Ingelsrud, L. H., Terwee, C. B., Terluin, B., Granan, L.-P., Engebretsen, L., Mills, K. A. G., Roos, E. M. (2018). Meaningful Change Scores in the Knee Injury and Osteoarthritis Outcome Score in Patients Undergoing Anterior Cruciate Ligament Reconstruction. *American Journal of Sports Medicine, 46*, 1120-1128.

TITLE PAGE

Title

Meaningful Change Scores in the Knee Injury and Osteoarthritis Outcome Score in Patients Undergoing Anterior Cruciate Ligament Reconstruction

Corresponding author

Ingelsrud, Lina Holm

Postal address: Department of Orthopaedic surgery, Hvidovre Hospital, Kettegaard Alle 30, 2650 Hvidovre, Denmark

E-mail: lina.holm.ingelsrud@regionh.dk

Telephone number: +45 60893375

Authors

Ingelsrud Lina Holm, PT, M.Sc., PhD student, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark AND Department of Orthopaedic Surgery, Copenhagen University Hospital Hvidovre, Copenhagen, Denmark

Terwee Caroline B., PhD, Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam Public Health research institute, Amsterdam, Netherlands

Terluin Berend, MD, PhD, Department of General Practice and Elderly Care Medicine, VU University Medical Center, Amsterdam, Netherlands

Granan Lars-Petter, MD, PhD, Department of Pain Management and Research, Division of Emergencies and Critical Care, Oslo University Hospital, Oslo, Norway

Engebretsen Lars, MD, Professor, Oslo Sports Trauma Research Center, Norwegian School of Sport Sciences, Oslo, Norway AND Department of Orthopaedic Surgery, Oslo University Hospital and Faculty of Medicine, University of Oslo, Norway

Mills Kathryn A. G., PT, PhD, Department of Health Professions, Macquarie University, Sydney, Australia

Roos Ewa M., PT, Professor, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

The study was performed at the Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

Acknowledgements

The study was funded by the Danish Rheumatism Association. We thank the staff at the Norwegian Knee Ligament Registry for help and guidance with data collection. Statistician Thomas Kallemose, Hvidovre Hospital, is acknowledged for his support with the statistical analyses.

ABSTRACT

Background

Meaningful change scores in the Knee injury and Osteoarthritis Outcome Score (KOOS) in patients undergoing anterior cruciate ligament (ACL) reconstruction have not yet been established.

Purpose

This study aimed to define the Minimal Important Change (MIC) for the KOOS after ACL reconstruction.

Study design: Prospective cohort study

Methods

KOOS scores and anchor questions with 7-point scales ranging from “better, an important improvement” to “worse, an important worsening” were completed post-operatively by randomly chosen participants from the Norwegian Knee Ligament Registry. Pre-surgery KOOS scores were retrieved from the registry. The MIC for improvement was calculated with anchor-based approaches using the predictive modeling method adjusted for the proportion of improved patients, the mean change method, and the receiver operating characteristic (ROC) method.

Results

Complete data for at least one of the KOOS subscales were obtained from 542 (45.3%) participants. Predictive modeling MIC values were 12.1 for the KOOS subscales of Sport and Recreational function and 18.3 for the knee-related Quality Of Life. These values aid in interpreting within-group improvement over time and can be used as responder criteria when comparing groups. The corresponding and much lower values for the subscales of Pain (2.5), Symptoms (-1.2) and Activities of Daily Living (2.4) are the results from patients reporting on average only mild problems with these domains pre-operatively. Although 4 to 10% of patients reported subscale-specific worsening, MIC deterioration calculations were not possible. The ROC MIC values were associated with high degrees of misclassification. Values obtained by the mean

change method were considered less reliable because these estimates are derived from subgroups of patients. Average KOOS change scores were approximately similar for patients reporting acceptable symptoms post-operatively and patients reporting important improvements on the anchor items after surgery.

Conclusion

KOOS users should apply subscale-specific cut-offs for meaningful improvement. Our results confirm using the subscales of Sport and Recreational function and knee-related Quality Of Life as primary patient-reported outcomes following ACL reconstruction. The predictive modeling approach gave the most robust estimates of MIC values. Our data suggests that reporting acceptable symptoms postoperatively corresponds to reporting an important improvement after ACL reconstruction.

Clinical relevance

Not applicable

Key terms

Minimal Important Change, Knee injury and Osteoarthritis Outcome Score, Anterior Cruciate Ligament Reconstruction

What is known about the subject

The statistical significance of between-group change score comparisons is often used to determine treatment effect in RCT studies. However, a statistically significant change score is not necessarily clinically meaningful or meaningful to patients. Meaningful scores are context-specific and no single cut-off can necessarily be applied for a questionnaire across contexts. Investigating interpretability characteristics in patient-reported outcome measures is crucial to improve the evaluation of treatment effect in research and clinical practice.

What this study adds to existing knowledge

The study concludes that MIC values are KOOS subscale-specific, and presents specific MIC values for the KOOS for use 6-24 months after an ACL reconstruction. The presented MIC values are intended to be used for interpretation of within-group evaluation of treatment effect and between-group comparisons through responder analyses. Additionally, the study highlights some careful considerations to make when investigating MIC values.

