Changes in the forgotten joint score after total knee arthroplasty: Minimal clinical important difference, minimal important and detectable change

Journal:	*The Bone & Joint Journal*
Manuscript ID	BJJ-2020-0396.R1
Manuscript Type:	Original Article
Keywords:	Knee, Arthroplasty, Forgotten Joint Score, Minimal Important difference

Changes in the forgotten joint score after total knee arthroplasty: Minimal clinical important difference, minimal important and detectable change

Abstract

Aims To identify the minimal clinically important difference (MCID) minimal important change (MIC) and minimal detectable change (MDC) in the forgotten joint score (FJS) according to patient satisfaction 6 months following total knee arthroplasty (TKA).

Methods During a one-year period 484 patients underwent a primary TKA and completed preoperative and six-month FJS and OKS. At six-months patients were asked "How satisfied are you with your operated knee?" Their response was recorded as: very satisfied, satisfied, neutral, dissatisfied or very dissatisfied. The difference between patients recording neutral (n=45) and satisfied (n=151) was used to define the MCID. MIC for a cohort was defined as the change in the FJS for those patients declaring their outcome as satisfied, whereas receiver operating characteristic curve analysis was used to determine the MIC for an individual. Distribution based methodology was used to calculate the MDC.

Results Using satisfaction as the anchor the MCID for the FJS was 16.6 (95% confidence intervals (CI) 8.9 to 24.3, p<0.001) and when adjusting for confounding this decreased to 13.7 points (95% CI 4.8 to 22.5, p<0.001). The MIC for the FJS for a cohort of patients was 17.7 points and for an individual patient was 10 points. The MDC90 for the FGS was 12 points; where 90% of patients scoring more than this will have experienced a real change that is beyond measurement error.

Conclusion The estimates for MCID and MIC can be used to assess whether there is clinical difference between two groups and whether a cohort/patient has had a meaningful change in their FJS, respectively. The MDC90 of 12 points suggests a value lower than this may fall within measurement error.

Level of Evidence Level III, diagnostic study.

Introduction

The Forgotten Joint Score (FJS) is a patient reported outcome measure (PROM) which can be used to assess the outcome of total knee arthroplasty (TKA). The FJS was described by Behrend et al[1] in 2012 and is now an established and validated PROM, demonstrating low ceiling and floor effects and responsiveness to change.[2] This has resulted in the FJS being increasingly reported as a PROM after TKA in the orthopaedic literature.[3] Despite the increasing use of the English version of FJS the authors are not aware of an establish minimally clinically important difference (MCID), which is required to power studies and allow the FJS to be used as a primary outcome measure in clinical trials.

The MCID is the minimal difference in a scoring measure that is perceived by the patient to be beneficial or harmful relative to those that perceive no change.[4-6] This is different from the minimal important change (MIC) which is the change in the scoring measure for a cohort or individual patient that perceive their improvement to be minimal.[6] These definitions are often used interchangeably and can cause confusion in the literature.[6-8] Four previous European studies proposed MIC values for the FJS score after TKA of between 8 and 14 points; all of which did not use the English version of FJS.[9-12] The MCID is still to be defined for the FJS, as this is needed to power studies and assess whether an intervention has made a clinically important difference between two groups of patients. The minimal detectable change (MDC) can be defined the smallest change in an individual's FJS that is likely to be beyond the measurement error and represents a true change i.e. a change in the FJS less than this may be due to measurement error and not clinically relevant.

The primary aim of this study was to identify MCID, MIC and MDC in the FJS 6-months after TKA. The secondary aims were to assess (1) the effect of preoperative patient case-mix variables and preoperative functional status on the MCID and (2) whether the MCID changed according to patient specific activities and symptoms assessed.

Patients and methods

During a 1-year period (2014) 517 patients undergoing primary TKA completing pre and 6-month postoperative questionnaires were identified retrospectively from a TKA database held at the study centre. The patient demographics, BMI and comorbidities were recorded at the pre-operative assessment. Categories of comorbidity include: myocardial infarction (MI), heart failure, peripheral vascular disease, stroke, dementia, chronic obstructive airways disease (COPD), connective tissue disease, peptic ulcer, diabetes, kidney disease and backpain, which were all recorded as dichotomous variables of yes or no. Of the 517 patients 484 (93.6%) completed pre and postoperative questionnaires in their entirety without any missing data/scores. There was no significant difference in preoperative gender (p=0.677), age (p=0.712), BMI (p=0.999), comorbidities (p>0.151), FJS (p=0.871) or Oxford knee score (OKS) (p=0.972) between those fully completing (n=484) and not completing (n=33) their questionnaires.

Outcomes measured

The FJS assesses joint awareness during activities of daily living (for example, climbing stairs, walking for more than 15 minutes, in bed at night etc).[1] It consists of 12 questions assessed using a five-point Likert response format. Item scores are summed and linearly transformed to a 0 to 100 scale, a high value reflecting the ability of the patient to forget about the affected/replaced joint during the activities of daily living.

