the optimal combination of predictors for the "rarely appropriate," "may be appropriate," and "appropriate" classifications. Cross-validation was not necessary because we studied the entire population of vignettes. Weighted κ was used to judge the extent of agreement between the AAOS system classification and the classification tree. We used IBM SPSS, Version 24 for all analyses.

We also conducted a sensitivity analysis by repeating the multinomial regression and classification tree analysis after excluding the 33 vignettes coded as "disagreement" in the AAOS full report.

RESULTS

Multinomial regression. Predictor variables with the largest coefficients were radiographic evaluation and age (Table 2). Vignettes classified with severe or moderate radiographic OA or elderly/middle age increased the odds of being classified as appropriate by an order of magnitude much larger than other predictor variables in the model. For example, when examining the likelihood of being classified as appropriate for hip arthroplasty, the OR for severe OA is more than 10^{192} times larger than function-limiting pain at rest or at night. Regression coefficients and OR showed a similar pattern for comparisons between "appropriate" and "rarely appropriate" classifications, and "may be appropriate" and "rarely appropriate classifications." That is, the variables of hip radiographic evaluation and age had substantially higher OR than other variables for both comparisons. The Nagelkerke r^2 for the model = 0.99, indicating near-perfect explanation of appropriateness classification despite exclusion of interactions.

In a sensitivity analysis in which 33 vignettes scored as "disagreement" by the expert voting panel had been removed, the multinomial regression was almost identical to the original analysis (Table 3).

Classification tree. The accuracy of the classification tree for

Table 1. Characteristics of the predictive criteria and appropriateness ratings for the AAOS hip arthroplasty clinical scenarios (n = 270).

Predictive Criterion Measurement Scale	N	%
Age		
Young	90	33.3
Middle-aged	90	33.3
Elderly	90	33.3
Function-limiting pain		
Pain walking moderate/long distances (e.g., > ¼ mile)	90	33.3
Pain walking short distances (e.g., ~ 2 city blocks)	90	33.3
Pain at rest/night	90	33.3
Radiographic evaluation of the hip		
Minimal osteoarthritis	54	20
Minimal osteoarthritis with acetabular dysplasia	54	20
Minimal osteoarthritis with femoroacetabular impingement	54	20
Moderate osteoarthritis	54	20
Severe osteoarthritis	54	20
Hip range-of-motion limitation		
Minimal	90	33.3
Moderate	90	33.3
Severe	90	33.3
Presence or absence of risk factors for negative outcome		
Modifiable risk factors present	135	50
No modifiable risk factors present	135	50
Total hip arthroplasty classification		
Appropriate	53	19.6
May be appropriate	97	35.9
Rarely appropriate	120	44.4
Expert panel voting results[*]		
Agreement	95	35.2
Disagreement	33	12.2
Neither agreement not disagreement	142	52.6

* Agreement indicated that ≤ 3 voting panelists rated outside of the 3-point range containing the median score for a given vignette. Disagreement was defined as ≥ 4 members' ratings falling between 1 and 3 and ≥ 4 members' ratings falling between 7 and 9 for a given vignette. Neither agreement nor disagreement occurred for other vignettes that fell between agreement and disagreement. AAOS: American Academy of Orthopaedic Surgeons.

correctly identifying AAOS appropriateness classifications was 87.8% (84.9% for "appropriate," 79.4% for "may be appropriate," and 95.8% for "rarely appropriate" ratings). The extent of agreement between the classification tree and AAOS classifications was weighted κ = 0.86, indicating almost-perfect agreement[5]. The most powerful predictor (i.e., most proximal in the tree) was radiographic hip OA severity evaluation and the next strongest predictor was age (Figure 1). Function-limiting pain was the final variable that entered the tree (see terminal nodes 12 and 13, Figure 1).