MAIN TEXT

INTRODUCTION

The Knee injury and Osteoarthritis Outcome Score (KOOS) is a frequently used disease-specific Patient Reported Outcome Measure (PROM) for measuring knee symptoms, function and quality of life in patients with anterior cruciate ligament (ACL) injury undergoing ACL reconstruction. The KOOS is available in ~50 languages and KOOS scores are monitored in national registries and international databases, and used in research and clinical practice.[3,10,17] The 37 studies evaluating KOOS' psychometric properties until January 2014 were summarized in a systematic review concluding adequate content validity, internal consistency, test-retest reliability, construct validity and responsiveness for age- and condition-relevant subscales.[1] How to interpret KOOS scores is however not straightforward, because a statistically significant change score is not necessarily clinically relevant or meaningful to patients.[24] The concept of Minimal Important Change (MIC), also known as minimal (clinically) important difference, has been termed and defined in a variety of ways in scientific literature. In recent years, emphasis has been placed on MIC being the smallest change in PROM score that is considered important by patients.[25]

There is limited knowledge about meaningful change scores for the KOOS. In 2003, Roos et al. suggested that a change exceeding 8-10 points represented a clinically important improvement.[17] However, interpretability characteristics of a PROM are considered to be context-specific.[6,9] For the KOOS, MIC values have been investigated for patients receiving rehabilitation after a total knee replacement[13] and non-operative treatment for knee osteoarthritis.[12,20] No previous studies have investigated meaningful change scores for the KOOS in patients undergoing an ACL reconstruction.

The purpose of this study, therefore, was to define MIC values for the KOOS after undergoing an ACL reconstruction, intended to be used for interpreting longitudinal change in KOOS subscale scores within one group of patients, or between groups with responder analysis.

METHODS

Study design and setting

We designed a prospective cohort study using registry data and an additional survey mailed to patients included in the Norwegian Knee Ligament Registry (NKLR). The NKLR gathers nationwide data on patients undergoing an ACL reconstruction. All public and private hospitals performing ACL reconstruction participated in the registry and voluntary compliance has been >85% since 2006.[4] The NKLR received approval from the Norwegian data inspectorate as an expansion of the Norwegian Arthroplasty Register concession.[3] Postoperative results on the patients' perceptions of treatment outcome have been published previously.[7]

Participants

Data were collected in 2012. A questionnaire was sent to 1197 randomly selected patients who were at three different follow-up timepoints after surgery: 397 at 6 (range 5-7) months, 400 at 12 (range 10-14) months and 400 at 24 (range 20-28) months post-operatively. At 6 months post-operatively, there were only 397 patients eligible for random data extraction. Patients who had undergone bilateral or revision ACL reconstruction were excluded. Individual patients' pre-operative data were retrieved from the NKLR database.

Variables/Questionnaires

An explanatory letter, paper-based questionnaires and a prepaid envelope were mailed to the patients. Reminders were sent after 2 months. Questionnaires included the KOOS and a set of anchor questions.

The KOOS includes five subscales: Pain, Symptoms, Activities of Daily Living (ADL), Sport and Recreational function (Sport/Rec), and knee-related Quality Of Life (QOL). Each KOOS item is scored from 0 to 4 and the total score for each separate subscale is transformed into 0 (worst) to 100 (best) scales. The 2012 rule for

handling missing items was used, allowing calculation of subscale scores when >50% of these subscale items were answered.[16]

When determining meaningful change values, self-reported anchor questions are considered optimal to best capture patient perspectives of important changes in health status.[2,8,9] Domain-specific questions asking for the importance of the change experienced within separate domains such as pain, symptoms, function and quality of life are suggested to improve the anchor's validity.[23] Thus, we used five anchor questions, one for each KOOS subscale. Patients rated their perceived importance of the experienced change on seven-point scales ranging from *"Worse, an important worsening"* to *"Better, an important improvement"*. Two additional anchor questions were asked: 1) whether or not patients had achieved acceptable symptoms, and if not, 2) whether they considered the treatment to have failed.[7] Both questions were answered "yes" or "no" (Appendix, Section 1).

Statistics

Patient demographics were presented as means with 95% confidence intervals (95% CI) around the mean for continuous variables and n (%) for categorical variables.

All analyses were performed on separate KOOS subscales. Patients were excluded from MIC analyses if a subscale score was missing pre-operatively or post-operatively or if the corresponding anchor question was missing.

The anchors' validity was evaluated with Spearman's correlation coefficients between the KOOS change scores and the respective subscale-specific anchors. Due to inconsistency in the literature and several MIC methods applied,[15,26] no pre-defined correlation level was set prior to performing the analyses.

Statistical analyses were performed with R (version 3.2.1, R-project.org).

Anchor-based MIC methods

Anchor-based methods involve anchoring the PROM change score to an external measure of important change such as a domain-specific anchor question. Several anchor-based MIC analyses have been proposed, using different statistical approaches to estimate the optimal cut-off for MIC.[2,9,15] In this study, the primary method was predictive modeling (MIC_{pred}) because it has been shown to be more precise compared to the frequently used Receiver Operating Characteristic (ROC) method.[22] Simulations have shown that both the MIC based on ROC analysis and the MIC_{pred} identify the mean of the hypothesized latent individual MICs in a sample when the proportion of improved patients is 50% and the scores are normally distributed. Both MICs will be biased when the proportion of improved patients differs from 50%, which is the case after ACL reconstruction where a larger proportion commonly report improvement. However, predictive modeling allows for the adjustment for proportions improved other than 50%[21]. To enable comparison to more traditional methods, we also applied the Mean Change method ($MIC_{MeanChange}$)[8] and the ROC method (MIC_{ROC}).[27] Detailed descriptions of these MIC calculation methods are presented in the appendix. MIC analyses were performed on pooled data from the 6, 12 and 24 months post-operative time-points.

KOOS mean change scores for patients reporting acceptable symptoms or treatment failure

Mean KOOS change scores were calculated for patients responding “yes” to having ‘acceptable symptoms’, and if not, “yes” to considering the treatment to have failed. If patients answered “no” to having achieved ‘acceptable symptoms’ and “no” to ‘treatment failure’, they were categorized to an ‘undecided’ intermediate group.