The OKS[13] was recorded preoperatively and at 6-months postoperatively. The OKS consists of twelve questions assessed on a Likert scale with values from 0 to 4. A summative score is then calculated where 48 is the best possible score (least symptomatic) and 0 is the worst possible score (most symptomatic). The MCID for the OKS is 5 points and is thought to represent a clinical difference between two groups of patients.[6]

The EuroQoL (EQ) general health questionnaire evaluates five domains (5D: assesses mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and was recorded preoperatively and at 6-months postoperatively.[14] The 3L version of the EuroQoL questionnaire was used, with the responses to the five domains being recorded at three levels of severity (no, slight problems, moderate, severe or unable/extreme problems). This index is on a scale of -0.594 to 1, where 1 represents perfect health, and 0 represents death. Negative values represent a state perceived as worse than death.[15]

Patient satisfaction was assessed by asking the question "How satisfied are you with your operated knee?". The response was recorded using a five-point Likert scale: very satisfied, satisfied, neither satisfied nor dissatisfied (simplified to neutral for the rest of manuscript), dissatisfied and very dissatisfied. Twelve further questions were asked relating to how their TKA has affected specific activities and symptoms which were rated using a five-point Likert scale: much worse, worse, the same, better and much better (Figure 1). These questions are based on the OKS questions.

MCID

The MCID was primarily defined according to patient satisfaction and secondarily by patient specific activities and symptoms. The MCID was calculated using two different methods: anchor-based. Using the anchor-based method the MCID was defined as the difference in the mean FJS change between patients responding with "satisfied" or "better" compared to those responding with "neutral" or "same" for level of satisfaction and patient specific activities and symptoms, respectively.[16, 17] Linear regression analysis was used to adjust for preoperative confounding variables to identify the MCID for the FJS.

MIC

The MIC for a cohort was defined as the change in the FJS for those patients declaring their outcome as satisfied. Receiver operating characteristic (ROC) curve

analysis was used to determine the MIC for an individual and was defined as the threshold value in the FJS that was predictive of patient satisfaction.

MDC

A distribution methodology based on the standard error of measurement was used to calculate the MDC90. The 90 indicates a 90% confidence interval that a change greater than this is real and not due to intrinsic variability of the FJS. The standard error of measurement (the range in which a patient's true score lies) is the error associated with the measuring tool. The standard error of measurement was calculated using the standard deviation (SD) for the change in the FJS (from the study cohort) and the reliability of the FJS: standard error of measurement = SD x √1-reliability. A previously established Cronbach's alpha of 0.97 for test re-test reliability for the FJS was used.[18, 19] The MDC was then calculated by multiplying the standard error of measurement by √2 (representing two separate occasions in which to measure change) and by a z value which represents the chosen confidence intervals (CI). To establish the 90% CI a value of 1.65, hence: MDC90 = standard error of measurement x √2 x 1.65.

Statistical analysis

Statistical Package for Social Sciences version 17.0 (SPSS Inc., Chicago, IL, USA) was used for all data analysis. Data was assessed for normality and parametric tests were appropriate. Linear variables were assessed using either unpaired Student's t-test, or one-way analysis of variance (ANOVA) with correction for multiple testing (Tukey). A Chi square test was used to assess gender and comorbidity differences between groups. ROC curve analysis was used to identify a threshold (point of maximal sensitivity and specificity) in the mean FJS change that was predictive of patient satisfaction. The area under the ROC curve (AUC) ranges from 0.5, indicating a test with no accuracy in distinguishing whether a patient is satisfied, to 1.0 where the test is perfectly accurate identifying all satisfied patients. To adjust for

confounding preoperative variables influencing the MCID linear regression analysis was used identify an adjusted MCID. Significance was set as a p-value of <0.05.

Ethics

Ethical approval was obtained from the regional ethics committee (Research Ethics Committee, South East Scotland Research Ethics Service, Scotland [16/SS/0026]) for analysis and publication of the presented data. The data collection was carried out in accordance with the GMC guidelines for good clinical practice and the Declaration of Helsinki.

Results

Study cohort characteristics

The study cohort consisted of 484 patients undergoing TKA with complete pre and postoperative data that met the inclusion criteria. This included 218 (45%) male patients and 266 (55%) female patients, with an overall mean age of 70.0 (range 27 to 91) years. There was a greater improvement in the FJS with increasing level of patient satisfaction at 6 months (Figure 2), but this was not significantly different between neutral, dissatisfied and very dissatisfied patients (Table I).

