

Patient-reported outcome measures (PROMs): Can they be used to guide patient-centered care and optimize outcomes in total knee replacement?

Published as: *Michelle Tew , Kim Dalziel , Philip Clarke , Anne Smith , Peter F Choong, Michelle Dowsey Patient-reported outcome measures (PROMs): Can they be used to guide patient-centered care and optimize outcomes in total knee replacement? Qual Life Res doi: 10.1007/s11136-020-02577-4. Online ahead of print.*

Abstract

Purpose

As patient-reported outcome measures (PROMs) are increasingly integrated into clinical practice, there is a need to translate collected data into valuable information to guide and improve the quality and value of patient care. The purpose of this study was to investigate health-related quality-of-life (QoL) trajectories in the five years following total knee replacement (TKR) and the patient characteristics associated with these trajectories. The feasibility of translating QoL trajectories into valuable information for guiding patient-centered care was also explored.

Methods

Data on patients who underwent TKR between 2006 and 2011 from a single-institution registry were extracted including patient-reported QoL (captured using the Short Form Survey (SF-12) instrument) up to 5 years post-surgery. QoL trajectories were modelled using latent class growth analysis. Quality-adjusted life-years (QALYs) were calculated to illustrate longer term health benefit. Multinomial logistic regression analyses were performed to examine the association between trajectory groups and baseline patient characteristics.

Results

After exclusions, 1,553 patients out of 1,892 were included in the analysis. Six unique QoL trajectories were identified; with differing levels at baseline and improvement patterns post-surgery. Only 18.6% of patients were identified to be in the most positive QoL trajectory (low baseline, large sustainable improvement after surgery) associated with the greatest gain in QALY. These patients were likely to be younger, have no co-morbidities and report greater pain at pre-surgery than most in other QoL trajectories.

Conclusions

Our findings demonstrate the importance of underlying heterogeneity in QoL trajectories, resulting in variable QALY gains. There is scope in translating routinely collected PROMs to improve shared decision making allowing for more patient engagement. However, further research is required to identify suitable approaches of its implementation into practice to guide clinical care and maximize patient outcomes.

Key words: knee replacement, patient reported outcomes, quality of life, patient-centered care

Key points for decision makers

- There is strong evidence indicating important heterogeneity in QoL trajectories in TKR patients indicating not all patients benefit from the surgical procedure in the same way.
- Knowledge of the combination of characteristics that predisposes a patient to trajectories with poor health gains can be useful in anticipating possible outcomes and mitigating such risks.
- Associating patient-reported outcomes such as QoL to patient characteristics can facilitate delivery of individualized health care as it allows patient engagement in shared decision making to help optimize outcomes.

Patient-reported outcome measures (PROMs): Can they be used to guide patient-centered care and optimize outcomes in total knee replacement?

1 Introduction

The value of patient-reported outcomes measures (PROMs) to evaluate outcomes after surgery is gaining recognition and the need to integrate these into clinical practice is becoming increasingly important [1-6]. For surgical interventions such as total knee replacement (TKR), patient-reported outcomes generally include pain, function and health-related quality-of-life (QoL). Patients typically experience a significant improvement in these outcomes within the first year following surgery and the effects tend to plateau in subsequent years [7-11]. Although this pattern of recovery is well-known, it is unclear if it can be universally applied because up to 20% of TKR patients do not gain clinically meaningful improvement following surgery [12, 13]. Recent research on longer-term functional outcomes identified a subgroup with delayed functional gains [14] which indicates that longer-term recovery patterns, and consequently effectiveness gained from surgery, differ considerably across patients. Therefore, it is important to better understand the longer-term implications of TKR, particularly in patients who experience poorer outcomes and whether these patients can be identified early to optimize their outcomes.

There is growing evidence that QoL is an important predictor of outcomes such as complications, hospitalisation and mortality [15-18]. This suggests that a better understanding of patients' QoL trajectories can reveal important information on disease progression and outcomes. QoL PROMs have the additional benefit of capturing the necessary information for cost-effectiveness analysis allowing decision makers to compare the value of health interventions and prioritize resource allocation. Further, associating patterns of QoL with patient characteristics may help identify groups for whom TKR may be of higher or lower value. This can help facilitate the rational deployment of TKR to those who stand to benefit the most while targeting others for more appropriate alternative interventions or management strategies. This is important as healthcare systems are transitioning from volume- to value-based health care as a means of improving sustainability of the healthcare system whilst also optimizing patient outcomes and experience [19, 20]. This is particularly relevant for surgical interventions such as TKR, which are performed in high volumes annually and are associated with considerable health care costs, amounting to \$11.8 billion in 2014 in the US alone [21].

In this study, we aimed to identify unique QoL trajectory groups for TKR patients from routinely collected PROMs, demonstrate the distinct variations in health gains and explore the individual characteristics related to group membership, using a rich data source with 5 years QoL data. By quantifying health gains using quality-adjusted life years (QALYs), a commonly used outcome in economic evaluations such as cost-effectiveness analyses, we demonstrate the feasibility of how QoL trajectories can be translated into valuable information for guiding patient-centered care. This will also provide a better understanding of the value of surgery across different trajectory groups.

2 Material and Methods

2.1 Data

The St. Vincent's Melbourne Arthroplasty Outcomes (SMART) Registry prospectively captures clinical and patient-reported outcomes in all patients undergoing elective hip and knee replacement at the study institution in Melbourne, Australia. The study institution is a tertiary referral centre for joint replacement surgeries and receives state-wide referrals. Registry data collection started in 1998 and to date, has recorded over 10,000 procedures with approximately 800 registered annually [11, 22, 23]. This dataset is ideal to answer the research question regarding longer term trajectories as at least 5 years of annual follow-up data are available. This included patients who had TKR between January 1, 2006 and December 31, 2011. Individuals were excluded if they had missing baseline QoL score, no QoL scores at all subsequent time points, underwent early revision or died within 2 years of surgery. Our analysis included patients with at least two QoL scores. For individuals that underwent bilateral knee surgery during the study period, only the most recent TKR was included in the analysis. Sensitivity analyses were conducted to assess the effect of our exclusion criteria.

Baseline data on patients were prospectively collected and included baseline socio-demographic and patient characteristics such as age, gender, body mass index (BMI), smoking status, co-morbidity measures such as Charlson Co-morbidity Index (CCI) and American Society of Anesthesiologist (ASA) Physical Status Classification. Cultural and linguistic diversity was measured via the need for an interpreter, socioeconomic status was measured via the Socio-Economic Index for Areas (SEIFA) [24] and geographical accessibility index (ARIA+) [25] reflected rurality. Clinical variables included bilateral knee surgery, Kellgren-Lawrence scale [26] describing radiographic severity of osteoarthritis and the Knee Society Scores (KSS) [27] subscales for pain and function.

2.2 Quality-of-life measurements

Patients completed the 12-item Short Form Survey (SF-12) prior to surgery and annually post-operatively. Baseline and annual QoL scores up to 5 years post-surgery were considered for analysis. SF-12 responses were transformed into utility values using a published algorithm [28]. A utility value is a general index of wellbeing used for economic evaluation where 1 is equivalent to 'full health' and 0 is equivalent to being 'dead' with scoring algorithms based on public preferences for health states.

