Supplementary material

Trajectory Analysis

Repeated trajectory analyses were performed to identify the latent classes by changing the number of groups from 2 to 4, with the same starting values calculated from the 1-group model. According to Muthén & Muthén (2000), there are general criteria for selecting optimal numbers of latent classes. These criteria include comparing relative fit across models, examining the "quality" of classification across models, and interpreting usefulness of latent class trajectories.

A large number of fit indices can be used when choosing among models. Examples include the Akaike information criterion (AIC; Akaike, 1987), consistent AIC (CAIC; Bozdogan, 1987), Bayesian information criterion (BIC, Schwartz, 1978), the Lo-Mendell-Rubin likelihood ratio test (LMR; Lo, Mendell, & Rubin, 2001), the bootstrap parametric likelihood ratio test (BLRT; Muthén, 2006), and the approximation of integrated classification likelihood using BIC (ICLBIC; Biernacki, Celeux, & Govaert, 2000). Although all of these fit indices are common (Bauer & Curran, 2003a; Bozdogan, 1987; Jung & Wickrama, 2008; Nagin, 1999), there are strengths and limitations of each. For example, AIC tends to overestimate the true number of classes, whereas BIC and CAIC tend to underestimate (Bauer & Curran, 2003a; McLachlan & Peel, 2000). Furthermore, BIC is largely insensitive to sample size (D'Unger, Land, McCall, & Nagin, 1998), a possible strength of the method. The bootstrapped parametric likelihood ratio test, a newer method, often outperforms alternatives including BIC (Nylund, et al., 2007).

Quality of classification across models is measured using entropy (Connell & Frye, 2006). Entropy ranges from 0.0 to 1.0, with values closer to 1.0 representing better classification. It provides a summary measure of the probability of membership for each individual for the class the individual most likely belongs to (Connell & Frye, 2006).

In this study, KOOS-PS, VAS and EQ-5D-5L scores did not show linear or quadratic growth trajectories over time (Supplementary Figs 3-5); therefore, the growth mixture models(GMMs) were fitted in a two-step approach to identify potential heterogeneity of KOOS-PS, VAS and EQ-5D-5L, where factor loadings of the slope terms, as well as variances and covariances of continuous latent variables were free estimates across groups. In addition, the shapes and optimal number of groups were determined by the following criteria: 1) AIC, BIC and aBIC decreased at least 20; 2) Entropy>0.7; 3)% Participants per class >7.426 4) The p-value of LMR and BLRT are <0.05. Random starts were used to avoid convergence towards local maxima.The final models were described as

$$Y_{it}|_{C_i} = \eta_{0i} + \eta_{1i}\alpha_t + \varepsilon_{it}$$

where Y_{it} is the outcome for individual i at time t in latent class c of the latent class variable C; α_t are factor loadings, for the recognizability of the models, we set the factor loadings 7 days after TKA and 1 month after TKA to 0 and 1, respectively, and the factor loadings at other time points are freely estimated across groups; the random intercepts and random slopes are expressed as

$$\eta_{ji}|_{C_i=c} = \alpha_{jc} + \boldsymbol{\gamma_{jc}^T X_i} + \boldsymbol{\xi_{ji}}$$

where $j = 0,1$; $\boldsymbol{X_i}$ is a q-dimensional vector of time-invariant covariates.