MCID using patient satisfaction as the anchor

There were 153 (31.6%) patients that declared they were satisfied and 44 (9.1%) patients that were neutral. There was no significant difference in the demographics, comorbidities or the preoperative EQ-5D and FJS between these groups, but the preoperative OKS score was significantly worse in the neutral group (Table II).The unadjusted MCID for the FJS was 16.6 (95% confidence intervals (CI) 8.9 to 24.3, p<0.001). When adjusting for confounding (all preoperative variables from Table II) using regression analysis the MCID for the FJS decreased to 13.7 (R^2=0.45, 95% CI 4.8 to 22.5, p<0.001).

MCID using patient specific activities and symptoms as the anchor

When the MCID was defined using patient specific activities and symptoms it varied from 1.9 points for general knee pain to 26.6 points for kneeling and getting up again

(Table III). However, the MCID was not statistically different for four of the activities/symptoms assessed. The ability to do household shopping was the question with the lowest MCID (7.2 points) that was also statistically significantly (p=0.021) different between the groups. Using this same question (household shopping) the MCID for the OKS was 4.9 (95% CI 2.2 to 7.5, p<0.001).

MIC estimate for single group over time

The mean change, pre to postoperative, in the FJS associated with a response of satisfied was 17.7 (95% CI 14.2 to 21.2) and was defined as the MIC (Table IV).

MIC estimate for individual patients

There were 416 (86%) patients that were satisfied with their outcome. The individual MIC was identified using ROC analysis to identify satisfied patients from dissatisfied patients according to change in their FJS. The change in the FJS was demonstrated to be a reliable and significant predictor of patient satisfaction with an AUC of 82% (95% CI 77 to 87, p<0.001) (Figure 3). The maximal point of sensitivity and specificity for predicting satisfaction, which was 77%, corresponded to a change in the FJS for 10 points or more (Figure 4).

MDC90

The standard deviation of the mean postoperative FJS was 30 points. In test retest reliability was defined as 0.97 (Cronbach's alpha coefficient).[18, 19] The standard error of measurement was 5.2 (standard deviation [30] multiplied by the square root of one minus the test retest reliability [0.173]). The MDC90 was then calculated to be 12 points (SEM x 1.41 x 1.65) i.e. 90% patient patients scoring more than this will have experienced a real change that is beyond measurement error.

Discussion

This defines the MCID, MIC and MDC for the English version of FJS score after TKA. Preoperative variables were shown to influence the MCID, using the anchor-based question of patient satisfaction, and when adjusted for the MCID reduced to 13.7 points. The MCID for the FJS was also found to vary between 1.9 to 26.6 points

depending on the focus of the anchor question used relating to patient specific activates and symptoms. The MIC for the FJS for a cohort of patients was 17.7 points and for an individual was 10 points i.e. a change greater than this represents a real change recognised by the group or individual, respectively. The MDC90 of 12 points suggests a value lower than this may fall within measurement error.

There are several limitations to the current study that relate to the relatively early timepoint of assessment (6-months) and using patient satisfaction as the anchor question. Hamilton et al[18] demonstrated a 4 point improvement in the FJS between 6- and 12-months, which could potentially influence the identified MCID from the current study. However, previous studies assessing the MCID in the OKS have found the same MCID at 6- and 12-months[6, 20], despite a similarly recognised improvement in the OKS over that time period.[21] Loth et al[22] recently assessed patient satisfaction at one-year and the associated FJS after TKA; from their published data their MCID would be 13.6 points at one year which is similar the 13.7 points identified in the current study after adjusting for confounding at 6-months. This suggests there may not be a change in the MCID between 6- and 12-months, but this should be assessed in future studies. Patient satisfaction changes with time and patient factors can also influence the rate of this change[23, 24], using satisfaction as the anchor question to define the MCID predisposes this to the same influences. However, the same argument could be made for whatever anchor question was used. The ideal anchor question for the FJS would be one that assesses the scores underlying psychometric construct of the patient's "awareness" of their joint. Joint awareness has been suggested to be an overarching value that encompasses pain and function in those with an arthroplasty and as such patient satisfaction is likely to be a good surrogate marker.[25, 26] In addition, the way patient satisfaction was measured was also a limitation of the study as the different between neutral and satisfied patients may not represent the "minimal" difference. Using the joint awareness question as the anchor with responses such as "same", "a little better",

"somewhat better" and "much better" may be more optimal to define the "minimal" difference between those that are the "same" from the those "a little better". A disadvantage of using such a response is that relatively few patients record their outcome as the "same" or "a little better" and this will limit the patient numbers available for analysis.[10, 12] However, those patients who declared their level of satisfaction as dissatisfied or neutral in the current study had similar changes in their FJS, which may indicate that they were similar group of patients and could be combined for analytical purposes should number be limited.