2.3 Statistical analysis

2.3.1 Latent class trajectory analysis

Latent class growth analysis (LCGA) was used to identify subgroups of patients according to their trajectory of QoL (described using utility values) pre-surgery and up to 5 years following TKR. LCGA is a semi-parametric technique used to classify distinct subgroups that follow a similar pattern of change over time hence appropriate for analyzing longitudinal data [29]. This means that patients exhibiting similar patterns of QoL are grouped forming sets of

homogenous classes. LCGA is able to accommodate missing data such that patients with missing QoL values at several time points are not excluded from the analysis thus minimizing the exclusion of patients [30, 31].

2.3.2 Identifying trajectory groups

As the number of potential trajectories is unknown, a series of models considering 1 to 8 classes were estimated. The censored normal model was selected as the most appropriate for the available data. The Bayes Information Criteria (BIC) is a commonly used criteria to assess model fit, where higher BIC values indicate better model fit [32]. The choice for optimal model was guided by a combination of factors including our research objective, goodness-of-fit statistics Akaike's Information Criteria (AIC), model interpretability, posterior group-membership probability diagnostics [31, 32]. The latter set of diagnostics included ensuring all groups displayed average group posterior probabilities above 0.7 [29] and odds of correct classification (OCC) were greater than 5 [32]. Patients were assigned to the trajectories which they had the highest posterior probability of membership.

2.3.3 Estimating QALYs

Quality-adjusted life-year (QALY) is a common metric used to measure health benefit and incremental outcomes are of interest for economic analysis to quantify the value of interventions [33]. QALYs for each QoL trajectories were calculated using patient-level utility values using the area under the curve method [34]. To quantify the effectiveness (health benefit) gained from the intervention, QALYs gained (incremental QALYs) were calculated for each patient assuming the patient experienced no change from baseline utility if the patient had not had a TKR [10, 35-38].

2.3.4 Multinomial logistic regression analysis

Based on assigned trajectories, multinomial logistic regression analysis weighted by probability of class membership was performed to examine the association between trajectory groups and baseline patient characteristics. The multivariable model included variables identified as potentially important discriminators of class membership in the univariable multinomial logistic regression analyses (those displaying associations at $P < 0.10$). Tests for collinearity were conducted with variance inflation factor (VIF) greater than 10 and tolerance of less than 0.1 considered to indicate the presence of multi-collinearity. The trajectory group with the highest incremental QALYs was used as reference category against which other trajectory groups were compared. All analyses were conducted using Stata SE14 (StataCorp, College Station, TX, USA), employing Traj plugin for LCGA.

3 Results

3.1 Study population

1,553 TKR patients were included in the analysis after 339 cases were excluded based on: missing baseline QoL (n=3), no follow-up QoL (n=36), early death (n=14), early revision

(n=32) and bilateral surgeries where the most recent surgery was already included (n=254). Table 1 displays the baseline patient characteristics who were on average 70.1 years (SD,8.5) and 67.4% were female and mean QoL utility of 0.56 (SD,0.11) pre-operatively. Of those included, complete QoL data from baseline and across all 5 years were available for 1,218 patients (78%).

3.2 Model selection

The model with six classes was chosen to achieve a balance between model parsimony and adequately identifying distinct QoL patterns to demonstrate heterogeneity within the cohort to provide insights on the longer-term QoL outcomes and the potential value of surgical intervention across different patient groups. The 6-class model produced six distinct QoL trajectories (Figure 1) and met all diagnostic tests criteria. The probability of membership for allocated class ranged between 0.78 and 0.85 and displayed OCC above the minimum value of 5 (full results can be found in Supplementary Material Tables S1 and S2). The addition of excluded individuals in the sensitivity analysis produced similar results.

3.3 Characterization of classes

The trajectories were characterized by 3 main phases; pre-surgery, post-operative improvement (period between pre-surgery and Year 1) and maintenance (after Year 1). Table 2 provides the description for each of the trajectories, total and incremental QALYs gained over the 5-year period.

Total QALYs of the trajectory with lowest QoL (Trajectory 1) was 2.62 (SD, 0.19) and the number of QALYs increased with higher utility values for subsequent trajectories. In terms of effectiveness gained from TKR, incremental QALYs were lowest for Trajectory 1 (0.16 (SD,0.35)) and greatest gain for trajectory 5 (1.42 (SD,0.40)). Although patients in Trajectory 6 had the greatest number of QALYs, estimated incremental gains from TKR were small at 0.39 (SD, 0.37) compared to most other trajectories.

3.4 Characterization of patients across trajectory groups

Baseline patient socio-demographic and clinical characteristics were compared across the 6 QoL trajectories and are provided in Supplementary Material Table S3. Patient characteristics differed across trajectories. The mean age of patients in Trajectories 3 and 5 was lower than in other trajectories. There was a higher proportion of females in trajectories reporting poorer QoL. Trajectory 1 had the largest proportion of patients reporting severe baseline pain (71.3%) and lowest baseline KSS function (26.7 (SD,19.8)) compared to others.

Although there was a low chance of collinearity (tolerance range between 0.79 and 0.98; mean VIF=1.14) when all variables were included, to achieve a parsimonious model, only patient characteristics displaying associations of $P < 0.10$ from the univariable regression models (Supplementary Material Table S4) were included in the final multivariable model. These were age, gender, BMI, interpreter, CCI, ASA, rurality, baseline KSS pain and function. Results from multivariable multinomial logistic regression are presented in Table 3

(Supplementary Material Figure S1), showing the relative risk of belonging in the respective trajectory for each patient characteristic.

Compared to patients with the greatest incremental QALY (Trajectory 5), patients with the lowest gains from TKR (Trajectory 1) are more likely to have co-morbidities, high ASA score, need an interpreter, more likely to report lower KSS function score (poorer mobility) and less likely to be in rural residence. Patients with moderate, sustained gains in Trajectory 2 are more likely to be older, female, require an interpreter, have co-morbidities, less likely to be in a rural residence, more likely to report lower KSS function score and are less likely to report severe pain compared to those with large gains (Trajectory 5). Patients exhibiting slow progressive improvement (Trajectory 3) were found to be more likely to have co-morbidities and report mild than moderate/severe pain compared to those whose improvement peaked earlier (Trajectory 5). Compared to Trajectory 5, patients in Trajectory 4 were older, have co-morbidities, and less likely to report moderate/severe pain. Patients consistently reporting high QoL (Trajectory 6) were likely to be older, less likely to report moderate/severe pain and more likely to report higher KSS function score compared to Trajectory 5. A summary of these findings is presented in Table 4.

4 Discussion

Using latent class growth analysis, we identified 6 distinct QoL trajectories indicating the presence of significant heterogeneity in QoL outcomes among TKR patients. Although most patients exhibited a trajectory profile that is commonly reported in the literature (improvement within 1 year followed by a plateau), the distinct difference observed in this study is that patients had variable gains following surgery and not all patients maintained the improvement. This highlights that patients will not universally achieve large QoL improvement following TKR as is commonly reported in the literature. Trajectory 5 (large sustainable improvement after surgery) was identified to be the most positive QoL trajectory with the greatest gain in QALYs. However, only 18.6% of the patients were classified in this trajectory and were likely to be younger, have no co-morbidities and report greater pain at pre-surgery than most in other QoL trajectories.