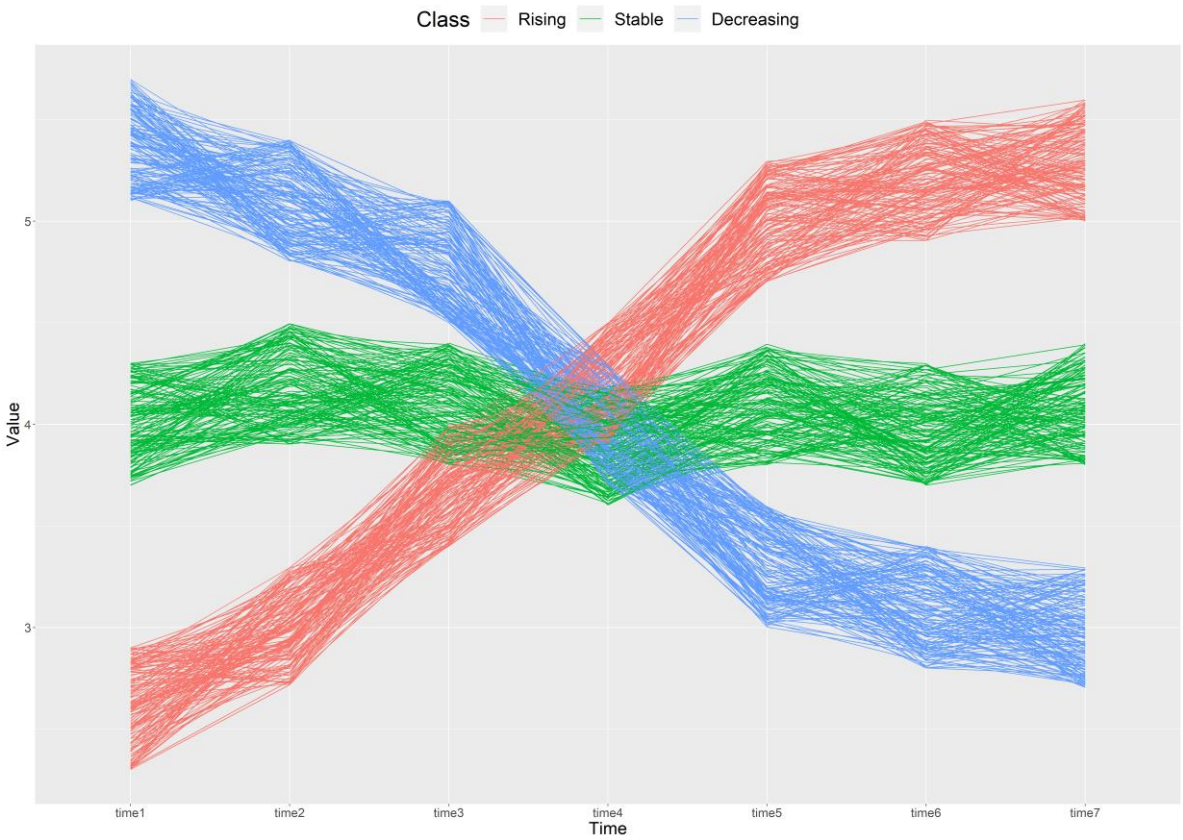

Figure S1. Schematic diagram of longitudinal data with group heterogeneity

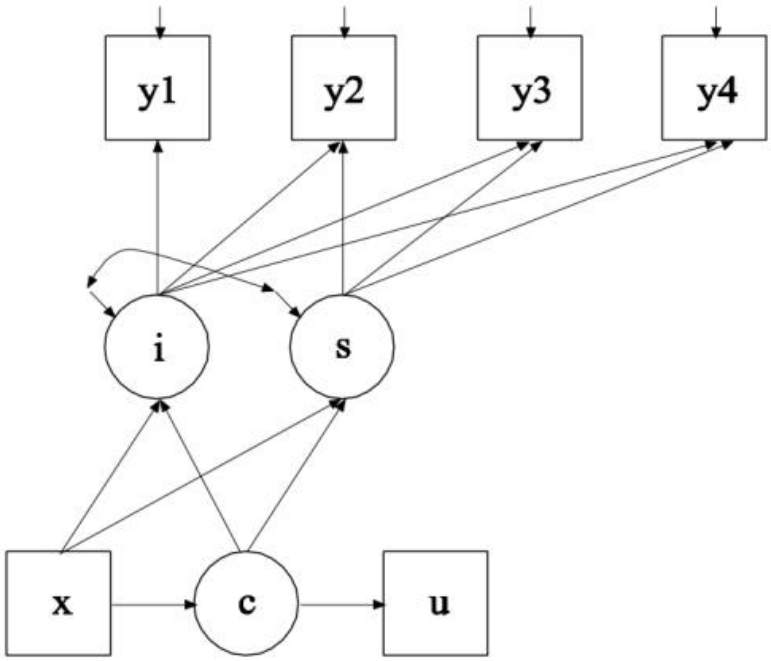

Figure S2. Schematic diagram of GMM with a categorical distal outcome

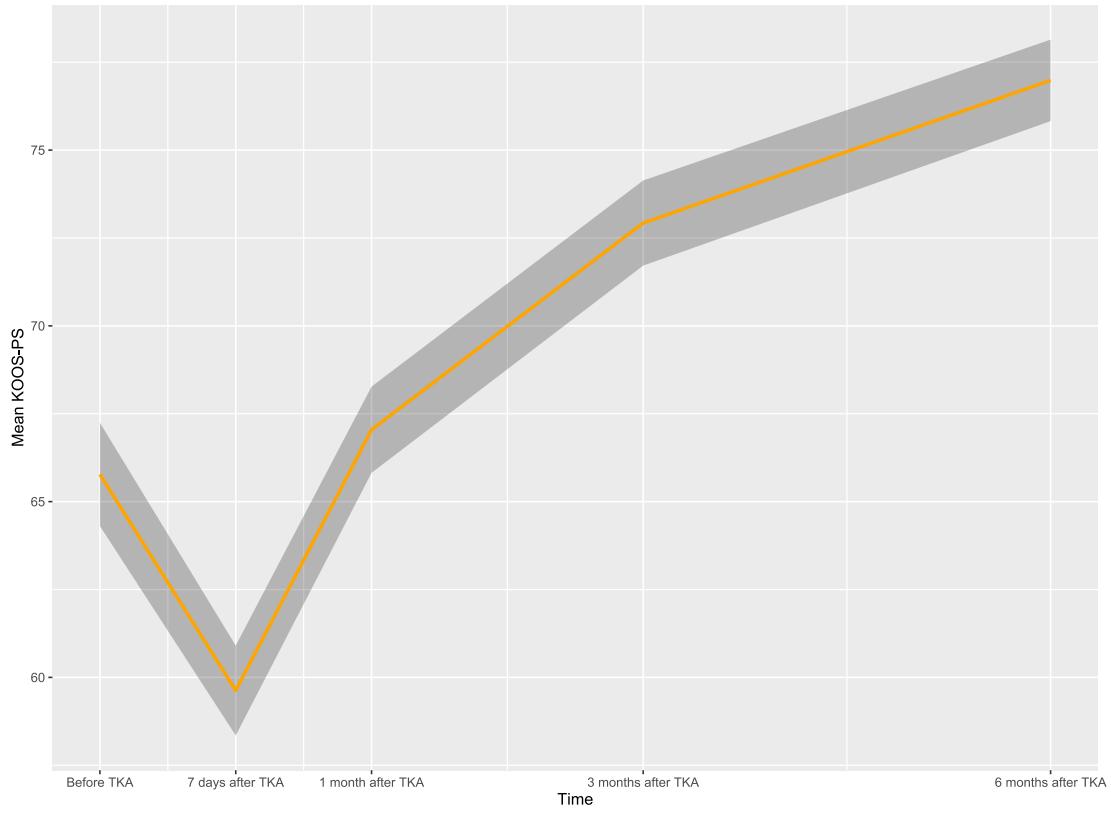

Figure S3. Mean growth trajectory of KOOS-PS

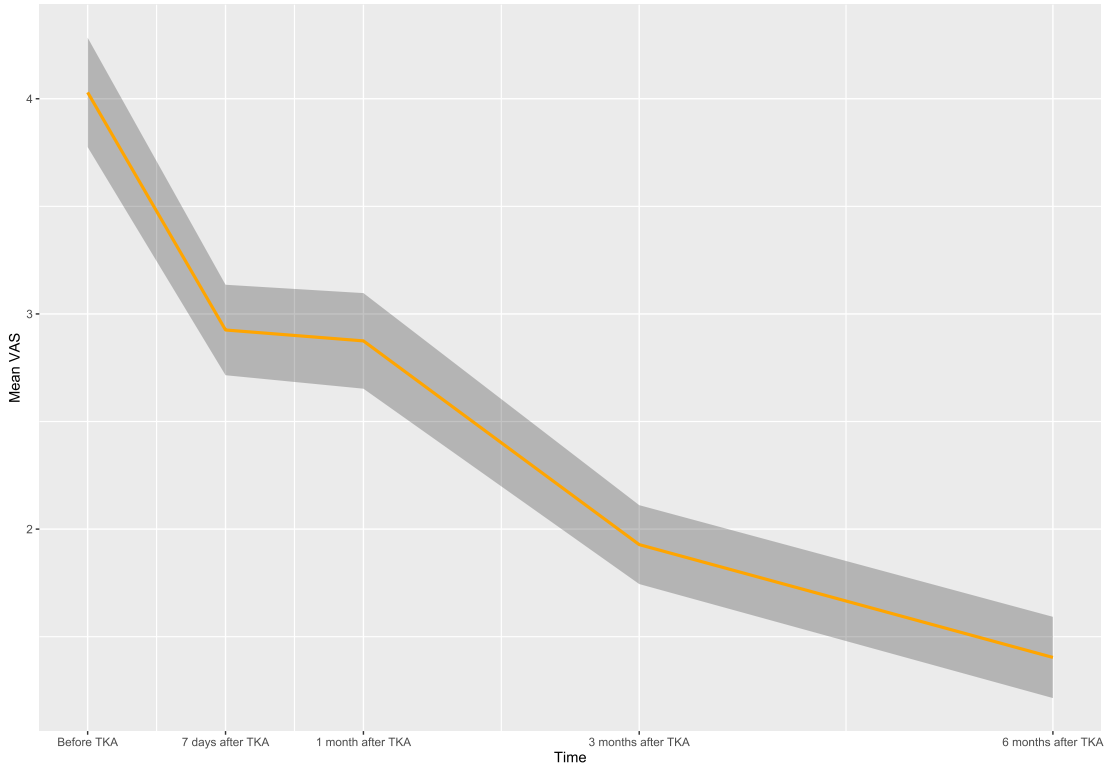

Figure S4. Mean growth trajectory of VAS

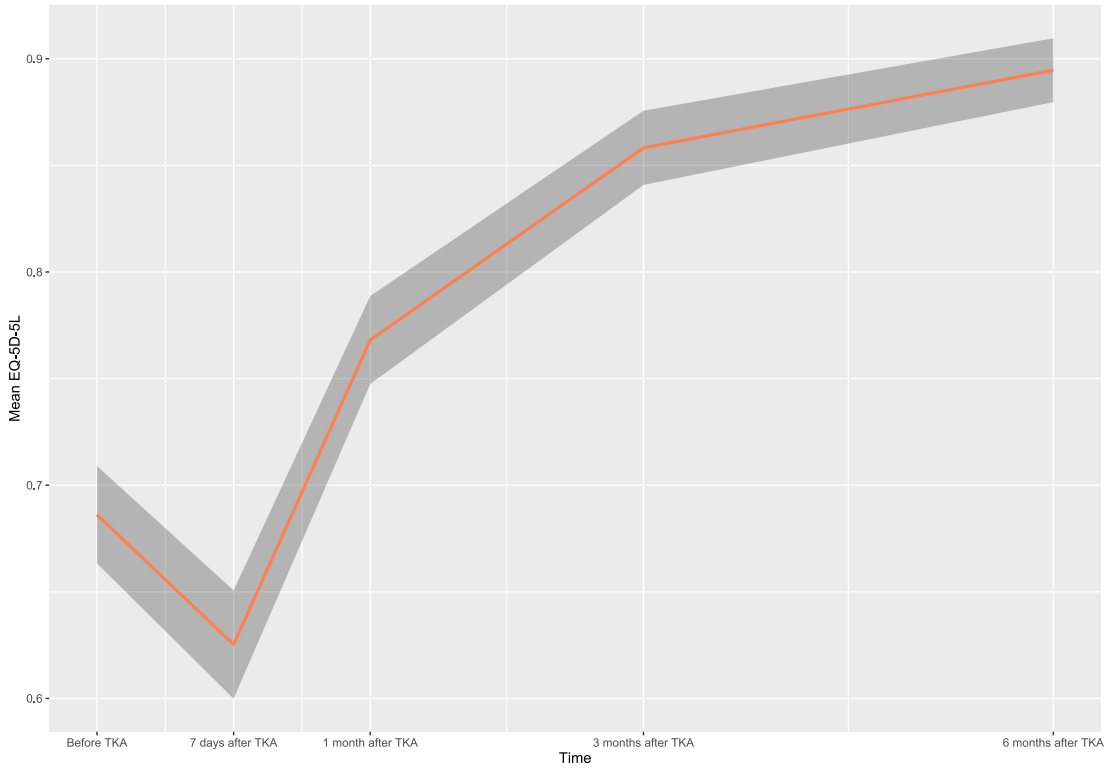

