The Role of Arthroscopy in the Management of Glenohumeral Osteoarthritis: A Markov Decision Model

Ulrich J. Spiegl, M.D., Scott C. Faucett, M.D., Marilee P. Horan, M.P.H., Ryan J. Warth, M.D., and Peter J. Millett, M.D., M.Sc.

Purpose: The purposes of this study were (1) to construct a theoretical Markov decision model to compare the total remaining quality-adjusted life-years following either arthroscopic management (AM) or total shoulder arthroplasty (TSA) for the treatment of glenohumeral osteoarthritis and (2) to determine the possible effects of age on the preferred treatment strategy. Methods: A Markov decision model was constructed to compare AM and TSA in patients with glenohumeral osteoarthritis. The rates of surgical complications, revision surgery, and death were derived from the literature and analyzed. The principal outcome measure was the mean total remaining quality-adjusted life-years after each treatment strategy. Sensitivity analyses were performed for age at the initial procedure, utilities, and transition probabilities. Results: This theoretical decision model showed that AM was the preferred strategy for patients younger than 47 years, TSA was the preferred strategy for patients older than 66 years, and both treatment strategies were reasonable for patients aged between 47 and 66 years. The model was highly sensitive to age at the index surgery, utilities of wellness states, survivorship, and the probability of failure after either AM or TSA. Conclusions: According to this theoretical decision model, AM was the preferred treatment strategy for patients younger than 47 years, primary TSA was the preferred treatment strategy for patients older than 66 years, and both treatment options were reasonable for patients aged between 47 and 66 years. Level of Evidence: Level II, economic and decision analysis.

For most patients with advanced glenohumeral osteoarthritis (OA), primary total shoulder arthroplasty (TSA) provides the most predictable outcome. Over the past decade, the role of primary TSA in young, active patients with glenohumeral OA has been questioned because of high rates of early glenoid component loosening that requires subsequent revision TSA.1-4 As a result, arthroscopic and other non-arthroplasty alternatives have been used as joint-preservation strategies. Such approaches can serve as bridging procedures with palliation of pain and delay the need for primary TSA in younger patients. However, comparisons of the potential benefits between arthroscopic management (AM) and primary TSA have not been performed. Therefore the purposes of this study were (1) to construct a theoretical decision model based on published literature to compare the total remaining quality-adjusted life-years (QALYs) between AM and TSA treatment strategies and (2) to determine the effects of age on the preferred treatment strategy. We hypothesized that AM would provide an increase in the total remaining QALYs when compared with TSA for patients younger than 50 years.

Methods

Markov Decision Process

In the context of this study, the Markov decision process uses a mathematical model to simulate the effects of different treatment options on clinical outcomes in terms of QALYs. To calculate QALYs, a decision tree is constructed and literature-derived probabilities for
each treatment outcome after the initial decision (in this case, AM v TSA) are combined (Fig 1). In other words, when a theoretical patient is treated using either method, there exists a probability for each possible outcome (e.g., revision AM, TSA, or death) for each year (or cycle) after the index procedure. At the end of each year (or cycle), theoretical patients are reassigned to a particular outcome (TSA, revision TSA, and so on) according to the calculated probability of each possible outcome. Each patient is also reassigned a health state (or utility) ranging from perfect health (1.0) to death (0.0) according to the treatment outcome of the previous year (or cycle). Treatment disutilities represent a deduction in health state, which reflects the probability of an adverse outcome (such as a complication or revision surgery). For each treatment strategy (AM or TSA), the summation of utilities and disutilities from the beginning of the model to a point when all patients have theoretically died represents the total remaining QALYs. Using this method, we aimed to model the effect of age on the total remaining QALYs in 2 cohorts of theoretical patients with glenohumeral OA who would undergo either AM or TSA. The preferred treatment strategy was defined as the treatment strategy (AM or TSA) that provided the highest calculated overall remaining QALYs.

