Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Responsiveness and Minimal Clinically Important Differences of Four Patient- Reported Outcome Measures Following Treatment of Knee Articular Cartilage Defects

M. Naomi Dugard
Robert Jones and Agnes Hunt Orthopaedic Hospital
Nicola J. Kuiper
Robert Jones and Agnes Hunt Orthopaedic Hospital
Jan Herman Kuiper
j.h.kuiper@keele.ac.uk

Robert Jones and Agnes Hunt Orthopaedic Hospital

Research Article

Keywords: Minimal Clinically Important Difference (MCID), Cell therapy, Articular cartilage repair, Lysholm score, International Knee Documentation Committee, Cincinnati Knee Rating System

Posted Date: November 19th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11527839/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Responsiveness and Minimal Clinically Important Differences of Four Patient-Reported Outcome Measures Following Treatment of Knee Articular Cartilage Defects

M. Naomi Dugard[a,b], Nicola J Kuiper[a,b], Jan Herman Kuiper*[a,b]

[a] Oswestry/Keele Orthopaedic Research group, The Robert Jones & Agnes Hunt Orthopaedic Hospital NHS Foundation Trust, Oswestry, SY10 7AG, UK

[b] Centre for Science Technology and Medicine, School of Life Sciences, Keele University, Keele, UK

***Denotes the corresponding author**

Jan Herman Kuiper, Robert Jones & Agnes Hunt Orthopaedic Hospital, Oswestry, Shropshire SY10 7AG, UK. Email: j.h.kuiper@keele.ac.uk

Background: Patient-reported outcome measures (PROMs) are essential for evaluating treatment efficacy. This study investigated the responsiveness of the Modified Lysholm Knee Scale and compared its performance with three established PROMs following treatment of articular cartilage defects of the knee.

Methods: A total of 233 patients (144 men, 89 women; mean age 34.6 ± 9.0 years) with symptomatic chondral defects of the medial femoral condyle (67%), lateral femoral condyle (17.6%), trochlea (7.7%) or patella (7.7%) were recruited. Participants completed the Modified Lysholm Knee Scale, original Lysholm Knee Scale, International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form and three scales of the Cincinnati Knee Rating System (Sports Activity (SA), Activities of Daily Living (ADL), Sports Function (SF)) before surgery and at a median follow-up of 15.1 months (range 10.9-27.3 months). Responsiveness and the minimal clinically important difference (MCID) were determined for each measure.

Results: The Modified Lysholm Scale demonstrated acceptable responsiveness (effect size = 0.84; standardised response mean = 0.70). Its change scores correlated strongly with those of the IKDC (r = 0.76), and stronger with those of the ADL scale than the SA scale (r=0.65 vs r=0.42). MCIDs derived using distribution- and anchor-based approaches were 4-10/13 points for the Modified Lysholm Scale, 5-12/14 for the original Lysholm Scale, 4-10/11 for the IKDC, and 6-15/15, 6-14/20 and 11-28/40 for three Cincinnati scales.

Conclusions: The modified Lysholm Knee scale demonstrated satisfactory responsiveness for assessing outcomes one year after treatment of articular cartilage defects using either cell-based therapies or alternative surgical interventions.

Key Words: Minimal Clinically Important Difference (MCID); Cell therapy; Articular cartilage repair; Lysholm score; International Knee Documentation Committee; Cincinnati Knee Rating System

Introduction

The primary aims of cartilage repair are to restore function, alleviate symptoms, and improve patients' quality of life. The treatment options for articular cartilage defects in the knee have expanded considerably over recent decades, with surgical procedures such as microfracture [1], osteochondral grafting [2], autologous chondrocyte implantation (ACI) [3,4] and cartilage transplantation [5] now widely used. Despite advances in surgical management, outcomes remain highly variable and are influenced by both patient-related factors, including age, body mass index, activity level, and lesion-related characteristics such as defect size, location, and chronicity [6,7].

As the focus of cartilage repair has shifted towards patient-centred outcomes, patient-reported outcome measures (PROMs) have become integral to treatment evaluation [8,9]. PROMs provide direct insight into patients' perceptions of pain, symptoms, function, and their ability to participate in daily activities. To be clinically useful, PROMs must demonstrate robust psychometric properties, including reliability, validity, and responsiveness to clinically meaningful change over time [10,11]. Responsiveness is particularly important when assessing treatment effectiveness, as more responsive instruments can detect change more efficiently and may reduce the sample size required in clinical studies [12,13]. However, the clinical importance of observed changes in PROM scores will depend on the minimal

clinically important difference (MCID). This is defined as the smallest change in an outcome measure that is meaningful to patients and may influence clinical decision-making [14,15]. MCID values facilitate interpretation of treatment effects, support shared decision-making, provide clinically meaningful benchmarks for evaluating outcomes following cartilage repair and are important in sample size calculations for clinical trials [16,17].