While much research has focused on identifying potential risk factors and integrating these to improve medical decision making, associating patient-reported outcomes such as QoL to these patient characteristics can facilitate delivery of individualized health care as it allows patient engagement in shared decision making to help optimize outcomes [20, 39]. The unique QoL trajectories identified in this study clearly show variations in the benefits of TKR; one-year post-surgery and in the longer term, and the combination of patient characteristics associated with each trajectory. Whilst there are limitations in employing the current findings to deterministically identify patient subgroup most likely to have poor outcomes, knowledge of the combination of characteristics (Table 4) that predisposes a patient to trajectories with poor health gains (for example, trajectories 1, 2 and 6) can be useful in anticipating possible outcomes and mitigating such risks. This may include managing pre-surgery expectations [40], personalizing self-management plans [41], careful planning in managing co-morbidities to optimize patients prior to surgery [42] and tailoring

pre-surgery management through mindfulness training to maximize outcomes in these patients [43].

There is also potential to use this information to improve post-surgical management to optimize care. Correlating patient characteristics with patient-reported QoL responses can help clinicians track progress and identify patients who are unlikely to obtain the maximum effectiveness from the treatment; for example, elderly female patients with moderate pain pre-surgery who consistently report low QoL may not benefit fully from the standard prescribed post-surgical management and may require an individualized approach. This gives both the patients and providers opportunity to engage and plan follow-up consultations based on goals and expectations for physical [44-46] or mental health [43] therapies to improve outcomes. Recognizing the variability in health trajectories could also enable patients to have realistic expectations, to better understand their clinical course and facilitate discussions with their surgeons [1]. This allows for the opportunity to tailor the evolving care post-surgery on an as-needed basis. While understanding the patient characteristics associated with these trajectories is important, it is acknowledged that beyond these characteristics, psychological factors such as pain-related beliefs and psychological distress can also influence TKR outcomes [47-49] and should be considered alongside.

To date, trajectory analysis on TKR patients have mostly focused on pain and function trajectories and have also demonstrated heterogeneity within the TKR population; commonly identifying the presence of a subgroup with poor pain and/or function outcomes comprising between 14% and 23% of the study cohort [11, 50, 51]. While it is unclear if patients with a low QoL trajectory (Trajectory 1) were non-responders or those reporting poor pain/function outcomes after surgery, some similarities in the characteristics of these patients were observed. Patients in Trajectory 1 had higher BMI, were more likely to be co-morbid, report severe pain, have low mental and physical well-being (Supplementary Material Table S3) which are consistent with the predictors of non-responders [22] or poor pain and function outcomes [11, 50]. For these patients, the prescribed standard surgical treatment and follow-up plans are unlikely to be adequate, thus resulting in poor patient outcomes and low value care. Therefore, by maximizing the use of PROMs to better understand potential QoL trajectories, clinicians can be better informed on how they may plan to manage subgroups with these characteristics and assign patients to more appropriate level of surveillance and better supportive care or alternative rehabilitation programs to optimize the outcomes of those who are truly experiencing low QoL long-term after surgery.

This study showed that improvement in QoL following surgery was observed to be the greatest among younger patients; Trajectory 5 and Trajectory 3 albeit over a longer period. Although the observed associations were statistically significant, they were relatively weak, and this could be due to the small number of patients under the age of 60 (approximately 12% of sample). Historically, younger patients are considered as less appropriate candidates compared to elderly patients due to the higher risk for revision [52]. This is likely related to duration of prostheses survivorship and higher levels of activity among younger patients [53]. While revision risk is an important consideration, the current study provides additional

insights. It may be useful for clinical practice to consider the potential benefits and value to be gained from the intervention when making surgical recommendations, particularly in younger patients [54]. Post-marketing surveillance and advances in technology have led to improvements in prostheses survivorship which have now reached 90% at 20 years and even 82% at 25 years [55]. Therefore, having to wait for advanced age to be suitable for surgery may represent a missed opportunity to improve an individual's well-being and labor force productivity.

As PROMs including QoL are increasingly recognized as an important consideration in clinical care, it is important these are routinely captured pre- and post-surgery using relevant tools to evaluate the effectiveness and value of intervening [56]. These findings also reinforce the need to encourage PROMs collection beyond the one-year post-surgery mark as delayed improvers (Traj 3) or diverging trends (e.g. Traj 4 and 5) can be indicators of sub-optimal care. Beyond routine collection of PROMs, there also needs to be considerations in integrating these into shared decision-making tools and identifying suitable approaches to implement these in practice to better guide clinical care and improve the value of surgery. Additionally, risk stratification is an important approach in advancing research [57], thus the ability to identify homogenous subgroups based on a combination of characteristics amongst a heterogenous cohort can be useful in selecting the right patients for trials of novel interventions allowing for a more targeted approach.

Because of the rapidly growing rates of utilization and large costs associated with TKR, the judicious use of scarce healthcare resources is ever more important to ensure sustainability for health insurers and health systems. Further, the appropriateness of the surgery in selected patients has also been called into question where studies showed up to one-third of TKRs were deemed to be inappropriate procedures [58, 59]. Therefore, it is important to target those whom we can maximize outcomes and improve value of care. We find patients reporting good QoL prior to surgery (Trajectory 6) were among those with small gains. Though it is uncertain if these patients have merely adapted to their condition hence report higher levels of QoL than others with the same condition [60], it may be important to understand the rationale for surgical intervention in these patients. While TKR is widely regarded as a cost-effective procedure in general, this raises the question if TKR is necessarily cost-effective for all patients. Some groups of patients may require additional care and healthcare services demands to improve their outcomes. This may be relevant to patients exhibiting poor long-term QoL outcomes with small gains in health benefit such as those in Trajectories 1, 2 and 6, which in combination contributes to a significant proportion (55%) of the cohort. Therefore, further research to quantify the healthcare needs and assess the cost-effectiveness across these sub-groups would be helpful in understanding the true value of surgery amongst the group of heterogenous TKR patients.

4.1 Limitations

Several limitations should be considered when interpreting these results. The generalizability of the findings could be limited as patients were from a single-center. However, the demographics of patients in this study closely reflect those reported in our National Joint

Replacement Registry [61]. It is acknowledged that changes to modifiable characteristics such as comorbidity over time can affect QoL trajectories [62]. However, it is difficult to ascertain the extent of this in the current study unless such information is also captured over time. QoL assessments can be subject to biases known as response shifts where patients could change the way they evaluate themselves and respond to surveys over time [63]. While studies have shown that changes in health outcomes were underestimated when response shifts were not accounted for in TKR patients [64, 65], another has shown that despite adjusting for large response shifts, it did not change the authors' clinical interpretation of the results [66]. In the context of our study where all patients were surveyed in the same manner across time, it is unlikely to change the conclusions drawn from our analysis. It is noted that our assumption of no change from baseline made in the calculation of incremental QALYs may result in an overestimation of QALYs as a result of regression to the mean [67]. Conversely, deterioration in QoL due to aging or absence of surgery may also result in an underestimation. The application of this assumption follows published economic evaluations [10, 34, 35, 37, 38]. Variables such as co-morbidity, ASA, KL scores and socio-economic indicators (SEIFA) were dichotomized to avoid small cell sizes which could reduce the sensitivity of our analysis.

5 Conclusion

There is strong evidence indicating important heterogeneity in QoL trajectories in TKR patients resulting in variable gains in QoL and QALYs across different trajectory groups. This indicates not all patients benefit from the surgical procedure in the same way. With the growing recognition to support patient-centered care, PROMs may have a particular usefulness when employed alongside patient characteristics for tracking and guiding clinical care to maximize patient outcomes and justifying costs of surgical intervention. Future research should focus on identifying approaches of its implementation into clinical practice.