The terminal nodes for each branch of the tree are labeled as nodes 6 through 13 (Figure 1). Terminal node 10 is a pure node, indicating there was no disagreement from the expert panel for the 54 vignettes classified as "rarely appropriate." Patients in vignettes in node 10 had minor radiographic hip OA and were < 40 years of age. In contrast, terminal node 8 is an example of a mixed terminal node, with most vignettes classified as "may be appropriate" for hip arthroplasty. The patients in these vignettes had severe radiographic hip OA and were aged < 40 years. The classification tree in our sensi-

tivity analysis (Figure 2) was very similar to the original regression tree. Differences at node 2 between the 2 analyses were likely due to a smaller sample size in the sensitivity analysis.

DISCUSSION

Lower extremity joint arthroplasty has substantially increased[6], which has led to strong interest in developing arthroplasty indication criteria[7,8]. Much like the knee arthroplasty appropriateness system developed by the AAOS, the hip arthroplasty system was based on a recently completed evidence synthesis[9]. Incorporation of new evidence in the appropriateness classification development process is a strength given that the most currently available RAND-based hip arthroplasty appropriateness system was developed in the late 1990s[4]. Despite the use of current evidence, we found that the hip system developed by AAOS relies almost exclusively on traditional hip arthroplasty indicators of hip OA severity and age to drive appropriateness. Function-limiting hip pain severity was a minor predictor in our multinomial

Table 2. Multinomial regression analysis of hip data from 270 vignettes with "rarely appropriate" as the reference category.

Classification	AAOS Prognostic/Predictor Variable	Rating for Each Prognostic/Predictor Variable	B	OR
Appropriate	Intercept		−1044.27	
	Function-limiting pain	Pain at rest or night	269.14	7.66×10^{116}
		Short distances	211.96	1.13×10^{92}
		Moderate to long distances	0[a]	
	Range-of-motion limitation	Severe	85.61	1.51×10^{37}
		Moderate	52.00	3.86×10^{22}
		Minimal	0[a]	
	Risk of negative outcome	No modifiable risk factors	153.92	7.05×10^{66}
		Modifiable risk factors	0[a]	
	Hip radiographic evaluation	Severe OA[b]	723.67	$> 1.60 \times 10^{308}$
		Moderate OA	327.61	1.90×10^{142}
		Minimal OA with FAI	3.71	40.66
		Minimal OA with AD	8.80	6624.08
		Minimal OA	0[a]	
	Age	Elderly	541.11	1.00×10^{235}
		Middle-aged	326.87	9.08×10^{141}
		Young	0[a]	
May be appropriate	Intercept		−82.54	
	Function-limiting pain	Pain at rest or night	24.97	7.00×10^{10}
		Short distances	28.56	2.54×10^{12}
		Moderate to long distances	0[a]	
	Range-of-motion limitation	Severe	25.70	1.45×10^{11}
		Moderate	21.24	1.68×10^{9}
		Minimal	0[a]	
	Risk of negative outcome	No modifiable risk factors	1.97	7.15
		Modifiable risk factors	0[a]	
	Hip radiographic evaluation	Severe OA	81.56	2.64×10^{35}
		Moderate OA	81.56	2.64×10^{35}
		Minimal OA with FAI	3.69	40.08
		Minimal OA with AD	8.76	6402.33
		Minimal OA	0[a]	
	Age	Elderly	54.04	2.94×10^{23}
		Middle-aged	22.81	8.04×10^{9}
		Young	0[a]	

[a] Parameter fixed to zero because it is redundant. [b] Computing the exponent of the coefficient for severe OA in the "appropriate" relative to "rarely appropriate" classification exceeds computational limits, producing a value of infinity. Therefore, a lower value than the coefficient was used to compute the reported OR. OA: osteoarthritis; AAOS: American Academy of Orthopaedic Surgeons; FAI: femoroacetabular impingement; AD: acetabular dysplasia.

regression and also played a minor role in the classification tree. Given the high importance of function-limiting pain to patients and surgeons[3,10], the minimal relevance of this variable in driving appropriateness classifications is concerning.