Figure S5. Mean growth trajectory of EQ-5D-5L

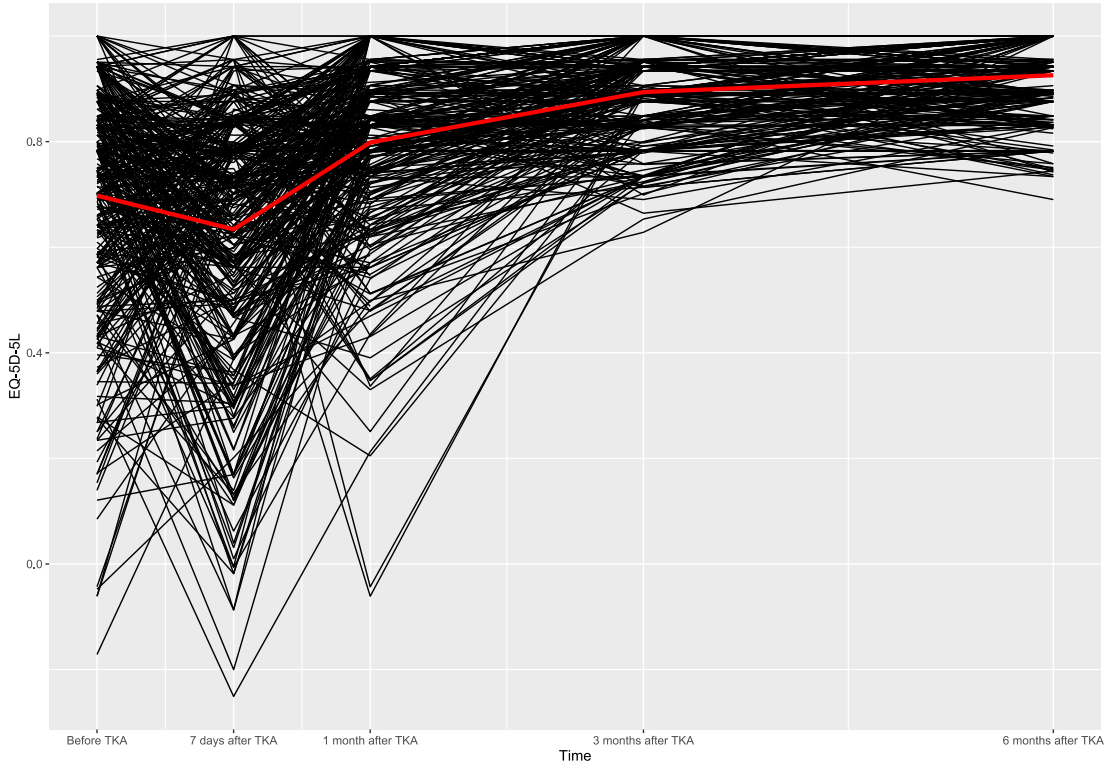

Figure S6. Growth trajectory of the "rising group" of EQ-5D-5L

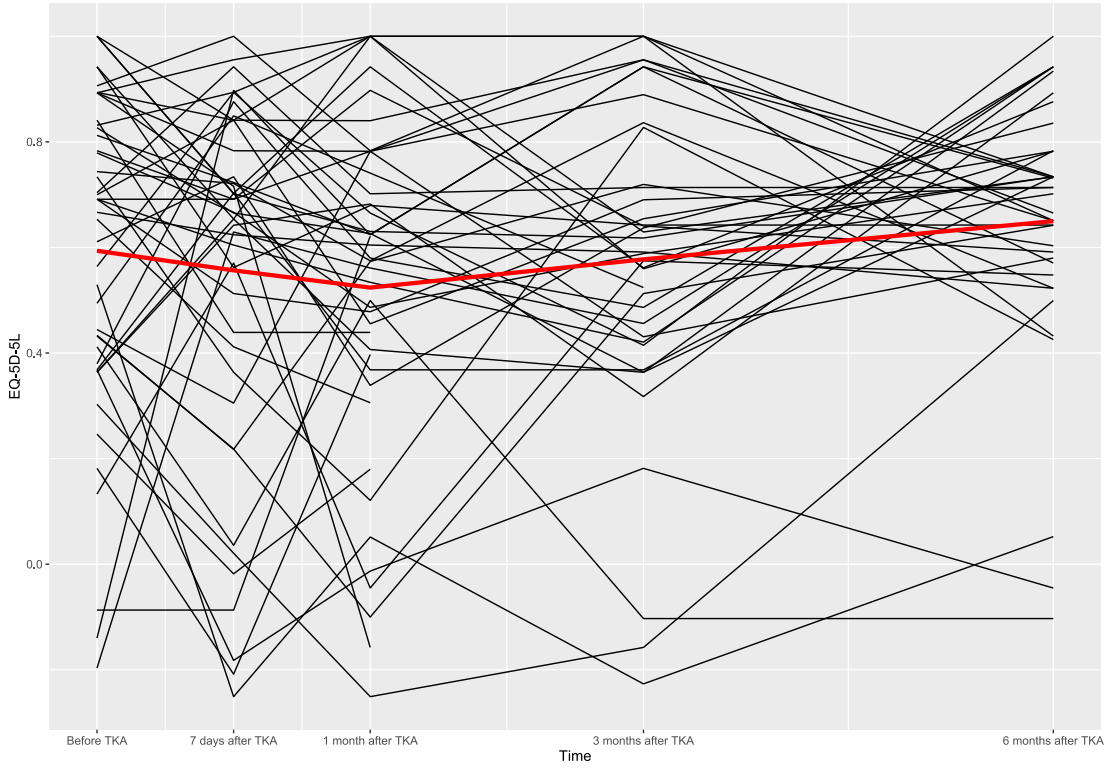

Figure S7. Growth trajectory of the " stable group" of EQ-5D-5L

Table S1. Results of GMM fit statistics with KOOS-PS as dependent variable (unconditional)

Models	AIC	BIC	aBIC	Entropy	LMR	BLRT	% Participants per class