Figure 1: QoL trajectory profiles and class membership for six-class model (coloured figure)

Legend

Traj 6 – High baseline, moderate sustained improvement (7%)

Traj 5 – Low baseline, large sustained improvement (19%)

Traj 4 – Low baseline, large unsustained improvement (18%)

Traj 3 – Low baseline, moderate improving (9%)

Traj 2 – Low baseline, moderate sustained improvement (30%)

Traj 1 – Low baseline, small sustained improvement (18%)

Table 1: Baseline characteristics

	No.	(%)
Total number of patients	1553	
Patient characteristics		
Age, mean (SD)	70.1	8.5
Female	1047	67.4
BMI, mean (SD)	32.8	6.0
Aetiology		
Osteoarthritis	1456	93.8
Other ^a	97	6.3
Bilateral surgery	382	24.6
Smoker		
No	1065	68.6
Ex	381	24.5
Yes	107	6.9
Interpreter	231	14.9
CCI		
0	875	56.3
1+	678	43.7
ASA		
1/2	901	58.0
3/4	652	42.0
KL		
<4	757	48.7
4	796	51.3
SEIFA deciles		
1-5	562	36.2
6-10	991	63.8
Rural residence	263	16.9
Patient-reported outcomes		
KSS pain ^b		
Mild	88	5.7
Moderate	596	38.4
Severe	869	56.0
KSS function ^b , mean (SD)	36.0	20.5
SF-12, mean (SD)		
PCS	23.1	8.0
MCS	45.6	16.1
Utility	0.56	0.11

^a Other aetiology includes rheumatoid arthritis and avascular osteonecrosis.

^b KSS pain scores were categorised as follows; none (50), mild occasional (45) mild on stairs (40), mild on walking (30), moderate occasional (20), moderate continual (10) and severe pain as (0) points. KSS function score assesses walking, stair ability and use of walking aids and ranges for 0 to 100 with a higher score indicating better function.

Abbreviations: ASA, American Society of Anesthesiologist; BMI, body mass index; CCI, Charlson Co-morbidity Index; KL, Kellgren-Lawrence scale; MCS, mental component score; PCS, physical component score; SEIFA, Socio-Economic Index for Areas; KSS, Knee Society Score.

Table 2: Description of each QoL trajectories by phases and estimated QALYs over 5-years

	Traj 1	Traj 2	Traj 3	Traj 4	Traj 5	Traj 6
Phases						
Pre-surgery QoL	Low	Low	Low	Low	Low	High
Post-surgery QoL improvement at year 1	Small	Moderate	Moderate	Large	Large	Moderate
Maintenance of trajectory after year 1	Maintained	Maintained	Improving	Declined	Maintained	Maintained
Measure of health gains						
Total QALY (SD) ^a	2.62 (0.19)	3.15 (0.17)	3.55 (0.20)	3.80 (0.19)	4.20 (0.20)	4.42 (0.20)
Incremental QALY (SD) ^{a b}	0.16 (0.35)	0.42 (0.46)	0.75 (0.47)	0.85 (0.43)	1.42 (0.40)	0.39 (0.37)

^a Complete case analysis

^b QALYs gained as a result of TKR assuming patient experienced no change from baseline utility if the patient not had a TKR
Abbreviations: QALY, quality-adjusted life-years; QoL, quality-of-life; Traj, trajectory

Table 3: Multivariable multinomial logistic regression showing relative risk (RRR) of belonging in each of the trajectory groups compared to Trajectory 5 (highest incremental QALYs/health gains)

Variable	Trajectory 1			Trajectory 2			Trajectory 3			Trajectory 4			Trajectory 6		
	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value
Age	1.00	0.98-1.03	0.73	1.03	1.01-1.05	0.004	1.00	0.97-1.03	0.90	1.03	1.01-1.06	0.001	1.04	1.01-1.07	0.02
Female	1.30	0.88-1.91	0.19	1.46	1.04-2.06	0.03	1.17	0.72-1.91	0.52	1.33	0.91-1.94	0.14	0.70	0.41-1.17	0.17
BMI	1.02	0.99-1.05	0.29	1.02	0.99-1.05	0.17	1.01	0.98-1.05	0.46	1.03	1.00-1.06	0.07	0.98	0.94-1.02	0.35
Interpreter	2.61	1.49-4.59	0.001	2.22	1.30-3.78	0.003	1.76	0.84-3.70	0.14	1.56	0.84-2.87	0.16	1.00	0.35-2.87	0.99
CCI															
0	1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a		
≥1	2.71	1.85-3.95	<.001	2.15	1.53-3.03	<.001	1.99	1.25-3.17	0.004	1.84	1.25-2.69	0.002	1.07	0.62-1.84	0.82
ASA score															
1/2	1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b		
3/4	2.10	1.44-3.08	<.001	1.29	0.92-1.81	0.15	1.33	0.84-2.12	0.22	1.32	0.90-1.95	0.15	0.71	0.39-1.29	0.26
Rural residence	0.47	0.28-0.80	0.005	0.66	0.44-0.99	0.04	0.68	0.37-1.23	0.20	1.02	0.67-1.56	0.91	0.96	0.54-1.73	0.90
KSS pain															
Mild	1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a		
Moderate	0.41	0.14-1.15	0.09	0.46	0.19-1.09	0.08	0.29	0.10-0.86	0.03	0.33	0.14-0.79	0.01	0.23	0.09-0.59	0.003
Severe	0.60	0.22-1.68	0.33	0.38	0.16-0.91	0.03	0.33	0.11-0.95	0.04	0.25	0.10-0.61	0.002	0.12	0.04-0.34	<.001

KSS function ^c	0.98	0.97-0.98	<.001	0.99	0.98-1.00	0.008	1.00	0.98-1.01	0.53	1.00	0.99-1.01	0.53	1.03	1.01-1.04	<.001
------------------------------	------	-----------	-------	------	-----------	-------	------	-----------	------	------	-----------	------	------	-----------	-------

^a Overall P Value < 0.001

^b Overall P Value = 0.001

^c Per point increase in KSS function score. The score assesses walking, stair ability and use of walking aids and ranges for 0 to 100 with a higher score indicating better function.

Abbreviations: ASA, American Society of Anesthesiologist; BMI, body mass index; CCI, Charlson Co-morbidity Index; KSS, Knee Society Score; RRR, relative risk ratio

Table 4: Patient characteristics associated with each of the trajectories compared to Trajectory 5 (highest incremental QALYs/health gains).

Trajectory 1	Trajectory 2	Trajectory 3	Trajectory 4	Trajectory 6
	Older Female		Older	Older
High ASA score (3 and above)				
Need interpreter	Need interpreter			
Have co-morbidities	Have co-morbidities	Have co-morbidities	Have co-morbidities	
Less likely to be in rural residence	Less likely to be in rural residence			
Report lower (below 40) KSS function score ^a	Report lower (below 40) KSS function score ^a			Report higher (above 40) KSS function score ^a
	Less likely to report severe pain ^a	Less likely to report moderate/severe pain ^a	Less likely to report moderate/severe pain ^a	Less likely to report moderate/severe pain ^a

^a KSS pain scores were categorised as follows; none (50), mild occasional (45) mild on stairs (40), mild on walking (30), moderate occasional (20), moderate continual (10) and severe pain as (0) points. KSS function score assesses walking, stair ability and use of walking aids and ranges for 0 to 100 with a higher score indicating better function.