The most powerful predictor was radiographic hip OA severity. When considering only hip OA severity when classifying appropriateness, if the vignette indicated the candidate had moderate or severe radiographic hip OA, 106/108 vignettes were judged to be either "appropriate" (n = 53) or "may be appropriate" (n = 53) for hip arthroplasty. Reliance on the presence of moderate to severe radiographic hip OA severity without other data to inform an appropriateness decision is, in our view, a substantial limitation of the system. For example, some persons with moderate or severe radiographic hip OA have either no pain or mild pain[11,12]. These persons would likely experience minimal or no benefit from hip arthroplasty while also being exposed substantial

cost and time loss as well as a risk, albeit a low risk, of serious adverse outcomes and substantial costs.

The minimal presence of function-limiting pain as a predictor of appropriateness in either the multinomial regression or the classification tree was expected given our findings from a prior study of knee arthroplasty[2]. The function-limiting pain variable improved classifications of "may be appropriate" and "rarely appropriate" for vignettes with minimal radiographic OA and age > 65 years (n = 54 vignettes), but this influence was, in our opinion, very small (Figure 1). It appears that the expert panel placed minimal emphasis on function-limiting pain when judging appropriateness, much like the AAOS knee arthroplasty appropriateness system[2]. Other evidence-based predictors used by the panel (i.e., hip motion limitation and the presence of modifiable prognostic variables) also did not influence classification in a meaningful way.

Some improvements were noted in the hip appropriateness

Table 3. Multinomial regression sensitivity analysis of hip data after n = 33 vignettes scored as "disagreement" by the expert voting panel have been removed. "Rarely appropriate" is the reference category.

Classification	AAOS Prognostic/Predictor Variable	Rating for Each Prognostic/Predictor Variable	B	OR
Appropriate	Intercept		–987.51	
	Function-limiting pain	Pain at rest or night	270.20	2.21×10^{117}
		Short distances	212.73	2.43×10^{92}
		Moderate to long distances	0[a]	
	Range-of-motion limitation	Severe	85.84	1.89×10^{37}
		Moderate	52.12	4.28×10^{22}
		Minimal	0[a]	
	Risk of negative outcome	No modifiable risk factors	154.53	1.30×10^{67}
		Modifiable risk factors	0[a]	
	Hip radiographic evaluation	Severe OA[b]	665.48	1.03×10^{289}
		Moderate OA	267.66	1.75×10^{116}
		Minimal OA with FAI	2.74	15.43
		Minimal OA with AD	–52.08	2.42^{-23}
		Minimal OA	0[a]	
	Age	Elderly	543.03	6.83×10^{235}
		Middle-aged	328.15	3.27×10^{142}
		Young	0[a]	
May be appropriate	Intercept		–82.29	
	Function-limiting pain	Pain at rest or night	24.89	6.46×10^{10}
		Short distances	28.47	2.31×10^{12}
		Moderate to long distances	0[a]	
	Range-of-motion limitation	Severe	25.63	1.35×10^{11}
		Moderate	21.20	1.60×10^{9}
		Minimal	0[a]	
	Risk of negative outcome	No modifiable risk factors	1.96	7.09
		Modifiable risk factors	0[a]	
	Hip radiographic evaluation	Severe OA	81.31	2.05×10^{35}
		Moderate OA	81.31	2.05×10^{35}
		Minimal OA with FAI	3.67	39.24
		Minimal OA with AD	8.71	6079.75
		Minimal OA	0[a]	
	Age	Elderly	53.87	2.49×10^{23}
		Middle-aged	22.75	7.59×10^{9}
		Young	0[a]	