There have been three previous studies, using the Danish, German and Italian versions of the FJS, that have defined the MIC after TKA.[9-11] Ingelsrud et al[10] used three methods to define the MIC which included predictive modelling and ROC curve analysis and found the MIC in the FJS to be 14 and 17 points, respectively. These values support the MIC of 17.7 points in the FJS suggested in the current study for a cohort of patients. Baumann et al[9] and Sansone et al[11] both used a simple rule of thumb method to estimate the MIC by using half of the SD of the FJS and demonstrated values for the postoperative MIC of 11 and 12 points, respectively. If the same rule of thumb was applied to the current cohort of patients with a SD of 30 points for change in the FJS at 6-months the MIC would have been 15 points, which is similar to defined MIC of 17.7 points using the mean change method. As the MDC90 was found to be 12 in the current study a score less than this may represent measurement error of the FJS and therefore a MIC of less than this may not be a reliable measure. However, if lower SD of 25 was used to calculate the MDC90, which has been reported buy two previous studies[9, 11], it would be reduced to 10 points.

Seven of the 12 patient specific activity and symptoms questions assessed in the current study demonstrated a MCID in the FJS after TKA lower than 13.7 points, which was the estimated MCID according satisfaction. The MCID in the FJS after TKA may therefore be lower than the 13.7 points identified in the current study.

However, for these seven questions only three demonstrated a statistically significantly difference of which the ability to do household shopping which had the lowest MCID of 7.2 points for the FJS. Using this same question the MCID in the OKS was found to be 4.9 points which is the currently accepted MCID after TKA.[6, 20] It is recognised that the MCID does change according to the focus of the anchor question and will also likely vary according to the number of responses to the chosen anchor question assessed.[20] the authors suggest the satisfaction derived MCID value of 13.7 points is the most reasonable estimate. However, to use a MCID of 7 points for the FJS after TKA, due to the high SD associated with the FJS (30 points in the current study), would result in an effect size of 0.23. This would translate into more than 580 patients needing to be randomised to power future clinical studies (Table IV). Whereas a MCID of 13.7 points in the FJS increases the effect size to 0.46 and reduces the number of patients required to power future studies.

The identified MCID may be specific for TKA, as previous authors have shown that patients undergoing medial unicompartmental knee arthroplasty (UKA) have a greater postoperative FJS relative to TKA patients.[3, 27] However, this may reflect the higher rate of patient satisfaction with UKA compared to TKA.[3] The MCID may also be language dependant with a previous study demonstrating significant differences in the improvement in the FJS after TKA at one year between Swiss and Scottish patients.[28] This suggests that the response trajectories and improvement values demonstrated in previous European studies may not directly translate to UK arthroplasty patients and UK derived values are from the current study are required.

Conclusion

The estimates for MCID and MIC can be used to assess whether there is clinical difference between two groups and whether a cohort/patient has had a meaningful change in their FJS, respectively. The MDC90 of 12 points suggests a value lower than this may fall within measurement error.

Competing Interest Statement

The authors declare no conflict of interest with the content of this study.

References

1. **Behrend H, Giesinger K, Giesinger JM, Kuster MS** The "forgotten joint" as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty* 2012;27:430-6.

2. **Adriani M, Malahias MA, Gu A, Kahlenberg CA, Ast MP, Sculco PK** Determining the Validity, Reliability, and Utility of the Forgotten Joint Score: A Systematic Review. *J Arthroplasty* 2019.

3. **Clement ND, Bell A, Simpson P, Macpherson G, Patton JT, Hamilton DF** Robotic-assisted unicompartmental knee arthroplasty has a greater early functional outcome when compared to manual total knee arthroplasty for isolated medial compartment arthritis. *Bone Joint Res* 2020;9:15-22.

4. **Jaeschke R, Singer J, Guyatt GH** Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407-15.

5. **King MT** A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res* 2011;11:171-84.

6. **Beard DJ, Harris K, Dawson J, Doll H, Murray DW, Carr AJ, Price AJ** Meaningful changes for the Oxford hip and knee scores after joint replacement surgery. *J Clin Epidemiol* 2015;68:73-9.

7. **Crosby RD, Kolotkin RL, Williams GR** Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395-407.

8. **Lydick E, Epstein RS** Interpretation of quality of life changes. *Qual Life Res* 1993;2:221-6.

9. **Baumann F, Ernstberger T, Loibl M, Zeman F, Nerlich M, Tibesku C** Validation of the German Forgotten Joint Score (G-FJS) according to the COSMIN checklist: does a reduction in joint awareness indicate clinical improvement after arthroplasty of the knee? *Arch Orthop Trauma Surg* 2016;136:257-64.

10. **Ingelsrud LH, Roos EM, Terluin B, Gromov K, Husted H, Troelsen A** Minimal important change values for the Oxford Knee Score and the Forgotten Joint Score at 1 year after total knee replacement. *Acta Orthop* 2018;89:541-7.

11. **Sansone V, Fennema P, Applefield RC, Marchina S, Ronco R, Pascale W, Pascale V** Translation, cross-cultural adaptation, and validation of the Italian language Forgotten Joint Score-12 (FJS-12) as an outcome

measure for total knee arthroplasty in an Italian population. *BMC Musculoskelet Disord* 2020;21:23.