**Transition Probabilities**

Transition probability was defined as the probability that a theoretical patient would move from one health state to another. Subsequently, the transition probabilities were used to determine the utility or disutility of

Fig 1. Health state diagram (i.e., decision tree) through which theoretical patients with glenohumeral OA were cycled. The distinct health states in the model were as follows: wellness after TSA or AM, revision TSA or re-revision TSA, and death. The health states at the end of the decision tree represent the new distribution of patients for the next cycle and vary according to the assigned transition probabilities.
each treatment option each time the patient was cycled through the model until theoretical death. In this study, transition probabilities, long-term outcomes, revision rates, and failure rates were derived from the literature, and when necessary, estimates were made to conform to the constraints of the model.\(^1\)\(^{1-22}\)

Long-term outcomes after primary TSA were estimated from multiple studies.\(^3\)^\(^{10,11,22}\) In this study, TSA with cemented glenoid components served as the reference for our model due to its survivorship when compared with procedures that involved non-cemented glenoid components.\(^12\)-\(^{15}\) The rates of failure, revision, and re-revision after primary TSA were determined from the literature and converted to yearly failure rates.\(^3\),\(^10,16-20\) Yearly failure rates for primary TSA after the first 10 years of implant survival were adjusted to approximately 50% of baseline based on the results presented by Cheung et al.\(^1\)

Long-term outcomes after AM were estimated based on several Level IV studies that reported 1- and 2-year survivorship rates between 85% and 92%,\(^5\)-\(^7\),\(^9\) Published failure rates after AM have been reported to range between 10% and 22% up to 34 months after the index surgical procedure.\(^7\)-\(^9\) These values corresponded with a yearly failure rate of approximately 8%. Therefore we conservatively estimated that the yearly transition probability for conversion to TSA after AM was approximately 10% at baseline.

**Mortality Rates**

Life expectancy and all-cause mortality rates were obtained from published life tables from the US National Center for Health Statistics.\(^23\) The perioperative mortality rate after AM in a young and active population was assumed to be zero. On the basis of available data, we assigned the probability of perioperative death after primary TSA, revision TSA, and re-revision TSA to 0.002, 0.004, and 0.008, respectively.\(^18\),\(^21,22\)

**Utilities**

A treatment utility was defined as a net-positive quantitative effect of a treatment option on the overall QALYs at the end of each cycle.

Treatment utility values for primary TSA were assigned based on the data presented by Mather et al.\(^22\) In their study, utility values for TSA and revision TSA were estimated to be 0.9 and 0.8, respectively. However, in our study we chose to reduce the estimated utility of TSA and revision TSA to 0.85 and 0.75, respectively, to reflect the reported inferior outcomes and poor satisfaction associated with TSA in a younger population.\(^1\)-\(^3\),\(^10,12,17,19,24-27\) In addition, an estimated utility value of 0.70 was assigned for re-revision TSA based on the continued functional deterioration in these patients when compared with patients treated with revision TSA.\(^27,28\)

The treatment utility value for AM was determined based on a comparison of frequency-weighted mean American Shoulder and Elbow Surgeons’ scores, degree of pain relief, and patient satisfaction between the AM\(^5\),\(^7\),\(^9\) and TSA\(^3\),\(^10,16,29-31\) treatment strategies. As a result of this comparison, we assigned a baseline treatment utility value of 0.82 for AM, slightly less than that of primary TSA.

**Disutilities**

A treatment disutility was defined as a 1-time net-negative quantitative effect of a treatment strategy on the overall QALYs at the end of each cycle as a result of reported complication rates and the diminished quality of life that generally occurs within the early postoperative period.\(^32\)

Treatment disutility for TSA was calculated based on a comparison of published outcomes between TSA and other joint replacement operations\(^33\),\(^34\) and the disutility values presented by Mather et al.\(^22\) In our study, primary TSA was assigned a disutility of −0.1, revision TSA was assigned a disutility of −0.2, and re-revision TSA was assigned a disutility of −0.25.

Treatment disutility for AM was determined based on anecdotal evidence provided by the senior author that patients appear to recover their shoulder function more quickly after arthroscopic rather than open procedures. As a result, AM was assigned a disutility value of −0.05, half of that which was assigned for primary TSA.

**Sensitivity Analyses**

The assigned parameters of our Markov decision model (transition probabilities, mortality rates, and utilities and disutilities) underwent both 1- and 2-way sensitivity analyses. The purpose of these analyses was to determine the effects of small changes in these assigned parameters on the overall remaining QALYs and thus the preferred treatment strategy. Although 1-way sensitivity analyses are capable of varying a single parameter (e.g., the assigned probability of revision TSA after primary TSA), 2-way sensitivity analyses are capable of varying 2 parameters simultaneously (e.g., patient age and the assigned probability of TSA after AM) to determine their overall effect on the remaining QALYs. By use of this analytical approach, the effect of age on the overall remaining QALYs was determined according to the initial treatment approach (AM or TSA).