References

1. Smith AB, Schwarze ML. Translating patient-reported outcomes from surgical research to clinical care. *JAMA surgery*. 2017;152(9):811-2.
2. Black N. Patient reported outcome measures could help transform healthcare. *BMJ : British Medical Journal*. 2013;346:f167.
3. Van Der Wees PJ, Nijhuis-Van Der Sanden MW, Ayanian JZ, Black N, Westert GP, Schneider EC. Integrating the use of patient-reported outcomes for both clinical practice and performance measurement: views of experts from 3 countries. *The Milbank quarterly*. 2014 Dec;92(4):754-75.
4. Aaronson NK, Snyder C. Using patient-reported outcomes in clinical practice: proceedings of an International Society of Quality of Life Research conference. *Quality of Life Research*. 2008 2008/12/01;17(10):1295-.
5. Austin E, LeRouge C, Hartzler AL, Segal C, Lavalley DC. Capturing the patient voice: implementing patient-reported outcomes across the health system. *Quality of Life Research*. 2020 2020/02/01;29(2):347-55.
6. Hsiao C-J, Dymek C, Kim B, Russell B. Advancing the use of patient-reported outcomes in practice: understanding challenges, opportunities, and the potential of health information technology. *Quality of Life Research*. 2019 2019/06/01;28(6):1575-83.
7. Baker P, Muthumayandi K, Gerrand C, Kleim B, Bettinson K, Deehan D. Influence of body mass index (BMI) on functional improvements at 3 years following total knee replacement: a retrospective cohort study. *PloS one*. 2013;8(3):e59079.
8. Nilsdotter A-K, Toksvig-Larsen S, Roos E. A 5 year prospective study of patient-relevant outcomes after total knee replacement. *Osteoarthritis and cartilage*. 2009;17(5):601-6.
9. Schilling C, Dowsey MM, Clarke PM, Choong PF. Using patient-reported outcomes for economic evaluation: getting the timing right. *Value in Health*. 2016;19(8):945-50.
10. Dakin H, Gray A, Fitzpatrick R, MacLennan G, Murray D, Group KT. Rationing of total knee replacement: a cost-effectiveness analysis on a large trial data set. *BMJ open*. 2012;2(1):e000332.
11. Dowsey M, Smith A, Choong P. Latent class growth analysis predicts long term pain and function trajectories in total knee arthroplasty: a study of 689 patients. *Osteoarthritis and cartilage*. 2015;23(12):2141-9.
12. Alzahrani K, Gandhi R, Petruccioli D, MAHOMED N. Prevalence of clinically significant improvement following total knee replacement. *The Journal of rheumatology*. 2011;38(4):753-9.
13. Quintana JM, Escobar A, Arostegui I, Bilbao A, Azkarate J, Goenaga JI, et al. Health-related quality of life and appropriateness of knee or hip joint replacement. *Archives of internal medicine*. 2006;166(2):220-6.
14. Lange JK, DiSegna ST, Yang W, Li W, Franklin PD. Using Cluster Analysis to Identify Patient Factors Linked to Differential Functional Gains After Total Knee Arthroplasty. *The Journal of Arthroplasty*. 2019 2019/08/23/.
15. Clarke PM, Hayes AJ, Glasziou PG, Scott R, Simes J, Keech AC. Using the EQ-5D index score as a predictor of outcomes in patients with type 2 diabetes. *Medical care*. 2009;47(1):61-8.
16. Mapes DL, Lopes AA, Satayathum S, Mccullough KP, Goodkin DA, Locatelli F, et al. Health-related quality of life as a predictor of mortality and hospitalization: the Dialysis Outcomes and Practice Patterns Study (DOPPS). *Kidney international*. 2003;64(1):339-49.
17. Ediebah DE, Quinten C, Coens C, Ringash J, Dancey J, Zikos E, et al. Quality of life as a prognostic indicator of survival: A pooled analysis of individual patient data from canadian cancer trials group clinical trials. *Cancer*. 2018;124(16):3409-16.
18. Hansen TB, Thygesen LC, Zwisler AD, Helmark L, Hoogwegt M, Versteeg H, et al. Self-reported health-related quality of life predicts 5-year mortality and hospital readmissions in patients with ischaemic heart disease. *European journal of preventive cardiology*. 2015;22(7):882-9.
19. Obama B. United States health care reform: progress to date and next steps. *Jama*. 2016;316(5):525-32.

20. Spertus JA, Ghaferi AA. Transforming the National Surgical Quality Improvement Program to the Delivery of Precision Medicine to Improve the Value of Surgical Care: Summary of the John R. Clarke Keynote Address for the Surgical Outcomes Club 2016 Annual Meeting Transforming NSQIP to the Delivery of Precision Medicine. *JAMA Surgery*. 2017;152(9):815-6.
21. McDermott K, Freeman W, Elixhauser A. Overview of operating room procedures during inpatient stays in US hospitals, 2014. *HCUP Statistical Brief*. 2017;233:1-18.
22. Dowsey MM, Spelman T, Choong PF. Development of a Prognostic Nomogram for Predicting the Probability of Nonresponse to Total Knee Arthroplasty 1 Year After Surgery. *J Arthroplasty*. 2016 Aug;31(8):1654-60.
23. Tew M, Dalziel K, Dowsey M, Choong PF, Clarke P. Exploring the Impact of Quality of Life on Survival: A Case Study in Total Knee Replacement Surgery. *Medical Decision Making*. 2020:0272989X20913266.
24. Australian Bureau of Statistics. Information Paper: An Introduction to Socio-Economic Indexes for Areas (SEIFA) - 2039.0. 2006 [Accessed May 10, 2019]; Available from: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/2039.0>
25. Australian Bureau of Statistics. Australian Statistical Geography Standard (ASGS): Volume 5 - Remoteness Structure - 1270.0.55.005. 2016 [Accessed May 10, 2019]; Available from: <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1270.0.55.005?OpenDocument>
26. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthritis. *Annals of the rheumatic diseases*. 1957 Dec;16(4):494-502.
27. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clinical orthopaedics and related research*. 1989 Nov(248):13-4.
28. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. *Medical care*. 2004;42(8):851-9.
29. Andruff H, Carraro N, Thompson A, Gaudreau P, Louvet B. Latent class growth modelling: a tutorial. *Tutorials in Quantitative Methods for Psychology*. 2009;5(1):11-24.
30. Jones BL, Nagin DS, Roeder K. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods & Research*. 2001 2001/02/01;29(3):374-93.
31. Nagin DS, Odgers CL. Group-based trajectory modeling in clinical research. *Annual review of clinical psychology*. 2010;6:109-38.
32. Nagin DS, NAGIN D. *Group-based modeling of development*: Harvard University Press; 2005.
33. Sanders GD, Neumann PJ, Basu A, Brock DW, Feeny D, Krahn M, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: second panel on cost-effectiveness in health and medicine. *Jama*. 2016;316(10):1093-103.
34. Hunter RM, Baio G, Butt T, Morris S, Round J, Freemantle N. An educational review of the statistical issues in analysing utility data for cost-utility analysis. *Pharmacoeconomics*. 2015;33(4):355-66.
35. Jenkins P, Clement N, Hamilton D, Gaston P, Patton J, Howie C. Predicting the cost-effectiveness of total hip and knee replacement: a health economic analysis. *Bone Joint J*. 2013;95(1):115-21.
36. Richardson G, Manca A. Calculation of quality adjusted life years in the published literature: a review of methodology and transparency. *Health economics*. 2004;13(12):1203-10.
37. Schilling CG, Dowsey MM, Petrie DJ, Clarke PM, Choong PF. Predicting the long-term gains in health-related quality of life after total knee arthroplasty. *The Journal of arthroplasty*. 2017;32(2):395-401. e2.
38. Waimann CA, Fernandez-Mazambroz RJ, Cantor SB, Lopez-Olivo MA, Zhang H, Landon GC, et al. Cost-effectiveness of total knee replacement: a prospective cohort study. *Arthritis care & research*. 2014;66(4):592-9.