12. **Holtz N, Hamilton DF, Giesinger JM, Jost B, Giesinger K** Minimal important differences for the WOMAC osteoarthritis index and the Forgotten Joint Score-12 in total knee arthroplasty patients. *BMC Musculoskelet Disord* 2020;21:401.

13. **Dawson J, Fitzpatrick R, Murray D, Carr A** Questionnaire on the perceptions of patients about total knee replacement. *J Bone Joint Surg Br* 1998;80:63-9.

14. **Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X** Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011;20:1727-36.

15. **Scott CEH, MacDonald DJ, Howie CR** 'Worse than death' and waiting for a joint arthroplasty. *Bone Joint J* 2019;101-B:941-50.

16. **Cook CE** Clinimetrics Corner: The Minimal Clinically Important Change Score (MCID): A Necessary Pretense. *J Man Manip Ther* 2008;16:E82-E83.

17. **Gerlinger C, Schmelter T** Determining the non-inferiority margin for patient reported outcomes. *Pharm Stat* 2011;10:410-3.

18. **Hamilton DF, Loth FL, Giesinger JM, Giesinger K, MacDonald DJ, Patton JT, Simpson AH, Howie CR** Validation of the English language Forgotten Joint Score-12 as an outcome measure for total hip and knee arthroplasty in a British population. *Bone Joint J* 2017;99-B:218-24.

19. **Thompson SM, Salmon LJ, Webb JM, Pinczewski LA, Roe JP** Construct Validity and Test Re-Test Reliability of the Forgotten Joint Score. *J Arthroplasty* 2015;30:1902-5.

20. **Clement ND, MacDonald D, Simpson AH** The minimal clinically important difference in the Oxford knee score and Short Form 12 score after total knee arthroplasty. *Knee Surg Sports Traumatol Arthrosc* 2014;22:1933-9.

21. **Browne JP, Bastaki H, Dawson J** What is the optimal time point to assess patient-reported recovery after hip and knee replacement? A systematic review and analysis of routinely reported outcome data from the English patient-reported outcome measures programme. *Health Qual Life Outcomes* 2013;11:128.

22. **Loth FL, Giesinger JM, Giesinger K, Howie CR, Hamilton DF** Single-item satisfaction scores mask large variations in pain, function and joint awareness in patients following total joint arthroplasty. *Eur J Orthop Surg Traumatol* 2020;30:267-74.

23. **Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ** Three groups of dissatisfied patients exist after total knee arthroplasty: early, persistent, and late. *Bone Joint J* 2018;100-B:161-9.

24. **Clement ND, Bardgett M, Weir D, Holland J, Gerrand C, Deehan DJ** The rate and predictors of patient satisfaction after total knee arthroplasty are influenced by the focus of the question. *Bone Joint J* 2018;100-B:740-8.

25. **Loth FL, Liebensteiner MC, Giesinger JM, Giesinger K, Bliem HR, Holzner B** What makes patients aware of their artificial knee joint? *BMC Musculoskelet Disord* 2018;19:5.

26. **Hamilton DF, Lane JV, Gaston P, Patton JT, MacDonald DJ, Simpson AH, Howie CR** Assessing treatment outcomes using a single question: the net promoter score. *Bone Joint J* 2014;96-B:622-8.

27. **Zuiderbaan HA, van der List JP, Khamaisy S, Nawabi DH, Thein R, Ishmael C, Paul S, Pearle AD** Unicompartmental knee arthroplasty versus total knee arthroplasty: Which type of artificial joint do patients forget? *Knee Surg Sports Traumatol Arthrosc* 2017;25:681-6.

28. **Giesinger JM, Giesinger K, Federico B, Howie CD, Hamilton DF** Differences in case mix and outcomes between Swiss and Scottish total knee arthroplasty patients. *Knee Surg Sports Traumatol Arthrosc* 2019.

Tables

Table I. Differences in the mean change (95% CI) in the FJS according to level patient satisfaction 6 months after TKA for the study cohort.

Versus	Level of Satisfaction				
	Very Satisfied	**Satisfied**	**Neutral**	**Dissatisfied**	**Very Dissatisfied**
Very Satisfied	-	29.8 (22.1 to 37.5) p<0.001	46.3 (34.0 to 58.7) p<0.001	45.4 (27.4 to 63.3) p<0.001	37.3 (3.2 to 71.5) p=0.024
Satisfied	-29.8 (-37.5 to -22.1) p<0.001	-	16.6 (3.7 to 29.5) p=0.004	15.6 (-2.8 to 34.0) p=0.140	7.5 (-26.8 to 42.0) p=0.974
Neutral	-46.3 (-58.7 to -34.0) p<0.001	-16.6 (-29.5 to -3.7) p=0.004	-	-1.0 (-21.7 to 19.8) p=0.999	-9.0 (-44.7 to 26.7) p=0.958
Dissatisfied	-45.4 (-63.3 to -27.4) p<0.001	-15.6 (-34.0 to 2.8) p=0.140	1.0 (-19.8 to 21.7) p=0.999	-	-8.0 (-46.0 to 30.0) p=0.978
Very Dissatisfied	-37.3 (-71.5 to -3.2) p=0.024	-7.5 (-42.0 to 26.8) p=0.974	9.0 (-26.7 to 44.7) p=0.958	8.0 (-30.0 to 46.0) p=0.978	-

p-values are for an ANOVA with Tukey correction for multiple testing

Table II. Patient demographics and pre-operative functional scores according group.