Class-1	**14672.858**	**14724.877**	**14683.626**	-	-		-
Class-2	14577.944	14665.976	14596.167	0.994	0.022	<0.001	1.485/98.515
Class-3	14526.081	14650.125	14551.758	0.744	0.001	<0.001	1.980/23.762/74.257
Class-4	14503.06	14663.12	14536.199	0.597	0.214	0.333	15.842/2.228/25.

	7	3					495/56.436

AIC: Akaike information criterion; BIC: the Bayesian information Criterion; aBIC: Sample-Size Adjusted BIC; Entropy: Indicator for evaluating classification quality; % Participants per class: proportion of participants per class; The best fitting model is highlighted in bold characters.

Table S2. Results of GMM fit statistics with Pain-VAS as dependent variable (unconditional)

Models	AIC	BIC	aBIC	Entropy	LMR	BLRT	% Participants per class
Class-1	8049.318	8101.336	8060.085	-	-		-
Class-2	**7786.911**	**7874.943**	**7805.134**	**0.702**	**<0.001**	**<0.001**	**22.030/77.970**
Class-3	7733.279	7857.323	7758.957	0.652	0.064	<0.001	27.970/55.446/16.584
Class-4	7679.609	7839.665	7712.741	0.796	0.296	0.326	10.644/14.851/14.604/59.901

AIC: Akaike information criterion; BIC: the Bayesian information Criterion; aBIC: Sample-Size Adjusted BIC; Entropy: Indicator for evaluating classification quality; % Participants per class: proportion of participants per class; The best fitting model is highlighted in bold characters.