[a] Parameter fixed to zero because it is redundant. [b] Computing the exponent of the coefficient for severe OA in the "appropriate" relative to "rarely appropriate" classification exceeds computational limits, producing a value of infinity. Therefore, a lower value than the coefficient was used to compute the reported OR. OA: osteoarthritis; AAOS: American Academy of Orthopaedic Surgeons; FAI: femoroacetabular impingement; AD: acetabular dysplasia.

system compared to the knee appropriateness system. For example, actual age ranges for the vignettes were provided in the hip system but not in the knee system and 3 non-surgeons served on the hip expert panel as compared to only 1 non-surgeon in the knee system. RAND recommends a diverse multidisciplinary panel to reduce bias risk. However, hip pain distribution and the presence of posttraumatic hip OA was not addressed, which is a limitation. Despite some improvements, our studies suggest an over-reliance on historically traditional predictors of age and OA severity and a lack of meaningful effect for function-limiting pain severity or for other contemporary evidence-driven prognostic measures such as psychological distress[13] or risks such as 30-day hospital readmission[14], both of which could affect the decision to undergo hip arthroplasty.

One limitation of our study is that we examined only the hip arthroplasty appropriateness classification system. The AAOS hip OA appropriate use criteria address a variety of other treatment decisions including the use of physical therapy and hip preservation surgery, among others[15]. These other treatment decisions addressed in the AAOS criteria may have greater utility than the hip arthroplasty criteria.

We found that the AAOS hip arthroplasty appropriateness classification system appears to be driven almost exclusively by age and radiographic hip OA severity and is therefore substantially limited and not likely to improve patient care, though further validation testing using actual patient outcome data is needed to ultimately judge the usefulness of the system. Assessment with actual patient data would allow for determination of both journey (i.e., change over time) and destination outcomes (i.e., outcome at a defined point in time)[16]. It does not appear that newer evidence had a meaningful effect on classification, and function-limiting pain plays a minor, and likely inconsequential, role. These
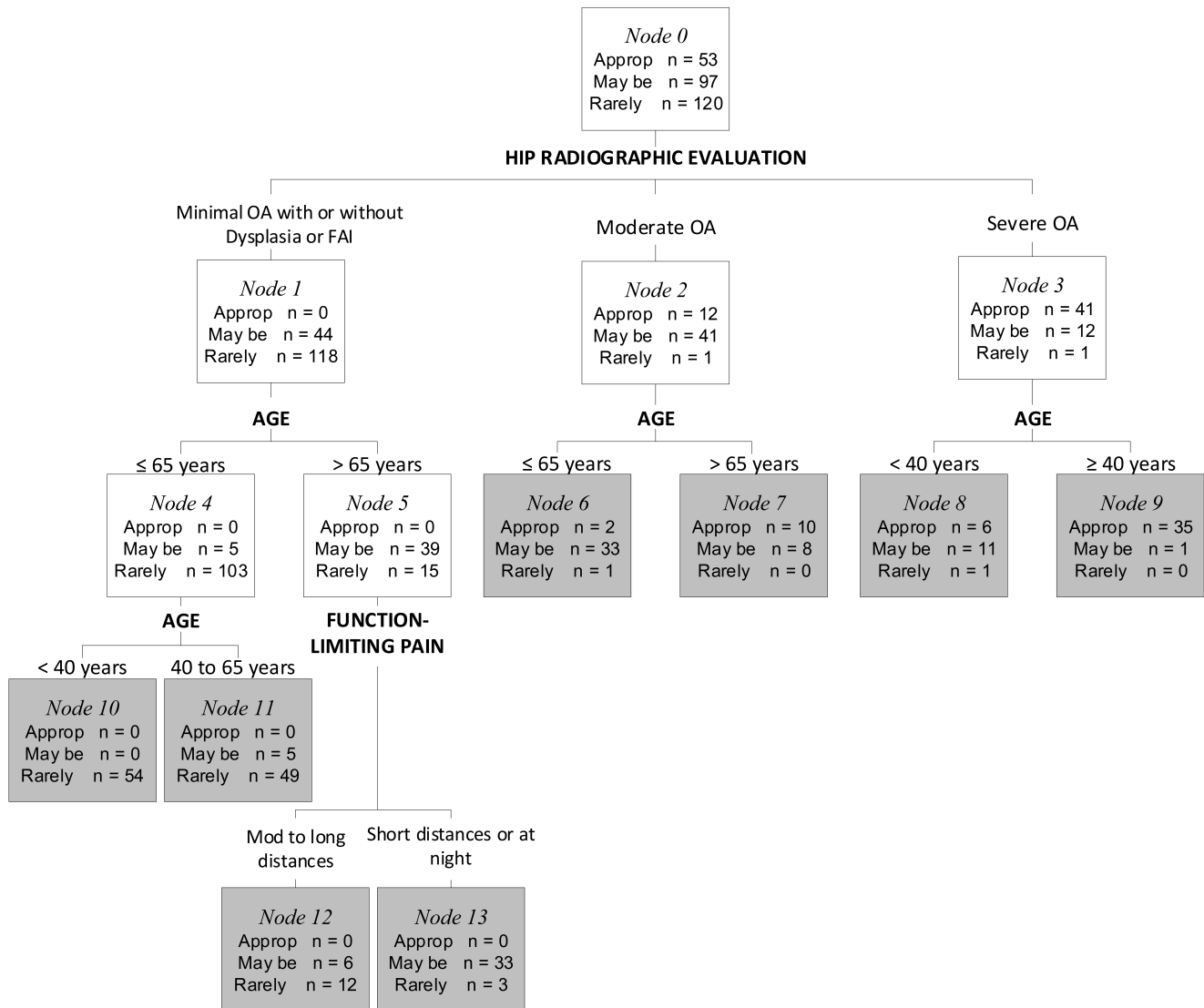
Figure 1. Classification tree. The branches of the tree are labeled based on the key variables that discriminated among the classifications and these are listed as hip radiographic evaluation, age, and function-limiting pain. The terminal nodes of each branch (nodes 6–13) indicate the final distributions of ratings of appropriate (Approp), may be appropriate (May be), and rarely appropriate (Rarely). Vignette sample sizes are reported in each box. FAI: femoroacetabular impingement; Mod to long: moderate to long. OA: osteoarthritis.