Demographic	Descriptive	Satisfaction		Odds ratio/ *Difference*	95% CI		p-value*
		Neutral (n=44)	Satisfied (n=153)		Lower	Upper	
Gender (M/F)	Male	19	74	OR 0.81	0.41	1.59	0.544
(n, % of group)	Female	25	79				
Mean Age (years: mean, SD)		69.8 (9.8)	69.0 (9.8)	*Diff 0.8*	-2.5	4.1	0.626**
BMI (kg/m^2: mean, SD)		30.4 (6.1)	30.8 (5.9)	*Diff 0.4*	-1.7	2.5	0.690**
Comorbidity	Myocardial Infarction	6	7	OR 0.30	0.10	0.96	0.033
(n, % of group)	Heart Failure	1	2	OR 0.57	0.05	6.43	0.570***
	Vascular disease	1	3	OR 0.86	0.09	848	0.897***
	Stroke	1	1	OR 0.28	0.02	4.62	0.345***
	Dementia	1	1	OR 0.28	0.02	4.62	0.345***
	COPD	3	6	OR 0.56	0.13	2.33	0.417***
	Connective tissue	7	21	OR 0.84	0.33	2.13	0.715
	Peptic ulcer	1	3	OR 0.86	0.09	8.48	0.897***
	Diabetes	9	14	OR 0.39	0.16	0.98	0.040
	Kidney disease	2	2	OR 0.28	0.04	2.03	0.180***
	Back pain	18	73	OR 1.32	0.67	2.60	0.425
Functional measures (mean, SD)							
Oxford Score	Pre-operative	21.5 (8.2)	21.0 (7.9)	*Diff 0.6*	-2.1	3.3	0.674**
EQ5D	Pre-operative	0.437 (0.321)	0.421 (0.312)	*Diff 0.016*	-0.090	0.122	0.762**
FJS	Pre-operative	14.1 (21.1)	11.0 (12.3)	*Diff 3.1*	-3.6	9.8	0.359**

*chi square test unless otherwise stated, **unpaired t-test, ***Fishers exact test

Table III. The MCID according to the anchor question used. The MCID was defined as the difference between those experiencing no change in their symptoms and those defining their outcome as better.

Descriptive	Response	Mean change in FJS (SD)	MCID	95% CI		p-value*
				Lower	Upper	
General knee pain	Same (n=31)	9.8 (23.7)	1.9	-6.5	10.3	0.658
	Better (n=148)	11.7 (21.1)				
Washing and drying yourself	Same (n=117)	15.2 (27.1)	4.5	-2.4	11.5	0.204
	Better (n=104)	19.8 (25.5)				
Getting into and out of cars and / or public transport	Same (n=100)	12.3 (28.8)	16.3	9.6	23.0	<0.001
	Better (n=168)	28.5 (25.7)				
Walking	Same (n=74)	7.0 (22.0)	12.3	6.2	18.5	<0.001
	Better (n=161)	19.3 (22.4)				
Getting up from a table	Same (n=105)	10.2 (21.9)	16.4	10.2	22.6	<0.001
	Better (n=187)	26.7 (27.8)				
Limping	Same (n=62)	11.5 (19.3)	5.2	-1.8	12.3	0.144
	Better (n=155)	16.7 (25.3)				
Kneeling and getting up again	Same (n=167)	23.5 (31.3)	26.6	18.9	34.3	<0.001
	Better (n=100)	50.1 (30.0)				
Pain in bed at night	Same (n=70)	11.8 (29.9)	11.0	3.8	19.0	0.003
	Better (n=160)	22.8 (23.8)				
Your usual work	Same (n=100)	9.8 (24.6)	13.8	7.8	19.9	<0.001
	Better (n=167)	23.6 (24.0)				
Your knee giving way suddenly	Same (n=58)	16.5 (26.2)	0.7	-7.4	8.7	0.868
	Better (n=138)	15.8 (26.1)				
Doing household shopping	Same (n=112)	13.8 (24.2)	7.2	1.1	13.3	0.021
	Better (n=137)	21.1 (24.3)				
Walking downstairs	Same (n=93)	13.0 (27.0)	15.6	8.8	22.3	<0.001
	Better (n=160)	28.6 (25.8)				

* unpaired t-test

Table IV. Mean pre and postoperative FJS and change in the score according to the patients level of satisfaction with their outcome at 6-month follow TKA.