**Results**

**Clinical Outcomes**

The baseline treatment utilities for the AM and TSA strategies were 0.82 QALYs per cycle and 0.85 QALYs per cycle, respectively. Baseline analysis showed that AM was the preferred strategy for patients aged up to...
When the failure rate within the 0.85), then TSA would have been the preferred treatment for TSA had been greater than 0.874 QALYs (baseline was between these 2 age groups. Similarly, if the treatment utility for AM was the preferred treatment strategy. For example, if the change in the overall remaining QALYs and, potentially, the probability of revision TSA after initial AM was calculated to be 45% whereas the opposite would be true for variables with a wider range of sensitivity values. Utilities and disutilities were defined as net-positive and net-negative effects of each treatment strategy, respectively, on the overall QALYs at the end of each cycle. Transition probabilities represent the probability that a theoretical patient would move between health states (e.g., the probability of treatment failure). Sensitivity values for each associated variable were also calculated: A variable would produce large changes in the resulting QALYs (high sensitivity), whereas the opposite would be true for variables with a smaller range indicated that even minor changes in the associated probabilities) was assigned a baseline value at the beginning of each cycle. Transition probabilities

| Markov Model | Baseline | Sensitivity
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilities*</td>
<td>OA</td>
<td>0.6</td>
</tr>
<tr>
<td>AM</td>
<td>0.82</td>
<td>0.7 to 1.0</td>
</tr>
<tr>
<td>TSA</td>
<td>0.85</td>
<td>0.7 to 1.0</td>
</tr>
<tr>
<td>Revision TSA</td>
<td>0.75</td>
<td>0.64 to 0.95</td>
</tr>
<tr>
<td>Re-revision TSA</td>
<td>0.7</td>
<td>0.5 to 0.9</td>
</tr>
</tbody>
</table>

Disutilities* AM | -0.05 | 0 to 0.1 | 5, 7, 9, 40 |
| TSA | -0.1 | -0.2 to 0 | 22, 27 |
| Revision TSA | -0.2 | -0.4 to 0 | 12, 27 |
| Re-revision TSA | -0.25 | 0.3 | 27, 39 |

| Transition probabilities | Treatment failure | AM | 10.0% | 7% to 13% | 5, 7-9, 40 |
| | TSA | 1 to 11 yr | 1.8% | 1% to 4% | 3, 10, 16 |
| | >11 yr | 0.9% | 1% to 2% | 3 |
| | Revision TSA | 1 to 11 yr | 4.6% | 3% to 7% | 19, 20 |
| | >11 yr | 2.3% | 1% to 4% | 11, 22 |

Mortality AM | 0% | 0% to 0.05% | 5, 7-9, 40 |
| TSA | 0.2% | 0% to 0.5% | 18, 21 |
| Revision TSA | 0.5% | 0% to 0.5% | 21 |
| Re-revision TSA | 1.0% | 0% to 2.0% | 22 |

Age* 40 yr | 0.003 | — | 23 |
| 50 yr | 0.006 | — | 23 |
| 60 yr | 0.013 | — | 23 |
| 70 yr | 0.032 | — | 23 |
| 80 yr | 0.087 | — | 23 |
| 90 yr | 0.226 | — | 23 |
| 100 yr | 1.000 | — | 23 |

NOTE. Each parameter (utilities, disutilities, and transition probabilities) was assigned a baseline value at the beginning of each cycle. Sensitivity values for each associated variable were also calculated: A smaller range indicated that even minor changes in the associated variable would produce large changes in the resulting QALYs (high sensitivity), whereas the opposite would be true for variables with a wider range of sensitivity values. Utilities and disutilities were defined as net-positive and net-negative effects of each treatment strategy, respectively, on the overall QALYs at the end of each cycle. Transition probabilities represent the probability that a theoretical patient would move between health states (e.g., the probability of treatment failure and subsequent TSA after AM was approximately 10.0% in this model). Published sources are displayed as a Reference number.

*Data are presented in QALYs.

Table 1. Summary of Literature-derived Parameters Used for Markov Decision Model

Sensitivity Analyses

Our decision model was very sensitive to age at initial surgery, the probability of treatment failure for both the AM and TSA strategies, and the treatment utilities after AM, TSA, and revision TSA (Table 1). In other words, small changes in the values assigned for these parameters would result in a large change in the overall remaining QALYs and, potentially, the preferred treatment strategy. For example, if the treatment utility for AM had been assigned a value of 0.81 QALYs rather than 0.82 QALYs, then age threshold at which AM would no longer be the preferred strategy would have dropped from 66 to 47 years. On the basis of this finding and the results of baseline analysis, the model suggested that AM was the preferred treatment strategy for patients younger than 47 years, TSA was the preferred treatment for patients older than 66 years, and both treatment options were reasonable for patients aged between these 2 age groups. Similarly, if the treatment utility for TSA had been greater than 0.874 QALYs (baseline was 0.85), then TSA would have been the preferred treatment strategy for all ages (Fig 3). AM was the preferred strategy when the failure rate within the first 10 years after TSA was greater than 1.5% (baseline was 4.6%) (Fig 4). In a 47-year-old patient, the probability of conversion to TSA after initial AM was calculated to be 45% whereas the probability of revision TSA after initial TSA was calculated to be 78%.