The Lysholm Knee Scale, originally developed to assess function following anterior cruciate ligament injury, has been applied across a range of knee disorders, including articular cartilage defects [18–20]. Following Rasch analysis, Smith et al. [21] demonstrated that the original Lysholm scale did not fully satisfy modern measurement criteria in patients with articular cartilage damage and consequently developed the Modified Lysholm Knee Scale. Removal of the swelling item improved the scale's fit to the Rasch model, resulting in a seven-item, unweighted questionnaire that provides a unidimensional measure of knee function. The Modified Lysholm assesses pain, instability, locking, stair climbing, limping, squatting and use of support, enabling scores to be interpreted on a continuous scale while reducing the influence of items that contribute little to measurement precision in patients with cartilage defects. This Rasch-derived modification improves score interpretability by ensuring that each item contributes meaningfully and equally to the underlying construct of knee function, thereby enhancing measurement precision and minimising bias associated with ordinal scoring systems [22]. The Modified Lysholm has demonstrated promising psychometric properties in patients with articular cartilage damage and has been proposed as an outcome measure for this population, although further validation studies have been recommended prior to its

widespread adoption in cartilage repair research [8,21]. Furthermore, contemporary guidance on the evaluation of PROMs emphasises the importance of rigorous assessment of reliability, validity and responsiveness before implementation in clinical research [11]. Despite its increasing use in cartilage repair studies, clinically meaningful thresholds for interpreting changes in Modified Lysholm scores remain less well defined.

The International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form is a widely used knee-specific patient-reported outcome measure that assesses symptoms, function, and sports activity across a broad range of knee disorders using a single summary score, making it straightforward to administer, interpret, and compare across studies [23,24]. Contemporary consensus recommendations for cartilage repair research recognise the IKDC as one of the most appropriate knee-specific outcome measures because of its demonstrated reliability, validity and responsiveness [8]. In contrast, the Cincinnati Knee Rating System (CKRS) was developed to provide a more comprehensive assessment of knee outcomes by incorporating information on symptoms, Sports Activity (SA), Activities of Daily Living (ADL), Sports Function (SF) and clinical examination findings [25]. Although the CKRS has been widely used in studies of ligament reconstruction and athletic populations, its multidimensional structure and inclusion of clinician-assessed components have limited its adoption relative to more contemporary patient-reported instruments [9]. Consequently, while the CKRS may provide additional insight into sports participation and athletic function, the IKDC is generally preferred in current clinical research because it offers a validated, patient-centred assessment of overall knee health and function with a single, easily

interpretable score and has been recommended as a core outcome measure in studies evaluating cartilage repair interventions [8].

Although the Modified Lysholm Knee Scale has demonstrated acceptable validity and reliability in patients with chondral disorders, evidence regarding its responsiveness following cartilage repair remains limited, and MCID thresholds have yet to be established for this population. Also, its performance relative to other commonly used knee-specific PROMs has not been fully evaluated. Therefore, the purpose of this study was to compare the responsiveness of the Modified Lysholm Knee Scale with that of the IKDC, the CKRS and the original Lysholm Knee Scale, following treatment of articular cartilage defects of the knee.

Methods

Source of study participants

This study was reported in accordance with the COSMIN reporting guideline for studies evaluating measurement properties of PROMs. The COSMIN checklist relevant to construct validity, responsiveness and interpretability is provided in **Appendix A** in **Supplementary Material**. This study utilised anonymised data from the ACTIVE trial (https://www.thelancet.com/protocol-reviews/04PRT-28), a multicentre randomised controlled study approved by the North Staffordshire Research Ethics Committee, UK (04/Q2604/10). The primary aim of ACTIVE was to determine whether ACI provided superior long-term clinical outcomes compared with the best alternative non-ACI surgical treatment for symptomatic isolated chondral defects of the knee that remained symptomatic following previous treatment. Patients were randomised to either ACI or the surgeon's pre-selected best alternative treatment. Alternative surgical treatments included debridement, abrasion, drilling, microfracture and mosaicplasty, selected according to clinical indication before randomisation. Randomisation was administered centrally by the University of Birmingham Clinical Trials Unit, and the five and ten-year trial results have been published [26,27].