Level of Satisfaction	**Mean preoperative FJS** (SD)	**Mean postoperative FJS** (SD)	**Mean change in FJS** (95% CI)	**p-value***
Very Satisfied (n=263)	10.0 (10.5)	57.5 (29.6)	47.5 (43.7 to 51.3)	<0.001
Satisfied (n=153)	11.0 (12.3)	28.8 (18.4)	17.7 (14.2 to 21.2)	<0.001
Neutral (n=44)	14.1 (21.1)	15.3 (14.8)	1.1 (-6.7 to 9.0)	0.769
Dissatisfied (n=19)	6.9 (6.4)	9.0 (12.8)	2.1 (-4.0 to 8.3)	0.477
Very Dissatisfied (n=5)	8.3 (5.7)	18.5 (24.0)	10.1 (-15.2 to 35.4)	0.328

*Paired t-test

Table V. Power calculations for total sample size according to MCID and required power. A two-way analysis for two independent groups (means) with a standard deviation of 30 and an alpha of 0.05 was used in all calculations.

Power	**MCID**		
	7	13.7	14
80%	n=580	n=154	n=148
85%	n=663	n=176	n=168
90%	n=774	n=204	n=196

Figure Legends

Figure 1. Sample to the portion of the 6-month patient questionnaire asking whether specific activities and symptoms have been affected after their TKA.

Figure 2. Mean change in the FJS at 6 months following TKA according to level of patient satisfaction (error bars represent 95% confidence intervals).

Figure 3. Receiver operating curve predicting satisfied patients using the Forgotten Joint Score. Area under the curve 82% (95% CI 77 to 87, p<0.001).

Figure 4. Sensitivity and specificity plot for predicting satisfied from neutral patients after TKA according to change in their FJS.

Please tell us how your procedure has affected the following:

	Much worse	Worse	The same	Better	Much better
General knee pain	☐	☐	☐	☐	☐
Washing and drying yourself	☐	☐	☐	☐	☐
Getting into and out of cars and/or public transport	☐	☐	☐	☐	☐
Walking	☐	☐	☐	☐	☐
Getting up from a table	☐	☐	☐	☐	☐
Limping	☐	☐	☐	☐	☐
Kneeling and getting up again	☐	☐	☐	☐	☐
Pain in bed at night	☐	☐	☐	☐	☐
Your usual work (including housework)	☐	☐	☐	☐	☐
Your knee giving way suddenly or' letting you down'	☐	☐	☐	☐	☐
Doing household shopping	☐	☐	☐	☐	☐
Walking down stairs	☐	☐	☐	☐	☐

Figure 1. Sample to the portion of the 6-month patient questionnaire asking whether specific activities and symptoms have been affected after their TKA.

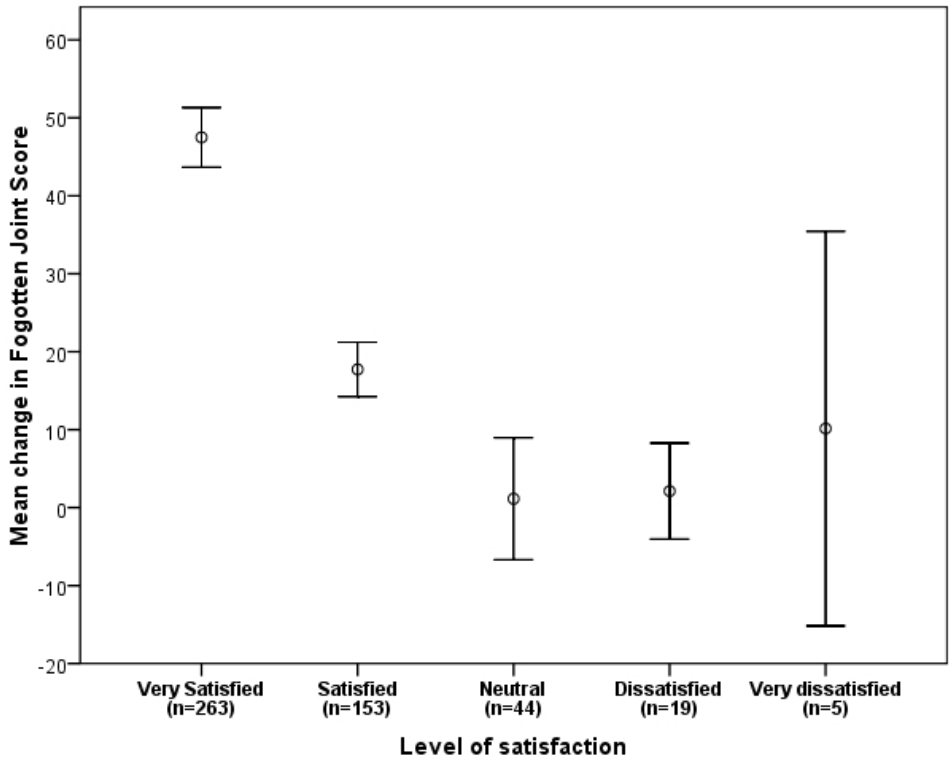

Figure 2. Mean change in the FJS at 6 months following TKA according to level of patient satisfaction (error bars represent 95% confidence intervals).