Sensitivity analyses

To investigate the impact of time from surgery, adjusted MIC_{pred} analyses were stratified for the 6-, 12- and 24-month follow-up subgroups. Additionally, since previous studies have presented baseline dependent MIC values,[26] we investigated the effect of pre-operative severity. Preoperative KOOS scores were included as interaction terms in the MIC_{pred} analyses and considered to be effect modifiers if p-values were <0.05.

RESULTS

Participants

Of the total 1197 randomly selected patients, 744 (62.3%) replied. Of those who replied, 202 (27.2%) were excluded from further analyses due to lack of any pre-operative KOOS subscale score or not answering any KOOS subscale or anchor questions post-operatively. As a result, the number of patients differed between subscales (Figure 1).

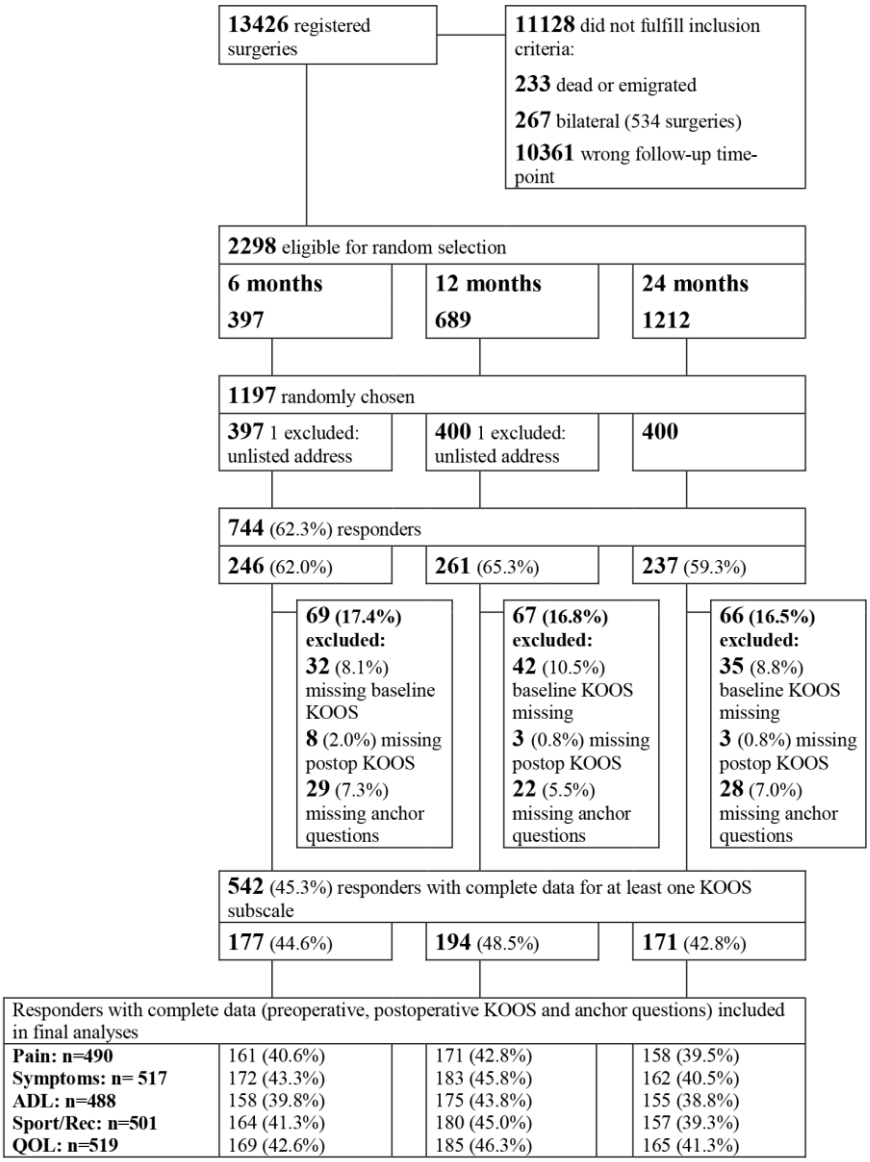

Figure 1. Study flow diagram.

Baseline characteristics and descriptive data

Included patients had a mean (SD) age of 29.9 (11.6) years and 52.6% were women. Responders with complete data were older and more were female than the responders without complete data and non-

responders combined. However, confidence intervals around the mean pre-surgery KOOS scores overlapped widely (Table 1).

Table 1. Pre-operative characteristics for responders with complete data for at least one KOOS subscale, responders without complete data and non-responders

	Responders with complete dataset for at least one KOOS subscale n=542[1]	Responders without complete dataset n=202[1]	Non-responders n=453[1]
	Mean ± SD[n]	Mean ± SD[n]	Mean ± SD[n]
Female n (%)	285 (52.6)	97 (48.0)	158 (34.9)
Age	29.9 (28.9;30.9)	28.5 (26.8; 30.1)	27.4 (26.5; 28.3)
KOOS Pain	72.4 (70.7;74.1)[515]	74.4 (69.9;78.8)[89]	71. 6 (69.5;73;8)[349]
KOOS Symptoms	71.1 (69.5;72.7)[530]	74.0 (70.2;77.8)[91]	71.4 (69.4-73.3)[359]
KOOS ADL[2]	81.7 (80.0;83.4)[508]	81.4 (76.9;85.9)[87]	80.2 (78.1;82.3)[346]
KOOS Sport/Rec[3]	40.2 (37.8;42.5)[518]	46.4 (40.0;52.9)[87]	42.6 (39.8;45.4)[351]
KOOS QOL[4]	33.9 (32.4;35.4)[535]	35.3 (30.8;39.8)[91]	35.2 (33.3;37.0)[366]

[1] The number of patients included in each analysis varies due to degree of missing data. The actual numbers included in the KOOS analyses are presented as [n] for each mean (95% CI) calculated.
[2] ADL: Activities of Daily Living
[3] Sport/Rec: Sport and Recreational function
[4] QOL: knee-related Quality Of Life

The percentages of patients reporting being importantly improved ranged from 71.3-78.7% and unchanged from 15.1-22.3%, across the five KOOS subscales(Figure 2).