Two-way sensitivity analyses were also performed to measure the effect of age and the rate of TSA per year

---

Fig 2. The total expected QALYs for TSA and AM were calculated at 1-year intervals between the ages of 35 and 90 years. The differences between these 2 values (TSA−AM) were then plotted according to age. The resulting line transitioned from negative to positive values (i.e., crossed the x-axis) when the theoretical patient reached 66 years, suggesting that TSA performed in patients older than 66 years may have a net-positive effect on the total expected QALYs whereas TSA performed in patients younger than 66 years may have a net-negative effect on the total expected QALYs.

66 years. According to the model, theoretical patients who underwent AM spent 8.8 years in the well state before requiring TSA and another 20 years in the well state before requiring revision TSA. Overall, TSA was the preferred treatment strategy for patients older than 66 years (Fig 2).

Sensitivity Analyses

Our decision model was very sensitive to age at initial surgery, the probability of treatment failure for both the AM and TSA strategies, and the treatment utilities after AM, TSA, and revision TSA (Table 1). In other words, small changes in the values assigned for these parameters would result in a large change in the overall remaining QALYs and, potentially, the preferred treatment strategy. For example, if the treatment utility for AM had been assigned a value of 0.81 QALYs rather than 0.82 QALYs, then age threshold at which AM would no longer be the preferred strategy would have dropped from 66 to 47 years. On the basis of this finding and the results of baseline analysis, the model suggested that AM was the preferred treatment strategy for patients younger than 47 years, TSA was the preferred treatment for patients older than 66 years, and both treatment options were reasonable for patients aged between these 2 age groups. Similarly, if the treatment utility for TSA had been greater than 0.874 QALYs (baseline was 0.85), then TSA would have been the preferred treatment strategy for all ages (Fig 3). AM was the preferred strategy when the failure rate within the first 10 years after TSA was greater than 1.5% (baseline was 4.6%) (Fig 4). In a 47-year-old patient, the probability of conversion to TSA after initial AM was calculated to be 45% whereas the probability of revision TSA after initial TSA was calculated to be 78%.

Two-way sensitivity analyses were also performed to measure the effect of age and the rate of TSA per year.
after AM (Fig 5). As the failure rate of AM increased, the threshold age for the AM strategy decreased. The results of the 2-way sensitivity analysis using the utility values for AM and TSA are presented graphically in Figure 6.

Discussion

This study was designed to provide a theoretical basis on which future studies could be developed to investigate the potential role of arthroscopy in the treatment of young patients with glenohumeral OA. According to our decision model, AM was the preferred treatment strategy for patients younger than 47 years, primary TSA was the preferred treatment strategy for patients older than

Fig 3. Sensitivity analyses for treatment utilities of AM, TSA, and revision TSA. (A) TSA was the preferred strategy when the calculated utility for AM was less than 0.796 QALYs. (B) TSA was the preferred strategy when the calculated utility for TSA was greater than 0.874 QALYs. (C) TSA was the preferred strategy when the calculated utility of revision TSA was greater than 0.82.

Fig 4. Sensitivity analysis for annual failure rate after TSA for period of up to 10 years postoperatively. If the annual revision rate was greater than 1.5%, then AM was the preferred strategy.

Fig 5. A 2-way sensitivity analysis was performed varying the age at surgery and the failure rate of AM. The shaded area indicates the strategy with the greatest expected QALYs for the given parameters.
66 years, and both treatment options were reasonable for patients aged between 47 and 66 years.