39. Lavalley DC, Chenok KE, Love RM, Petersen C, Holve E, Segal CD, et al. Incorporating Patient-Reported Outcomes Into Health Care To Engage Patients And Enhance Care. *Health affairs (Project Hope)*. 2016 Apr;35(4):575-82.
40. Cross M, Lapsley H, Barcenilla A, Parker D, Coolican M, March L. Patient Expectations of Hip and Knee Joint Replacement Surgery and Postoperative Health Status. *The Patient: Patient-Centered Outcomes Research*. 2009 2009/03/01;2(1):51-60.
41. Miller JL, Teare SR, Marlett N, Shklarov S, Marshall DA. Support for living a meaningful life with osteoarthritis: a patient-to-patient research study. *The Patient-Patient-Centered Outcomes Research*. 2016;9(5):457-64.
42. Dhatariya K, Levy N, Kilvert A, Watson B, Cousins D, Flanagan D, et al. NHS Diabetes guideline for the perioperative management of the adult patient with diabetes. *Diabetic Medicine*. 2012;29(4):420-33.
43. Dowsey M, Castle D, Knowles S, Monshat K, Salzberg M, Nelson E, et al. The effect of mindfulness training prior to total joint arthroplasty on post-operative pain and physical function: A randomised controlled trial. *Complementary therapies in medicine*. 2019 Oct;46:195-201.
44. Davis AM, Kennedy D, Wong R, Robarts S, Skou ST, McGlasson R, et al. Cross-cultural adaptation and implementation of Good Life with osteoarthritis in Denmark (GLA:D™): group education and exercise for hip and knee osteoarthritis is feasible in Canada. *Osteoarthritis and Cartilage*. 2018 2018/02/01;26(2):211-9.
45. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews*. 2015(1).
46. Skou ST, Roos EM. Good Life with osteoArthritis in Denmark (GLA: D™): evidence-based education and supervised neuromuscular exercise delivered by certified physiotherapists nationwide. *BMC musculoskeletal disorders*. 2017;18(1):72.
47. Jones AR, Al-Naseer S, Bodger O, James ETR, Davies AP. Does pre-operative anxiety and/or depression affect patient outcome after primary knee replacement arthroplasty? *The Knee*. 2018 2018/12/01;25(6):1238-46.
48. Riddle DL, Wade JB, Jiranek WA, Kong X. Preoperative pain catastrophizing predicts pain outcome after knee arthroplasty. *Clinical Orthopaedics and Related Research®*. 2010;468(3):798-806.
49. Vissers MM, Bussmann JB, Verhaar JA, Busschbach JJ, Bierma-Zeinstra SM, Reijman M. Psychological factors affecting the outcome of total hip and knee arthroplasty: a systematic review. *Seminars in arthritis and rheumatism*; 2012: Elsevier. p. 576-88.
50. Dumenci L, Perera R, Keefe F, Ang D, Slover J, Jensen M, et al. Model-based pain and function outcome trajectory types for patients undergoing knee arthroplasty: a secondary analysis from a randomized clinical trial. *Osteoarthritis and cartilage*. 2019.
51. Lenguerrand E, Wylde V, Gooberman-Hill R, Sayers A, Brunton L, Beswick AD, et al. Trajectories of pain and function after primary hip and knee arthroplasty: the ADAPT cohort study. *PloS one*. 2016;11(2):e0149306.
52. Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, et al. The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. *The Lancet*. 2017;389(10077):1424-30.
53. Kurtz SM, Lau E, Ong K, Zhao K, Kelly M, Bozic KJ. Future young patient demand for primary and revision joint replacement: national projections from 2010 to 2030. *Clinical orthopaedics and related research*. 2009 Oct;467(10):2606-12.
54. Goh GS-H, Liow MHL, Bin Abd Razak HR, Tay DK-J, Lo N-N, Yeo S-J. Patient-Reported Outcomes, Quality of Life, and Satisfaction Rates in Young Patients Aged 50 Years or Younger After Total Knee Arthroplasty. *The Journal of Arthroplasty*. 2017 2017/02/01;32(2):419-25.
55. Evans JT, Walker RW, Evans JP, Blom AW, Sayers A, Whitehouse MR. How long does a knee replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. *The Lancet*. 2019;393(10172):655-63.

56. Halawi MJ, Cote MP, Savoy L, Williams VJ, Lieberman JR. The Effect of Payer Type on Patient-Reported Outcomes in Total Joint Arthroplasty Is Modulated by Baseline Patient Characteristics. *The Journal of Arthroplasty*. 2019;34(6):1072-5.
57. Kent DM, Hayward RA. Limitations of Applying Summary Results of Clinical Trials to Individual PatientsThe Need for Risk Stratification. *JAMA*. 2007;298(10):1209-12.
58. Riddle DL, Jiranek WA, Hayes CW. Use of a validated algorithm to judge the appropriateness of total knee arthroplasty in the United States: a multicenter longitudinal cohort study. *Arthritis & rheumatology (Hoboken, NJ)*. 2014 Aug;66(8):2134-43.
59. Escobar A, Quintana JM, Aróstegui I, Azkárate J, Güenaga JI, Arenaza JC, et al. Development of explicit criteria for total knee replacement. *International journal of technology assessment in health care*. 2003;19(1):57-70.
60. Groot W. Adaptation and scale of reference bias in self-assessments of quality of life. *Journal of Health Economics*. 2000 2000/05/01/;19(3):403-20.
61. (AOANJRR) AOANJRR. Hip, Knee & Shoulder Arthroplasty Annual Report 2018. Adelaide: AOA; 2018. [Accessed April 24, 2019]
62. Xuan J, Kirchoerfer LJ, Boyer JG, Norwood GJ. Effects of comorbidity on health-related quality-of-life scores: an analysis of clinical trial data. *Clinical Therapeutics*. 1999 1999/02/01/;21(2):383-403.
63. Sprangers MAG, Schwartz CE. Integrating response shift into health-related quality of life research: a theoretical model. *Social Science & Medicine*. 1999 1999/06/01/;48(11):1507-15.
64. Razmjou H, Schwartz CE, Yee A, Finkelstein JA. Traditional assessment of health outcome following total knee arthroplasty was confounded by response shift phenomenon. *Journal of Clinical Epidemiology*. 2009;62(1):91-6.
65. Zhang X-H, Li S-C, Xie F, Lo N-N, Yang K-Y, Yeo S-J, et al. An Exploratory Study of Response Shift in Health-Related Quality of Life and Utility Assessment Among Patients with Osteoarthritis Undergoing Total Knee Replacement Surgery in a Tertiary Hospital in Singapore. *Value in Health*. 2012 2012/01/01/;15(1, Supplement):S72-S8.
66. Razmjou H, Yee A, Ford M, Finkelstein JA. Response shift in outcome assessment in patients undergoing total knee arthroplasty. *JBJS*. 2006;88(12):2590-5.
67. Schilling C, Petrie D, Dowsey MM, Choong PF, Clarke P. The Impact of Regression to the Mean on Economic Evaluation in Quasi-Experimental Pre–Post Studies: The Example of Total Knee Replacement Using Data from the Osteoarthritis Initiative. *Health economics*. 2017;26(12):e35-e51.