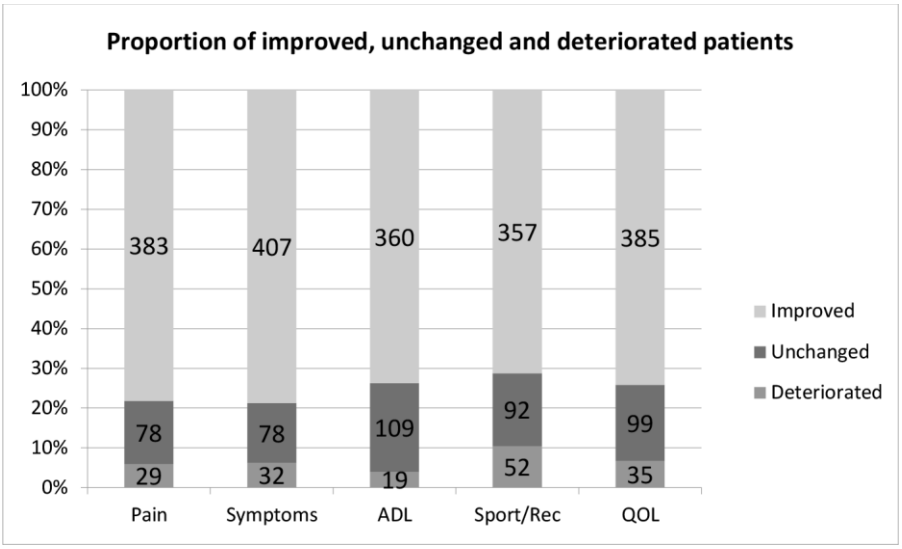

Figure 2. Percentages of improved, unchanged or deteriorated patients for each KOOS subscale are given on the y-axis. Numbers given in each column represent the number of patients reporting to be improved, unchanged or deteriorated.

The percentages of deteriorated patients ranged from 3.9-10.4%. Due to the comparatively small number of deteriorated patients, MIC deterioration analyses were not performed. KOOS mean change scores for each anchor response category ranging from better to worse are presented in Figure 3.