findings have substantial importance because the AAOS hip appropriateness classification system is freely available worldwide to both patients and clinicians.

REFERENCES

1. Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA appropriateness method user's manual [Internet. Accessed January 25, 2019]. Available from: www.rand.org/pubs/monograph_reports/MR1269.html
2. Riddle DL, Perera RA. Appropriateness and total knee arthroplasty: an examination of the American Academy of Orthopaedic Surgeons appropriateness rating system. Osteoarthritis Cartilage 2017;25:1994-8.
3. Hawker G, Bohm ER, Conner-Spady B, De CC, Dunbar M, Hennigar A, et al. Perspectives of Canadian stakeholders on criteria for appropriateness for total joint arthroplasty in patients with hip and knee osteoarthritis. Arthritis Rheumatol 2015;67:1806-15.
4. Quintana JM, Arostegui I, Azkarate J, Goenaga JI, Guisasola I, Alfageme A, et al. Evaluation by explicit criteria of the use of total hip joint replacement. Rheumatology 2000;39:1234-41.
5. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977;33:159-74.
6. Kurtz SM, Ong KL, Lau E, Bozic KJ. Impact of the economic downturn on total joint replacement demand in the United States: updated projections to 2021. J Bone Joint Surg Am 2014;96:624-30.
7. Katz JN, Winter AR, Hawker G. Measures of the appropriateness of elective orthopaedic joint and spine procedures. J Bone Joint Surg Am 2017 15;99:e15.
8. Riddle DL, Ghomrawi H, Jiranek WA, Dumenci L, Perera RA, Escobar A. Appropriateness criteria for total knee arthroplasty: additional comments and considerations. J Bone Joint Surg Am 2018;100:e22.
9. American Academy of Orthopaedic Surgeons. Management of