Acknowledgements:

A version of this paper was previously presented to the July 2019 meeting of the Health Economists' Study Group. We would like to thank Catherine Henderson and others present who commented on the paper.

Funding:

This study was supported by the Australian National Health and Medical Research Council (NHMRC) funded Centre for Research Excellence in Total Joint Replacement (1116325).

Michelle Dowsey holds a National Health and Medical Research Council of Australia Career Development Fellowship (1122526).

Peter Choong holds a National Health and Medical Research Council of Australia Practitioner Fellowship (1154203).

Peter Choong, Michelle Dowsey, Anne Smith and Philip Clarke are recipients of a National Health and Medical Research Council Centre for Research Excellence Grant in Total Joint Replacement (1116325).

Michelle Tew is jointly supported by the NHMRC funded Centre for Research Excellence in Total Joint Replacement (1116325) and Centre for Improving Cancer Outcomes Through Enhanced Infection Services (1116876), Melbourne Research Scholarship and Australian Research Council Centre of Excellence in Population Ageing Research.

Conflicts of Interest:

None reported for Philip Clarke, Kim Dalziel, Anne Smith and Michelle Tew. Dr. Dowsey reports grants from National Health & Medical Research Council, during the conduct of the study; grants from Medacta International, grants from Medibank Foundation and grants from MSK Australia outside of the submitted work. Professor Chong reports grants from National Health & Medical Research Council during the conduct of the study; fees from Stryker Corporation and Depuy, Johnson & Johnson outside of the submitted work.

Ethics Approval:

This study was approved by the St. Vincent's Hospital Melbourne Human Research Ethics Committee (HREC) (LNR/17/SVHM/136).

Consent to Participate:

This study involved the use of anonymised data extracted from a registry and was granted a waiver of informed consent.

Consent for Publication:

Not applicable

Availability of data and material:

Not applicable

Code availability:

Not applicable

Supplementary Material

Table S1: Comparison of fit statistics for models containing 1 to 8 classes

No. of classes	BIC	AIC	Min PP	Max PP	Proportion in each class (%)							
					1	2	3	4	5	6	7	8
1	4142.2	4155.6		1.00	100							
2	5842.0	5868.7	0.96	0.96	52.6	47.4						
3	6087.1	6127.2	0.85	0.94	31.7	35.1	33.2					
4	6149.8	6203.3	0.81	0.87	18.8	33.2	26.5	21.5				
5	6233.6	6300.4	0.76	0.88	17.8	10.9	32.6	15.7	22.9			
6 ^a	6279.0	6359.2	0.78	0.85	18.4	29.6	9.1	17.8	18.6	6.5		
7	6291.3	6384.9	0.69	0.82	15.1	9.4	25.8	17.6	15.2	10.9	6.1	
8	6299.5	6406.5	0.68	0.83	5.1	10.8	27.4	12.2	16.4	15.4	6.6	6.1

^a Model selected

Abbreviations: AIC, Akaike Information Criteria; BIC, Bayes Information Criteria; PP, average posterior probability of class membership

Table S2: Group membership probability diagnostics for models containing 1 to 8 classes

No. of classes	QoL trajectory #	No. in each trajectory	Average posterior probability	Odds of correct classification	Probability of group membership	Total probability
2	1	817	0.964	24.154	0.526	0.525
	2	736	0.963	28.631	0.474	0.475
3	1	492	0.899	19.248	0.317	0.315
	2	545	0.849	10.416	0.351	0.351
	3	516	0.935	28.835	0.332	0.333
4	1	292	0.839	22.482	0.188	0.191
	2	516	0.825	9.485	0.332	0.327
	3	411	0.807	11.606	0.265	0.265
	4	334	0.873	25.004	0.215	0.217
5	1	276	0.845	25.206	0.178	0.182
	2	170	0.763	26.190	0.109	0.120
	3	507	0.815	9.068	0.326	0.310
	4	244	0.776	18.564	0.157	0.162
	5	356	0.884	25.749	0.229	0.226
6 ^a	1	286	0.849	25.395	0.181	0.183
	2	481	0.815	9.601	0.315	0.301
	3	121	0.785	34.208	0.097	0.112
	4	279	0.817	19.505	0.186	0.180
	5	292	0.780	18.385	0.162	0.160
	6	94	0.822	72.466	0.060	0.064
7	1	234	0.818	25.283	0.151	0.152
	2	146	0.692	21.692	0.094	0.107
	3	400	0.726	7.653	0.258	0.240
	4	273	0.816	20.735	0.176	0.170
	5	236	0.757	17.362	0.152	0.148
	6	169	0.775	28.253	0.109	0.120
	7	95	0.819	69.333	0.061	0.063
8	1	79	0.825	88.222	0.051	0.063
	2	168	0.681	17.582	0.108	0.118
	3	425	0.777	9.264	0.274	0.248
	4	190	0.694	16.281	0.122	0.125
	5	254	0.821	23.381	0.164	0.161
	6	239	0.756	17.050	0.154	0.146

7	103	0.725	37.029	0.066	0.076
8	95	0.823	71.223	0.061	0.063

^a Model selected

Table S3: Patient profiles by QoL trajectories (Traj)

	Traj 1		Traj 2		Traj 3		Traj 4		Traj 5		Traj 6		P Value ^a
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	No.	(%)	
	286	18.42	481	30.97	121	7.79	279	17.97	292	18.80	94	6.05	
Patient characteristics													
Age, mean													<.001
(SD)	70.4	9.0	71.1	8.8	68.5	8.7	70.6	7.4	68.1	8.4	70.2	8.0	
Female	206	72.0	348	72.4	81	66.9	186	66.7	184	63.0	42	44.7	<.001
BMI, mean													
(SD)	33.7	6.7	33.0	6.1	32.9	5.8	32.9	5.5	32.3	5.8	30.7	4.5	0.001
Aetiology													0.55
Osteoarthritis	264	92.3	449	93.4	112	92.6	262	93.9	278	95.2	91	96.8	
Other ^b	22	7.7	32	6.7	9	7.4	17	6.1	14	4.8	3	3.2	
Bilateral													
surgery	57	19.9	118	24.5	28	23.1	84	30.1	68	23.3	27	28.7	0.101
Smoker													0.31
No	193	67.5	337	70.1	85	70.3	184	66.0	202	69.2	64	68.1	
Ex	66	23.1	108	22.5	26	21.5	82	29.4	75	25.7	24	25.5	
Yes	27	9.4	36	7.5	10	8.3	13	4.7	15	5.1	6	6.4	
Interpreter	66	23.1	91	18.9	16	13.2	32	11.5	21	7.2	5	5.3	<.001
CCI													<.001
0	127	44.4	252	52.4	64	52.9	158	56.6	209	71.6	65	69.2	
1+	159	55.6	229	47.6	57	47.1	121	43.4	83	28.4	29	30.9	
ASA													<.001
1/2	122	42.7	271	56.3	70	57.9	163	58.4	203	69.5	72	76.6	
3/4	164	57.3	210	43.7	51	42.2	116	41.6	89	30.5	22	23.4	
KL													0.16