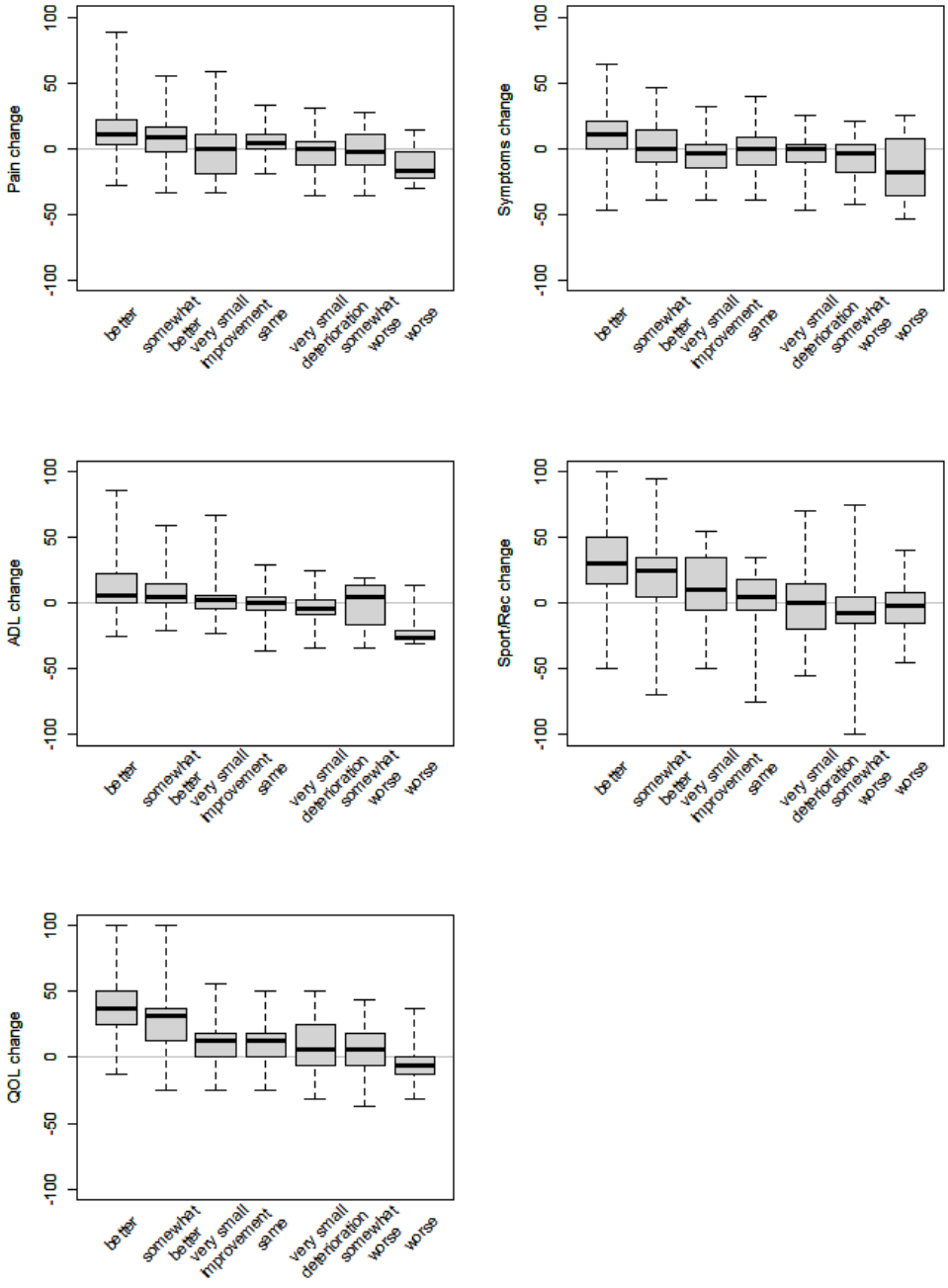

Figure 3: KOOS mean change scores by anchor question response category ranging from better to worse. Horizontal bars represent the median, the box represents the interquartile range and the whiskers represent the highest and lowest scores.

MIC improvement values

The correlations between anchor questions and KOOS change scores were 0.53 for QOL, 0.41 for Sport/Rec, 0.39 for Symptoms and 0.32 for Pain and ADL.

MIC_{pred} improvement (95% CI) values were 12.1 (9.3-14.8) for Sport/Rec and 18.3 (16.0-20.6) for QOL (Figure 4 and Table 2), when calculated using pooled data from patients at 6, 12 and 24 months post-operatively, and adjusted for the proportions of improved patients. The corresponding values for Pain, Symptoms and ADL were 2.5 (0.4; 4.5), -1.2 (-3.2; 0.8) and 2.4 (0.7; 4.1).

Sensitivity analyses performed separately on the 6, 12, and 24 months post-operative scores resulted in non-significant differences of less than 2.4 points in adjusted MIC_{pred} values (Appendix, Section 3). Furthermore, preoperative KOOS scores interacted non-significantly with KOOS change in the predictive modeling analyses (data not shown), suggesting no baseline dependency of MIC_{pred} values.

$MIC_{MeanChange}$ values were higher than MIC_{pred} values for all subscales (Figure 4 and Table 2). Due to generally flattened ROC curves and low areas under the curves, selecting the best ROC cut-off point was highly affected by minor chance fluctuations in the sample (Appendix, Section 4).

Table 2. MIC improvement values obtained by predictive modeling adjusted for percentages of improved patients, and as obtained by the Mean Change method.

KOOS	Predictive modeling MIC		Mean Change MIC	
	MIC_{pred}[1]	95% CI[2]	$MIC_{MeanChange}$	95% CI[3]
Pain	**2.5**	0.4; 4.5	7.9	4.8; 11.1
Symptoms	**-1.2**	-3.2; 0.8	1.2	-1.7; 4.1
ADL[4]	**2.4**	0.7; 4.1	8.1	4.9; 11.2
Sport/Rec[5]	**12.1**	9.3; 14.8	21.7	17.3; 26.2
QOL[6]	**18.3**	16.0; 20.6	27.3	24.3; 30.3

[1] MIC_{pred} values are adjusted for the proportion of improved patients
[2] 95% Confidence Intervals (CI) calculated using 1000 bootstrap replications, reported as 0.025-0.975 quantiles
[3] 95% CI calculated as $Mean_{change} \pm 1.96\left(\frac{SD_{change}}{\sqrt{n}}\right)$
[4] ADL: Activities of Daily Living
[5] Sport/Rec: Sport and Recreational function
[6] QOL: knee-related Quality Of Life

MIC values in comparison to KOOS mean change scores for patients with acceptable symptoms and treatment failure

To put MIC_{pred} values in context, we compared them to other longitudinal and cross-sectional determinants of outcome following ACL reconstruction. MIC_{pred} values were smaller than KOOS mean change scores for patients reporting important improvements and for those reporting acceptable symptoms after surgery. For Sport/Rec and QOL, MIC_{pred} values were similar to mean change scores for patients being undecided about treatment outcome, but larger than for patients experiencing no important changes (Figure 4 and Appendix, Section 5).