Node 0
Approp   n = 53
May be   n = 64
Rarely   n = 120

HIP RADIOGRAPHIC EVALUATION

Minimal OA with or without Dysplasia or FAI

Node 1
Approp   n = 0
May be   n = 28
Rarely   n = 118

Moderate OA

Node 2
Approp   n = 12
May be   n = 24
Rarely   n = 1

Severe OA

Node 3
Approp   n = 41
May be   n = 12
Rarely   n = 1

AGE

≤ 65 years

Node 4
Approp   n = 0
May be   n = 5
Rarely   n = 103

> 65 years

Node 5
Approp   n = 0
May be   n = 23
Rarely   n = 15

AGE

< 40 years

Node 8
Approp   n = 6
May be   n = 11
Rarely   n = 1

≥ 40 years

Node 9
Approp   n = 35
May be   n = 1
Rarely   n = 0

AGE

< 40 years

Node 10
Approp   n = 0
May be   n = 0
Rarely   n = 54

40 to 65 years

Node 11
Approp   n = 0
May be   n = 5
Rarely   n = 49

FUNCTION-LIMITING PAIN

Mod to long distances

Node 12
Approp   n = 0
May be   n = 6
Rarely   n = 12

Short distances or at night

Node 13
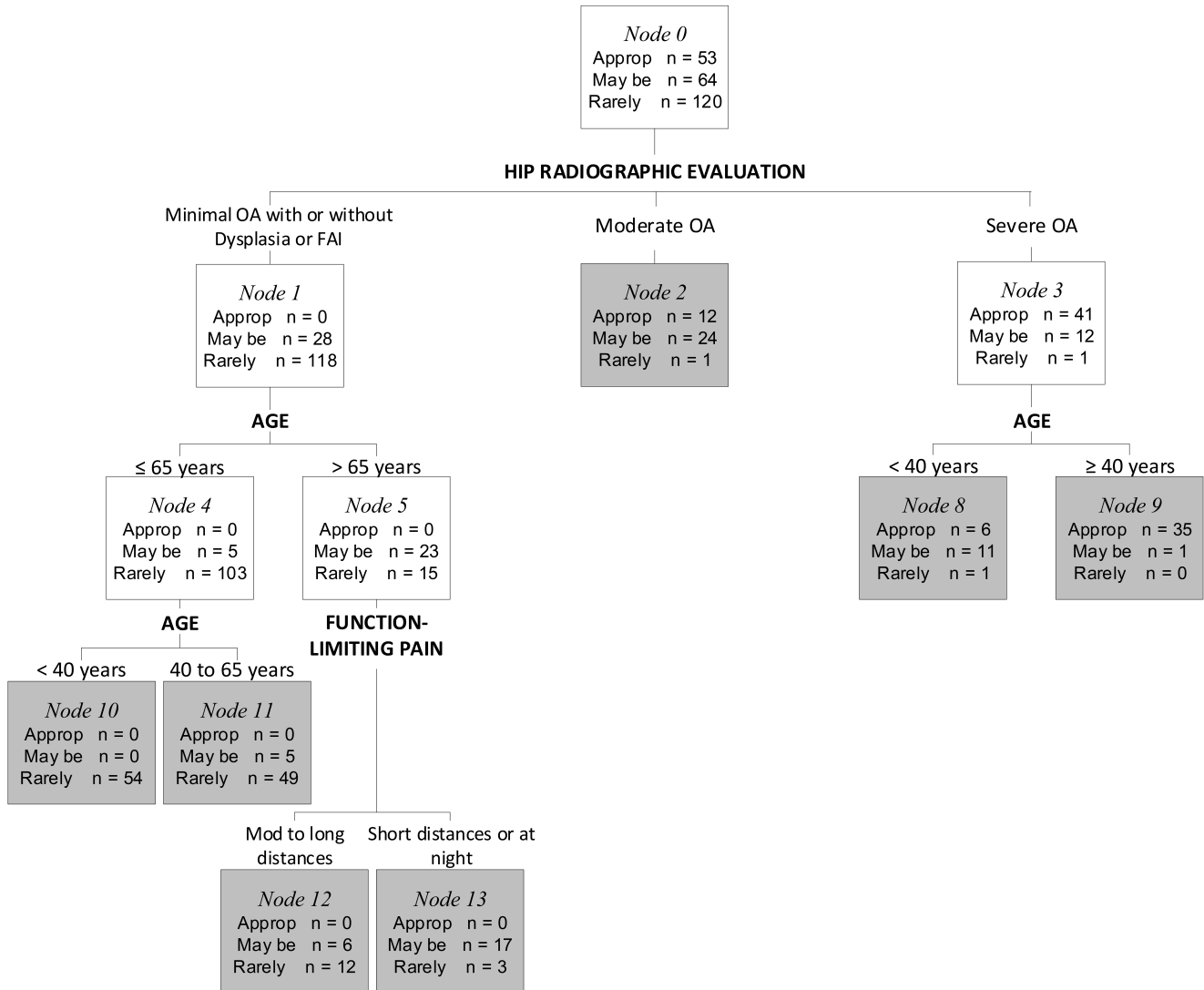Approp   n = 0
May be   n = 17
Rarely   n = 3