<4	147	51.4	219	45.6	69	57.0	142	51.3	138	47.4	41	43.6	
4	139	48.6	261	54.4	52	43.0	135	48.7	153	52.6	53	56.4	
SEIFA deciles													0.37
1-5	93	32.5	174	36.2	38	31.4	110	39.4	108	37.0	39	41.5	
6-10	193	67.5	307	63.8	83	68.6	169	60.6	184	63.0	55	58.5	
Rural residence	27	9.4	64	13.3	20	16.5	62	22.2	66	22.6	24	25.5	<.001
Patient-reported outcomes													
KSS pain ^c													
Mild	9	3.2	24	5.0	8	6.6	23	8.2	8	2.7	16	17	<.001
Moderate	73	25.5	188	39.1	42	34.7	124	44.4	118	40.4	51	54.3	
Severe	204	71.3	269	55.9	71	58.7	132	47.3	166	56.9	27	28.7	
KSS function ^c , mean (SD)	26.7	19.8	32.9	19.7	38.0	20.2	40.1	20.1	40.2	18.4	52.2	17.3	P<.001
SF-12, mean (SD)													
PCS	21.8	7.2	22.8	7.5	23.5	7.4	22.9	8.3	22.1	7.6	31.4	9.3	P<.001
MCS	36.7	14.8	43.2	15.2	45.2	15.3	50.0	15.3	48.7	15.3	63.3	5.6	P<.001
Utility	0.49	0.08	0.55	0.09	0.56	0.10	0.59	0.09	0.56	0.08	0.81	0.07	P<.001

^a Comparisons across trajectory groups using chi2 test for categorical variables and one-way ANOVA for continuous variables

^b Other aetiology includes rheumatoid arthritis and avascular osteonecrosis.

^c KSS pain scores were categorised as follows; none (50), mild occasional (45) mild on stairs (40), mild on walking (30), moderate occasional (20), moderate continual (10) and severe pain as (0) points. KSS function score assesses walking, stair ability and use of walking aids and ranges for 0 to 100 with a higher score indicating better function.

Abbreviations: ASA, American Society of Anesthesiologist; BMI, body mass index; CCI, Charlson Co-morbidity Index; KL, Kellgren-Lawrence scale; MCS, mental component score; PCS, physical component score; REF, reference; RRR, relative risk ratio; SEIFA, Socio-Economic Index for Areas; KSS, Knee Society Score; Traj, trajectory.

Table S4: Univariable multinomial logistic regression showing relative risk (RRR) of belonging in each of the trajectory groups compared to Trajectory 5 (highest incremental QALYs)

Variable	Trajectory 1			Trajectory 2			Trajectory 3			Trajectory 4			Trajectory 6		
	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value	RRR	95% CI	P Value
Age	1.03	1.01-1.05	0.004	1.04	1.03-1.06	0.000	1.01	0.98-1.03	0.519	1.04	1.02-1.05	0.000	1.03	1.01-1.06	0.018
Female	1.53	1.07-2.19	0.019	1.50	1.10-2.06	0.011	1.14	0.72-1.80	0.569	1.17	0.82-1.66	0.390	0.49	0.30-0.79	0.003
BMI	1.04	1.01-1.07	0.004	1.02	1.00-1.05	0.083	1.02	0.99-1.06	0.206	1.02	1.00-1.05	0.098	0.95	0.91-0.99	0.007
Aetiology															
OA	1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a			1 [Reference] ^a		
Other ^f	1.50	0.75-3.04	0.254	1.36	0.71-2.62	0.358	1.42	0.59-3.43	0.433	1.13	0.54-2.37	0.753	0.59	0.16-2.15	0.421
Bilateral surg	0.84	0.56-1.25	0.387	1.14	0.80-1.61	0.463	1.06	0.63-1.77	0.825	1.42	0.97-2.08	0.073	1.34	0.79-2.28	0.280
Smoking status															
No	1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b			1 [Reference] ^b		
Ex	0.92	0.62-1.37	0.689	0.81	0.57-1.15	0.231	0.76	0.45-1.29	0.308	1.16	0.80-1.70	0.437	1.00	0.58-1.72	0.988
Yes	1.87	0.95-3.67	0.069	1.38	0.73-2.62	0.325	1.63	0.69-3.83	0.262	0.84	0.39-1.85	0.674	1.14	0.42-3.10	0.795
Interpr.	3.84	2.25-6.54	0.000	3.00	1.80-4.99	0.000	1.94	0.96-3.92	0.064	1.67	0.93-3.01	0.088	0.70	0.25-1.95	0.497
CCI															
0	1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c		
≥1	3.32	2.33-4.72	0.000	2.33	1.69-3.19	0.000	2.12	1.35-3.32	0.001	2.01	1.41-2.87	0.000	1.15	0.69-1.93	0.585

ASA																
score																
1/2	1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c			1 [Reference] ^c			
3/4	3.13	2.21-4.43	0.000	1.75	1.28-2.40	0.000	1.66	1.06-2.61	0.026	1.69	1.19-2.41	0.003	0.68	0.40-1.19	0.178	
KL																
<4	1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			
4	0.86	0.62-1.20	0.373	1.08	0.80-1.45	0.630	0.74	0.48-1.14	0.175	0.84	0.60-1.18	0.322	1.19	0.74-1.92	0.470	
SEIFA																
deciles																
1 to 5	1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			1 [Reference] ^d			
6 to 10	1.25	0.88-1.77	0.219	1.00	0.73-1.36	0.983	1.39	0.87-2.21	0.166	0.91	0.64-1.28	0.573	0.84	0.52-1.36	0.479	
Rural residence	0.35	0.21-0.56	0.000	0.51	0.34-0.74	0.001	0.60	0.34-1.03	0.076	0.90	0.60-1.35	0.618	1.10	0.64-1.91	0.725	
KSS pain																
Mild	1 [Reference] ^e			1 [Reference] ^e			1 [Reference] ^e			1 [Reference] ^e			1 [Reference] ^e			
Moderate	0.46	0.17-1.26	0.131	0.46	0.20-1.08	0.076	0.29	0.10-0.84	0.023	0.30	0.13-0.72	0.007	0.17	0.07-0.43	0.000	
Severe	0.95	0.35-2.55	0.912	0.46	0.20-1.07	0.071	0.36	0.13-1.01	0.052	0.23	0.10-0.55	0.001	0.07	0.03-0.18	0.000	
KSS function ^g	0.99	0.996-0.997	0.000	0.99	0.997-0.999	0.000	0.999	0.998-1.000	0.219	1.000	0.999-1.001	0.934	1.004	1.002-1.005	0.000	

^a Overall P Value = 0.64

^b Overall P Value = 0.20

^c Overall P Value < 0.001

^d Overall P Value = 0.26

^e Overall P Value = 0.005

^f Other aetiology includes rheumatoid arthritis.

^g Per point increase in KSS function score. The score assesses walking, stair ability and use of walking aids and ranges for 0 to 100 with a higher score indicating better function.

Abbreviations: ASA, American Society of Anesthesiologist; BMI, body mass index; CCI, Charlson Co-morbidity Index; Interpr., Interpreter; KL, Kellgren-Lawrence scale; RRR, relative risk ratio; SEIFA, Socio-Economic Index for Areas; KSS, Knee Society Score; Traj, trajectory.

Figure S1: Relative risk ratio with 95% confidence intervals (multivariable multinomial logistic regression) of being in each trajectory vs. reference trajectory (Trajectory 5) for each covariate in the model.