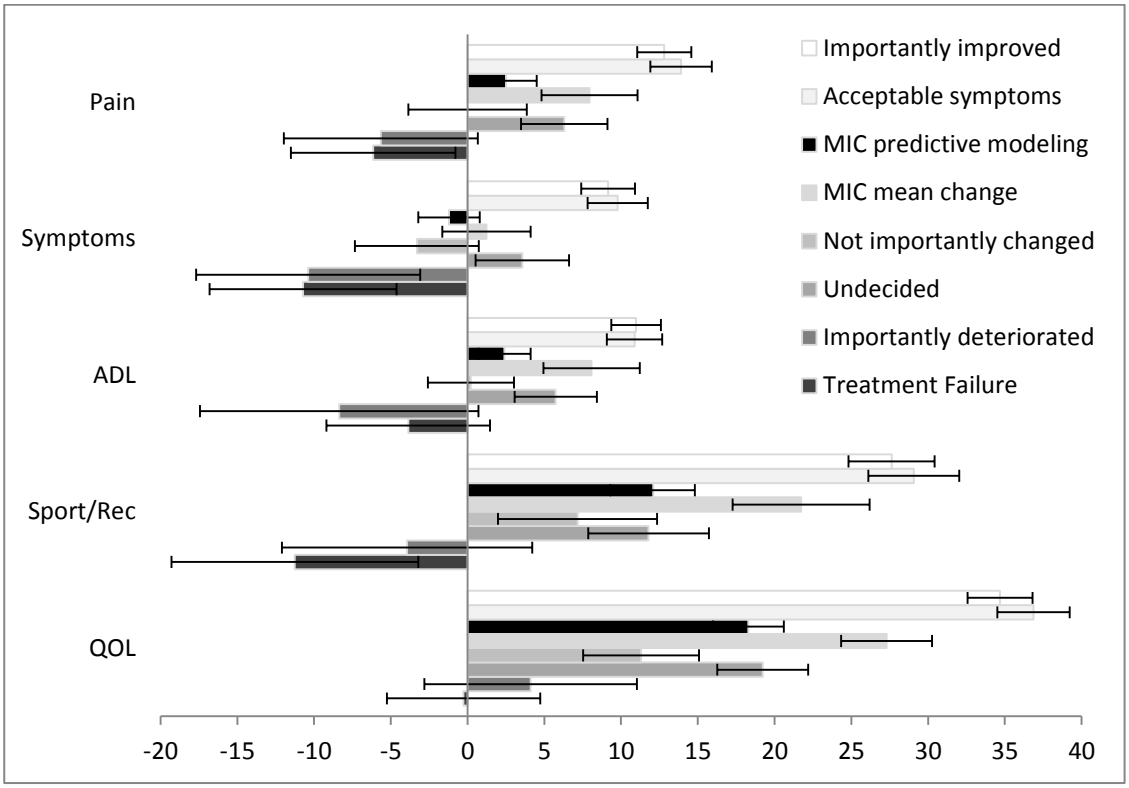

Figure 4. MIC values in comparison to other determinants of outcome following ACL reconstruction for the five KOOS subscales, respectively. Predictive modeling MIC values and mean change MIC values compared to mean change scores for patients who report to be 'importantly improved', 'importantly deteriorated' or 'unchanged' and mean change scores for patients reporting 'acceptable symptoms' or being 'undecided' regarding the outcome after ACL reconstruction. Error bars represent 95% CI.

DISCUSSION

Summary of findings

This study proposes estimates for the interpretation of meaningful improvement in KOOS scores after an ACL reconstruction. We found that these estimates are not dependent on the time to follow-up and, therefore, can be applied 6-24 months following ACL reconstruction. The MIC values for the subscales Pain, Symptoms and ADL were lower (-1.2 to 2.5) due to, on average, only mild problems pre-operatively (mean subscale scores 71-82) suggesting limited room for post-operative improvement compared to Sport/Rec and QOL with pre-operative mean scores of 40 and 34 (Table 1). However, the proportions of patients who consider themselves importantly improved in the subscales of Pain, other Symptoms and ADL are comparable to the subscales Sport/Rec and QOL, implying that the correspondingly smaller MIC values may still be a true reflection of what the average patient considers to be a minimal important improvement. While it is important to acknowledge that some ACL-injured patients actually report problems with pain, other symptoms and ADL function, the overall limited room for improvement in the Pain, Symptoms and ADL subscales with treatment confirm the previous recommendation to use the KOOS subscales Sport/Rec and QOL as primary outcomes following ACL reconstruction.[1]

Comparison of three different MIC methods

Different MIC analysis methods resulted in quite different MIC values. In line with previous simulation studies, predictive modeling was more precise than the ROC and mean change methods, giving smaller 95% CI around the MIC values.[21,22] Other benefits of predictive modeling are the reduced sensitivity to low correlation with the anchor question and the capacity to adjust for when the proportions of improved patients differ from 50%.[21]

Due to the low to moderate correlations, however comparable to those observed in other studies,[12,23] we found that using the Youden principle for defining the 'best' ROC cut-off point gave somewhat arbitrary MIC_{ROC} values because a large range of cut-off points were associated with approximately the same

276 relatively large degrees of misclassification. MIC_{ROC} values are not recommended for further use but are
277 presented in the Appendix for those with a methodological interest.

278 We found larger $MIC_{MeanChange}$ than adjusted MIC_{pred} values for all subscales, which is to be expected
279 because the adjusted MIC_{pred} reflects the mean of the individual MICs in a sample whereas the $MIC_{MeanChange}$
280 represents the mean change score of a subgroup whose perceived change has exceeded their individual
281 MICs.[22] $MIC_{MeanChange}$ values are considered less credible because the calculations are based on subgroups of
282 patients, while the MIC_{pred} values are calculated using the whole patient sample.[22,23] Additionally, the mean
283 change method is dependent on a reasonable correlation between the change in score and the anchor
284 item. Furthermore, since $MIC_{MeanChange}$ is the mean change in the subgroup of patients having minimal
285 important improvement, assuming normally distributed data, only half of the patients who reported a
286 minimal important improvement would actually be characterized as responders. Hence, $MIC_{MeanChange}$ is
287 considered less suitable to define responders.[11]

288 Further investigation is needed to confirm whether the predictive modeling approach is capable of
289 producing reliable MIC values in circumstances where the ROC and mean change methods are
290 inappropriate due to the scores' distributional characteristics, low correlation between change in score and
291 anchor question and, especially following surgical treatment, the proportion of improved patients being
292 greater than 50%.

293 **Comparison to previous studies**

294 Three previous studies have proposed MIC estimates for the KOOS in older populations with knee
295 osteoarthritis undergoing rehabilitation, all of which have used the ROC and/or mean change
296 methods.[12,13,20] These studies differ from our study with regard to patient group, intervention and MIC
297 methodology used. Since MIC values are context-specific, it is less meaningful to compare the current MIC
298 values determined in young adults to surgically reconstructed ACL-deficient knees to those obtained in
299 studies of older and less physically active people having had their knee joints replaced.

Understanding the MIC concept relative to other outcome cut-off points

To facilitate the understanding of how the MIC concept relates to other relevant cut-off points for interpreting outcomes from an intervention, we displayed MIC_{pred} and $MIC_{MeanChange}$ values together with mean KOOS change scores for those reporting different levels of change post-operatively, and for those who reported acceptable postoperative symptoms, who considered the treatment had failed, or who were undecided about their treatment outcome (Figure 4). The finding that average changes in KOOS scores were approximately similar for patients reporting acceptable symptoms post-operatively and patients reporting important improvements on the anchor items after surgery implies that reporting acceptable symptoms corresponds to perceiving an important improvement after ACL reconstruction. We consider it important to acknowledge that these values do not represent an optimal post-operative outcome or readiness to return to sport.