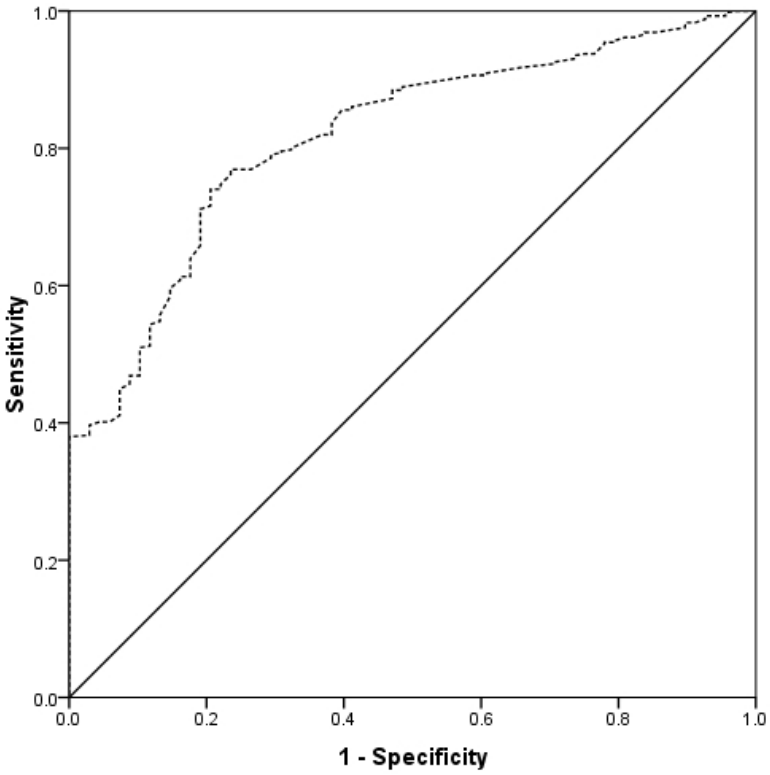

Figure 3. Receiver operating curve predicting satisfied patients using the Forgotten Joint Score. Area under the curve 82% (95% CI 77 to 87, p<0.001).

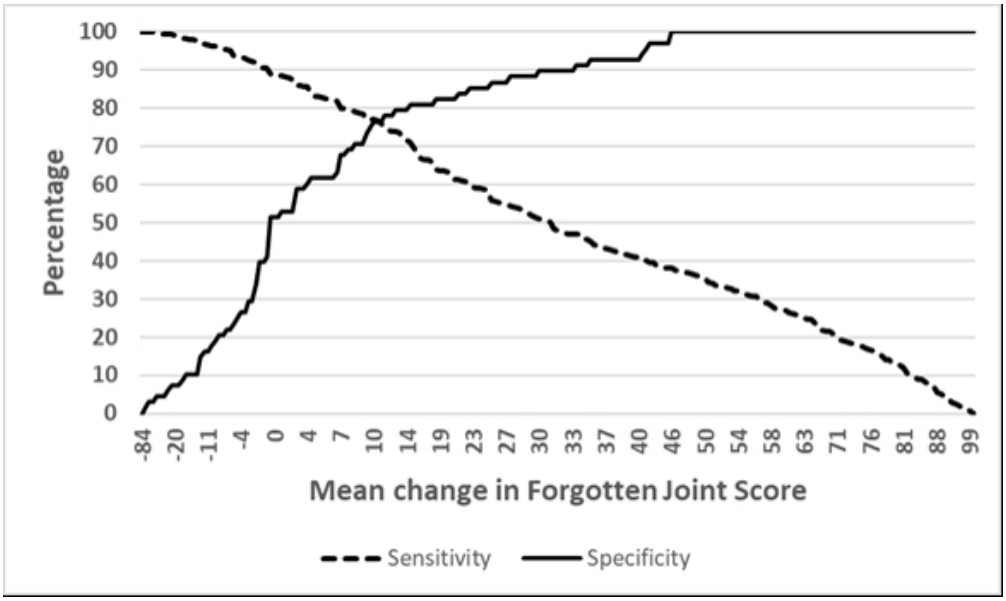

Figure 4. Sensitivity and specificity plot for predicting satisfied from neutral patients after TKA according to change in their FJS.