A convenience sample of eligible patients with symptomatic chondral defects of the medial femoral condyle, lateral femoral condyle, trochlea or patella was analysed. Eligibility required a symptomatic lesion considered suitable for both ACI and at least one alternative surgical treatment, together with failure of previous treatment at least 12 months before recruitment. Patients with bilateral defects, lesions larger than 12 cm^2, prior total meniscectomy, patellar malalignment, osteoarthritis, or a low

likelihood of compliance with rehabilitation or follow-up were excluded. Data from patients previously included in the Rasch analysis used to develop the Modified Lysholm Knee Scale were excluded to ensure independence between instrument development and validation cohorts.

Patient-Reported Knee Outcome Scales

Table 1 shows the characteristics of the four PROMs used in this study. Patients completed the outcome questionnaires at baseline, defined as within 3 months before randomisation and surgery. A 3-month pre-operative assessment window was considered sufficiently close to surgery to reflect baseline knee status while accommodating variation in the interval between recruitment and surgery and minimising missing data. PROMs were repeated 12 months after surgery. At the 12-month follow-up, patients also completed a global assessment of satisfaction with the operation using a five-point Likert scale. Participants were asked: "Please answer the following question only after you have had your operation." Response options were I am extremely pleased with the operation and would recommend it; I am pleased with the operation; I am no different to before the operation; I am worse than before the operation; and I am much worse than before the operation and would not recommend it. Responses were scored from 4 (extremely pleased) to 0 (much worse), with higher scores indicating greater satisfaction.

Table 1. Characteristics of the patient-reported outcome measures (PROMs) evaluated in this study

Measure	Construct assessed	Items (n)	Scoring approach	Score range	Interpretation
Lysholm Knee Score	Knee symptoms and functional limitations (pain, instability, locking, swelling, stair climbing, limping, squatting, use of support)	8	Weighted item scores summed to produce a total score	0-100	Higher scores indicate better knee function; 100 indicates no symptoms or functional limitations
Modified Lysholm Knee Score	Knee symptoms and functional limitations (pain, instability, locking, stair climbing, limping, squatting, use of support)	7	Swelling item omitted; remaining items summed (maximum 24 points) and rescaled to a 0-100 metric	0-100	Higher scores indicate better knee function; 100 indicates no symptoms or functional limitations
International Knee Documentation Committee (IKDC) Subjective Knee Evaluation Form	Knee symptoms, function, and sports activity	18	Responses transformed according to a standardised scoring algorithm to produce a single summary score	0-100	Higher scores indicate better knee function; 100 indicates no symptoms or functional limitations
Cincinnati Knee Rating System (CKRS)	Symptoms, Activities of Daily Living (ADL), Sports Activity (SA), and Sports-Specific Knee Function (SF)	13 domains, including ADL, SA & SF	Composite score can be derived from patient-reported assessment, symptoms, radiology, and surgical assessment	ADL: 0-120 SA: 0-100 SF: 120-300	Higher scores indicate better knee function and activity level

Statistical analyses

Responsiveness was assessed by testing three predefined hypotheses and quantified using the effect size (ES) and standardised response mean (SRM). First, changes in scores between baseline and 1-year follow-up for all four instruments were expected to follow a gradient according to patient satisfaction. Second, changes in the modified and original Lysholm scores were hypothesised to correlate positively with changes in the IKDC score. Third, changes in the two Lysholm scores and the IKDC score were expected to correlate more strongly with the Cincinnati ADL scale than with the Cincinnati SA scale.

ES was calculated as the mean change from baseline to follow-up divided by the standard deviation of the baseline score. ES values of 0.20-0.49 were considered small, 0.50-0.79 moderate, and ≥0.80 large [28]. The SRM was calculated as the mean change from baseline to follow-up divided by the standard deviation of the change score [29]. SRM values were interpreted using the same thresholds: small (≥0.20), moderate (≥0.50), and large (≥0.80) [30]. Spearman's rho correlation coefficients were used to assess the cross-sectional relationships between the Modified Lysholm, Lysholm, IKDC and CKRS outcome measures, and their change scores.

Floor and ceiling effects for the scores were evaluated by calculating the proportion of patients achieving the minimum and maximum possible scores at both the pre-operative and 12-month post-operative assessments. Consistent with established

criteria, floor and ceiling effects were considered acceptable if fewer than 15% of participants achieved the lowest or highest possible score, respectively [10].