Another explanation for the similarity in mean change scores in those having acceptable post-operative symptoms and those being importantly improved is that patients value their post-operative state more than the actual change when responding to the anchor questions. In line with previous research, the anchor questions in this study correlated more with the post-operative KOOS scores than the KOOS change scores (Appendix, section 6), which could be caused by response shift and recall bias.[9,19] When responding to the anchor questions, patients are required to retrospectively consider what degree their state has changed and make a judgement of importance. The response shift theory implies that patients may change their criteria for how they judge their own state, leading to paradoxical responses to the anchor response questions when compared to the degree of score change found in the PROM.[19] One such example could be if a patient presented with an unchanged KOOS Sport/Rec score, but still reported to be importantly improved. Recall bias implies that patients may not remember their initial state, and consequently base their judgement of important change on their post-operative state more than the actual change.[9] To what degree recall bias and response shift affect the MIC estimations is unclear.[9,19]

Limitations

This study is limited by the low response rate with less than half of the randomly selected patients included in the final analyses. The responders were older and more often female, which may limit the generalizability of our results to the younger male ACL-reconstructed population. However, the mean age difference was only 2 years. In support of the pre-operative differences being negligible, confidence intervals around the mean pre-operative KOOS scores overlapped widely (Table 1), suggesting that responders and non-responders did not differ with respect to their knee-specific functional state.

Another limitation is that even though 4-10% of patients reported worse outcomes for the five KOOS subscales following surgery, we were not able to estimate MIC values for deterioration. We have previously reported that when asked post-operatively, approximately 10% considered the ACL reconstruction to have failed and their KOOS scores corresponded to moderate to severe problems on average.[7] A responder analysis from a randomised controlled trial should present both the proportion of importantly improved and the proportion of importantly deteriorated patients to help interpret treatment effects.[5] In theory, two treatments could result in the same rates of importantly improved patients, but one treatment presents a larger proportion of importantly deteriorated patients, which is an important aspect to include in shared decision-making.

Lastly, even though our findings suggest no baseline dependency of MIC_{pred} values, further adjustment using more elaborate methods for proportions of patients who reported being improved may result in differences between severity subgroups in adjusted MIC_{pred} values. Nevertheless, the proposed MIC_{pred} estimates are considered applicable for interpreting group level results for the ACL-reconstructed population.

Implications of findings

The presented MIC values of 12.1 for Sport/Rec and 18.3 for QOL can aid in interpreting average within-group improvements, and in defining responders (i.e. individuals who experience an adequate treatment effect). A responder analysis facilitates interpretation of results from clinical trials by presenting the proportion of responders in each group.[6,11,14] In a clinical context, the individual patient is capable of defining what is important to him/her,[9] although the MIC values may serve as references to what the 'average patient' undergoing an ACL reconstruction would deem important. Due to the smaller room for improvement for the average patient after an ACL reconstruction in the subscales Pain, Symptoms, and ADL, the content validity for these subscales may be questioned and we do not consider the estimates useful for interpreting within-group change over time, nor as responder criteria. Based on this study, we are not able to recommend estimates for future sample size calculations or interpretation of difference in mean change scores between groups of patients. Such minimal important difference (MID) values are much more complex to derive, involving not only perceived changes in pain and functional status, but more importantly, value judgements about the costs and risks involved in the comparison treatments.[18]

CONCLUSION

In conclusion, we found that meaningful score changes vary across KOOS subscales. MIC values calculated with predictive modeling were 12.1 for Sport/Rec and 18.3 for QOL. Predictive modeling yielded more robust MIC estimates than the ROC and the mean change methods. Our findings confirm the previous recommendation that following ACL reconstruction, the KOOS subscales Sport/Rec and QOL are preferred as primary outcomes.

REFERENCES

1. Collins NJ, Prinsen CAC, Christensen R, Bartels EM, Terwee CB, Roos EM. Knee Injury and Osteoarthritis Outcome Score (KOOS): Systematic review and meta-analysis of measurement

properties. *Osteoarthr Cartil*. 2015;24(8):1317-1329. doi:10.1016/j.joca.2016.03.010.

2. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol*. 2003;56(5):395-407. doi:10.1016/S0895-4356(03)00044-1.

3. Granan L-P, Bahr R, Steindal K, Furnes O, Engebretsen L. Development of a National Cruciate Ligament Surgery Registry. *Am J Sports Med*. 2008;36(2):308-315. doi:10.1177/0363546507308939.

4. Granan LP, Baste V, Engebretsen L, Inacio MCS. Associations between inadequate knee function detected by KOOS and prospective graft failure in an anterior cruciate ligament-reconstructed knee. *Knee Surgery, Sport Traumatol Arthrosc*. 2014:1-6. doi:10.1007/s00167-014-2925-5.

5. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ*. 1998;316(7132):690-693. doi:10.1136/bmj.316.7132.690.

6. Guyatt GH, Osoba D, Wu AW, Wyrwich KW, Norman GR. Methods to explain the clinical significance of health status measures. *Mayo Clin Proc*. 2002;77(4):371-83. doi:10.1016/S0025-6196(11)61793-X.

7. Ingelsrud LH, Granan L-P, Terwee CB, Engebretsen L, Roos EM. Proportion of Patients Reporting Acceptable Symptoms or Treatment Failure and Their Associated KOOS Values at 6 to 24 Months After Anterior Cruciate Ligament Reconstruction: A Study From the Norwegian Knee Ligament Registry. *Am J Sports Med*. 2015;43(8):1902-1907. doi:10.1177/0363546515584041.

8. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415. doi:10.1016/0197-2456(89)90005-6.

9. King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert Rev Pharmacoecon Outcomes Res*. 2011;11(2):171-184. doi:10.1586/erp.11.9.

10. Magnussen R a, Granan L-P, Dunn WR, et al. Cross-cultural comparison of patients undergoing ACL reconstruction in the United States and Norway. *Knee Surg Sports Traumatol Arthrosc*. 2010;18(1):98-105. doi:10.1007/s00167-009-0919-5.

11. McLeod LD, Coon CD, Martin SA, Fehnel SE, Hays RD. Interpreting patient-reported outcome results: US FDA guidance and emerging methods. *Pharmacoeconomics Outcomes Res - Expert Rev*. 2012;11(2):163-169. doi:10.1586/erp.11.12.Interpreting.

12. Mills KAG, Naylor JM, Eyles JP, Roos EM, Hunter DJ. Examining the Minimal Important Difference of Patient-reported Outcome Measures for Individuals with Knee Osteoarthritis: A Model Using the Knee Injury and Osteoarthritis Outcome Score. *J Rheumatol*. 2016;43(2):395-404. doi:10.3899/jrheum.150398.

13. Monticone M, Ferrante S, Salvaderi S, Motta L, Cerri C. Responsiveness and minimal important changes for the Knee Injury and Osteoarthritis Outcome Score in subjects undergoing rehabilitation after total knee arthroplasty. *Am J Phys Med Rehabil*. 2013;92(10):864-70. doi:10.1097/PHM.0b013e31829f19d8.

14. Patrick D, Guyatt G, Acquadro C. Patient-reported outcomes. In: Higgins J, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (Updated March 2011)*. The Cochrane Collaboration; 2011. www.handbook.cochrane.org.

15. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol*. 2008;61(2):102-109. doi:10.1016/j.jclinepi.2007.03.012.

16. Roos E. www.koos.nu. www.koos.nu. Accessed February 1, 2016.

17. Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. *Health Qual Life Outcomes*. 2003;1:64. doi:10.1186/1477-7525-1-64.

18. Schulz KF, Grimes D a. Epidemiology 1 - Sample size calculations in randomised trials : mandatory and mystical. *Lancet*. 2005;365(9467):1348-53. doi:10.1016/S0140-6736(05)61034-3.

19. Schwartz CE, Powell VE, Rapkin BD. When global rating of change contradicts observed change:

418 Examining appraisal processes underlying paradoxical responses over time. *Qual Life Res*.
419 2017;26(4):847-857. doi:10.1007/s11136-016-1414-3.

420 20. Singh JA, Luo R, Landon GC, Suarez-Almazor M. Reliability and Clinically Important Improvement
421 Thresholds for Osteoarthritis Pain and Function scales: A Multicenter study. *J Rheumatol*.
422 2014;41(3):509-515. doi:10.3899/jrheum.130609.

423 21. Terluin B, Eekhout I, Terwee CB. The anchor-based minimal important change, based on receiver
424 operating characteristic analysis or predictive modeling, may need to be adjusted for the proportion
425 of improved patients. *J Clin Epidemiol*. 2017. doi:10.1016/j.jclinepi.2016.12.015.

426 22. Terluin B, Eekhout I, Terwee CB, de Vet HCW. Minimal important change (MIC) based on a predictive
427 modeling approach was more precise than MIC based on ROC analysis. *J Clin Epidemiol*. March 2015.
428 doi:10.1016/j.jclinepi.2015.03.015.

429 23. Terwee CBB, Roorda LDD, Dekker J, et al. Mind the MIC: large variation among populations and
430 methods. *J Clin Epidemiol*. 2010;63(5):524-534. doi:10.1016/j.jclinepi.2009.08.010.

431 24. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health
432 status questionnaires: distinction between minimally detectable change and minimally important
433 change. *Health Qual Life Outcomes*. 2006;4(54). doi:10.1186/1477-7525-4-54.

434 25. de Vet HCW, Beckerman H, Terwee CB, Terluin B, Bouter LM. Definition of Clinical Differences. *J*
435 *Rheumatol*. 2004;36(11):434-435.

436 26. de Vet HCW, Foumani M, Scholten MA, et al. Minimally important change values of a measurement
437 instrument depend more on baseline values than on the type of intervention. *J Clin Epidemiol*.
438 2015;68(5):518-24. doi:10.1016/j.jclinepi.2014.07.008.

439 27. De Vet HCW, Ostelo RWJG, Terwee CB, et al. Minimally important change determined by a visual
440 method integrating an anchor-based and a distribution-based approach. *Qual Life Res*.

2007;16(1):131-142. doi:10.1007/s11136-006-9109-9.

FIGURE LEGENDS

Figure 1. Study flow diagram.

Figure 2. Percentages of improved, unchanged or deteriorated patients for each KOOS subscale are given on the y-axis. Numbers given in each column represent the number of patients reporting to be improved, unchanged or deteriorated.

Figure 3: KOOS mean change scores by anchor question response category ranging from better to worse. Horizontal bars represent the median, the box represents the interquartile range and the whiskers represent the highest and lowest scores.

Figure 4. MIC values in comparison to other determinants of outcome following ACL reconstruction for the five KOOS subscales, respectively. Predictive modelling MIC values and mean change MIC values compared with mean change scores for patients who report to be 'importantly improved', 'importantly deteriorated' or 'unchanged' and mean change scores for patients reporting 'acceptable symptoms' or being 'undecided' regarding the outcome after ACL reconstruction. Error bars represent 95% CI.