Figure 2. Classification tree for the sensitivity analysis. The branches of the tree are labeled based on the key variables that discriminated among the classifications: hip radiographic evaluation, age, and function-limiting pain. The terminal nodes of each branch (shaded in grey) indicate the final distributions of ratings of appropriate (Approp), may be appropriate (May be), and rarely appropriate (Rarely). Vignette sample sizes are reported in each box. FAI: femoroacetabular impingement; Mod to long: moderate to long; OA: osteoarthritis.

osteoarthritis of the hip. Evidence-based clinical practice guideline [Internet. Accessed February 1, 2019]. Available from: www.aaos.org/uploadedFiles/PreProduction/Quality/Guidelines_and_Reviews/OA%20Hip%20CPG_3.13.17.pdf

10. Hawker GA, Stewart L, French MR, Cibere J, Jordan JM, March L, et al. Understanding the pain experience in hip and knee osteoarthritis—an OARSI/OMERACT initiative. Osteoarthritis Cartilage 2008;16:415-22.

11. Birrell F, Lunt M, Macfarlane G, Silman A. Association between pain in the hip region and radiographic changes of osteoarthritis: results from a population-based study. Rheumatology 2005; 44:337-41.

12. Iidaka T, Muraki S, Akune T, Oka H, Kodama R, Tanaka S, et al. Prevalence of radiographic hip osteoarthritis and its association with hip pain in Japanese men and women: the ROAD study. Osteoarthritis Cartilage 2016;24:117-23.

13. Duivenvoorden T, Vissers MM, Verhaar JA, Busschbach JJ, Gosens T, Bloem RM, et al. Anxiety and depressive symptoms before and after total hip and knee arthroplasty: a prospective multicentre study. Osteoarthritis Cartilage 2013;21:1834-40.

14. Siracuse BL, Chamberlain R. Preoperative scale for determining surgical readmission risk after total hip replacement. JAMA Surg 2016;151:701-9.

15. American Academy of Orthopaedic Surgeons. Appropriate use criteria for the management of osteoarthritis of the hip. [Internet. Accessed January 25, 2019.] Available from: www.aaos.org/uploadedFiles/PreProduction/Quality/AUCs_and_Performance_Measures/appropriate_use/OA Hip AUC.pdf

16. Katz JN. Editorial: appropriateness of total knee arthroplasty. Arthritis Rheumatol 2014;66:1979-81.