Table S3. Results of GMM fit statistics with EQ-5D-5L as dependent variable (with KOOS-PS_I, KOOS-PS_S and VAS_C as covariates)

Models	AIC	BIC	aBIC	Entropy	LMR	BLRT	% Participants per class
Class-1	-1408.208	-1356.190	-1397.440	-	-		-
Class-2	**-1719.877**	**-1591.832**	**-1693.372**	**0.841**	**<0.001**	**<0.001**	**11.139/88.861**
Class-3	-1823.518	-1635.451	-1784.587	0.868	0.059	<0.001	83.811/5.446/10.644
Class-4	-1920.748	-1672.661	-1869.394	0.766	0.240	<0.001	11.881/44.802/2.723/40.594

The growth mixture models(GMMs) were fitted in a two-step approach to identify potential heterogeneity of KOOS-PS, VAS and EQ-5D-5L, where factor loadings of the slope terms, as well as variances and covariances of continuous latent variables were free estimates across groups. First, two unconditional GMMs were fitted to KOOS-PS and VAS respectively. Next, A conditional GMM was fitted to the EQ-5D-5L score; where the KOOS-PS_I (the intercept term of KOOS-PS), KOOS-PS_S (the slope term of KOOS-PS) and VAS_C(the categorical latent variable of VAS) from the first step were entered as covariables; AIC: Akaike information criterion; BIC: the Bayesian information Criterion; aBIC: Sample-Size Adjusted BIC; Entropy: Indicator for evaluating classification quality; % Participants per class: proportion of participants per class;The best fitting model is highlighted in bold characters.

Table S4. Parameter estimation results for the two class model of EQ-5D-5L

		Estimated value	Standard error	t-value	p-value
Stable group（n=45）					
Intercept term	KOOS-PS_I	0.021	0.004	5.779	<0.001
	KOOS-PS_S	-0.022	0.017	-1.304	0.192
	VAS_C	-0.019	0.067	-0.283	0.779
	Mean residuals	-0.446	0.221	-2.015	0.044
	Residual variance	0.023	0.009	2.461	0.014
Slope term	KOOS-PS_I	0.002	0.003	0.647	0.518
	KOOS-PS_S	-0.007	0.009	-0.831	0.406
	VAS_C	-0.014	0.017	-0.855	0.397
	Mean residuals	-0.103	0.182	-0.569	0.569
	Residual variance	0.003	0.008	0.330	0.742
Residual covariance	Intercept with slope	0.000	0.004	-0.035	0.972
Rising group（n=359）					
Intercept term	KOOS-PS_I	0.030	0.003	11.617	<0.001
	KOOS-PS_S	-0.020	0.005	-4.389	<0.001
	VAS_C	0.006	0.018	0.359	0.72
	Mean residuals	-1.012	0.149	-6.776	<0.001
Slope term	KOOS-PS_I	-0.015	0.002	-7.531	<0.001
	KOOS-PS_S	0.020	0.004	5.197	<0.001

	VAS_C	0.009	0.011	0.840	0.401
	Mean residuals	0.955	0.121	7.920	<0.001
Residual covariance	Intercept with slope	0.002	0.002	1.194	0.232
Logistic regression	KOOS-PS_I	0.869	0.057	-2.278	0.023
	KOOS-PS_S	0.590	0.084	-4.906	<0.001
	VAS_C	7.389	0.336	5.952	<0.001

The correlation coefficient matrix for the latent variables in the rising group is not positive definite, resulting in negative residual variance estimates, so the residual variance of the intercept and slope terms in the rising group are not shown in this paper. Logistic regression: The category latent variables for EQ-5D-5L are the dependent variables (stable group is the reference group) and KOOS-PS_I, KOOS-PS_S and VAS_C are the independent variables. In the stable group (n = 45), The factor loadings of the slope term (EQ-5D-5L_S) at each time are 0.327, 0, 1, 2.115, and 2.809, respectively. In the rising group (n = 359), the factor loadings of the slope term (EQ-5D-5L_S) at each time are 0.487, 0, 1, 1.538, 1.715 respectively.