AM of glenohumeral OA in young patients is a palliative treatment approach that has historically been successful in delaying the need for arthroplasty. In most studies the goals of treatment were to provide symptomatic relief by stabilizing chondral defects, eliminating mechanical crepitation, and releasing capsular contractures. Weinstein et al.\textsuperscript{40} reported good results in a series of 25 patients with a mean age of 46 years (range, 27 to 72 years) who underwent glenohumeral debridement, lavage, and subacromial bursectomy. However, results were less satisfactory in patients with more advanced joint degeneration. Richards and Burkhart\textsuperscript{41} performed arthroscopic debridement with the addition of capsular releases in 8 young patients with glenohumeral OA (mean age, 56 ± 12 years). In their study, an improvement in range of motion was shown with a mean symptom-free period of approximately 9 months. Van Thiel et al.\textsuperscript{9} showed significant pain relief after glenohumeral debridement in 55 of 71 patients (77.5\%) after a mean follow-up period of approximately 27 months. In a study by Millett et al.,\textsuperscript{5} 30 shoulders with advanced glenohumeral OA (mean age, 52 years) underwent a comprehensive arthroscopic management procedure that included extensive debridement, capsular releases, humeral osteoplasty, axillary neurolysis, and open subpectoral biceps tenodesis as a means of delaying the need for arthroplasty. Of the 30 shoulders, 6 (20.0\%) required arthroplasty after a mean of 1.9 years after the index surgery. After a mean follow-up period of 2.6 years, patients who did not progress to arthroplasty showed significant improvements in American Shoulder and Elbow Surgeons’ scores ($P < .001$) and pain levels ($P < .05$), with a median patient satisfaction rating of 9 of 10. In addition, survivorship was calculated to be 92\% at 12 months and 85\% at 24 months.

On the other hand, primary TSA in patients younger than 60 years is commonly associated with inferior clinical outcomes. Lo et al.\textsuperscript{30} randomized 42 patients to receive either hemiarthroplasty or TSA for advanced glenohumeral OA and compared their clinical outcomes at regular intervals until final follow-up at 24 months postoperatively. With specific regard to the TSA group, the mean age was 70.4 ± 9.0 years at the time of surgery. At final 24-month follow-up, the mean Constant score was 68.9 ± 18.4 in the TSA group. In a younger population, Raiss et al.\textsuperscript{31} evaluated the clinical outcomes of 21 patients with a mean age of 55 years (range, 37 to 60 years) who underwent primary TSA for advanced glenohumeral OA. After a mean follow-up period of 7 years (range, 5 to 9 years), the mean Constant score was 64.5 without evidence of glenoid component loosening. Denard et al.\textsuperscript{10} studied the clinical outcomes in 50 young patients (aged <55 years) with glenohumeral OA who were treated with primary TSA. In their study 5- and 10-year construct survivorship data were calculated to be 98\% and 62.5\%, respectively, in patients younger than 55 years. On the basis of these data, it appears that construct failure is more likely to occur between 5 and 10 years postoperatively when primary TSA is performed in younger patients. In addition, lower patient satisfaction ratings in younger patients after primary TSA have also been reported.\textsuperscript{10,16}

The question remains as to when arthroplasty should be considered in young patients with glenohumeral OA. Of course, there are other factors to be considered in the decision-making process. These include etiology, individual anatomy, patient symptoms, and patient expectations, in addition to considerations involving implant durability, future options, and the possible need for joint salvage procedures. Despite the necessary differences in individual treatment options, this study provides a basis for future research on the topic.

Limitations

Given our study design, there are several limitations that warrant discussion to help prevent misinterpretation of our results. First, we were only able to consider the chronologic age of the theoretical patient cohort. However, individual life expectancy and the desired level of activity are parameters that should be considered when making treatment decisions. Second, although most studies that reported the outcomes after AM did not use identical techniques, the results were combined and reported as frequency-weighted means. For example, Weinstein et al.\textsuperscript{40} performed glenohumeral debridement and bursectomy; van Thiel et al.\textsuperscript{9} performed debridement with the addition of capsular releases; and Millett et al.\textsuperscript{5} performed debridement with the addition of capsular

![Fig 6. A 2-way sensitivity analysis was performed varying the utilities of AM and TSA. The shaded area indicates the strategy with the greatest expected QALYs for the given parameters.](Image)
releases, osteophyte removal, and axillary neurolysis. Third, the estimation of utility values for TSA was based on Level IV studies, including those that evaluated quality of life after total hip arthroplasty. Fourth, we did not make adjustments according to the differences in reported outcomes based on surgeon experience or volume.

Conclusions

According to our theoretical decision model, AM was the preferred treatment strategy for patients younger than 47 years, primary TSA was the preferred treatment strategy for patients older than 66 years, and both treatment options were reasonable for patients aged between 47 and 66 years.

References

29. Gartsman GM, Roddie TS, Hammerman SM. Shoulder arthroplasty with or without resurfacing of the glenoid in