The MCID was estimated using the distribution-based and the anchor-based method [31,32]. The distribution-based approach defines the MCID as one-half of the standard deviation of the change between the pre-operative and follow-up scores [31]. However, sometimes the definition 0.2xSD is also used. The anchor-based approach used the self-reported rating of the operation as an anchor. For this to be valid, the correlation coefficient between change in score and anchor value should be at least 0.3 [32]. To improve the precision of the estimate, the entire cohort was dichotomised [33], where scores of 3 or 4 indicated satisfaction and scores of 0 to 2 indicated no improvement or deterioration. Receiver operating characteristic (ROC) analysis was then performed using the change score from baseline to follow-up. The MCID was defined as the optimal ROC cut-off point, determined using the Youden criterion [32,34]. The optimal cut point was determined using a parametric approach if the change score was normally distributed, otherwise a generalised additive model was used [35]. The area under the ROC curve (AUC) was used to interpret the discriminative ability of the scales in identifying satisfied patients, using 0.60 and 0.75 as thresholds between poor, possibly helpful and clearly useful discrimination [36]. Confidence intervals of the anchor-based MCID and AUC were determined using 2000 bootstrap samples [35].

Statistical significance was defined as a two-tailed $p < 0.05$, or when comparing correlation coefficients by marginally or non-overlapping 95% confidence intervals. Missing data were handled using listwise deletion. Normality of change scores was

assessed using quantile-quantile plots. All analyses were performed using R version 4.6 (R Foundation for Statistical Software, Vienna, Austria) and the package "cutpointr" [35].

Results

A total of 233 patients (144 men and 89 women) with symptomatic cartilage defects were recruited. Lesions were located on the medial femoral condyle in 156 patients (67.0%), lateral femoral condyle in 41 (17.6%), trochlea in 18 (7.7%), and patella in 18 (7.7%). Patients were aged 16-56 years (mean 34.6 years, SD 9.0). Pre-operative scores were compared with scores obtained at a median follow-up of 12.6 months (range 9.9-18.0 months) following surgery. Baseline and post-op scores of all scales were not normally distributed, but the change scores of both Lysholm scales and the IKDC score were. All six patient-reported outcome measures demonstrated improvement following surgery (**Table 2**).

Table 2. Pre-operative and post-operative clinical outcome scores, effect size and standardised response mean (SRM)

Outcome measure	N	Baseline score Mean (SD)	Post-operative score Mean (SD)	Mean change (SD)	Effect size	SRM
Lysholm Scale	186	47.0 (19.0)	62.7 (24.8)	15.8 (23.0)	0.83	0.68
Modified Lysholm Scale	186	51.3 (17.4)	65.9 (22.9)	14.6 (20.9)	0.84	0.70
IKDC	178	42.6 (12.8)	54.3 (21.0)	11.8 (19.0)	0.92	0.62
Cincinnati SA	179	41.8 (28.1)	52.9 (29.1)	11.1 (13.5)	0.40	0.36
Cincinnati ADL	182	69.4 (27.8)	84.0 (29.5)	14.6 (28.2)	0.53	0.52
Cincinnati SF	180	151.1 (37.7)	178.3 (57.3)	27.2 (56.5)	0.72	0.48

Responsiveness was moderate to high across all instruments, with effect sizes ranging from 0.40 to 0.92 and standardised response means ranging from 0. 36 to 0.70 (**Table 2**). The greatest responsiveness was observed for the IKDC (ES = 0.92, SRM = 0.62) and the Modified Lysholm (ES = 0.84, SRM = 0.70), whilst the Cincinnati SA scale demonstrated the lowest responsiveness (ES = 0.40, SRM = 0.36). At one-year follow-up, 185 patients completed the satisfaction assessment. Overall, 104 patients (56.2%) reported being pleased or extremely pleased with the outcome of their operation. Specifically, 57 patients (30.8%) reported being extremely pleased and willing to recommend the procedure, whilst 47 (25.4%) were pleased. Forty-two patients (22.7%) reported no change in their symptoms, 34 (18.4%) felt worse than before surgery, and 5 (2.7%) reported being much worse and would not recommend the procedure.

Floor effects were only observed for the three Cincinnati domains at baseline and one-year follow-up and were above the recommended 15% threshold at both time points for the SF component (**Table 3**). Ceiling effects were minimal for all scales other than the Cincinnati activities of daily living scale, where 15.4% of patients achieved the maximum score. Ceiling effects were identical for both Lysholm scales.

Table 3. Floor and ceiling effects of the Lysholm, Modified Lysholm, IKDC and Cincinnati outcome measures at baseline and post-operative follow-up

Effect	Time point	Original Lysholm Scale	Modified Lysholm Scale	IKDC	CKRS		
					SA	ADL	SF
Floor	Baseline (n [%])	0 [0]	0 [0]	0 [0]	6 [3.4]	6 [3.3]	81 [45.0]
	Post-operative (n [%])	0 [0]	0 [0]	0 [0]	4 [2.2]	3 [1.6]	63 [35.0]
Ceiling	Baseline (n [%])	1[0.5]	1[0.5]	0 [0]	2 [1.1]	3 [1.6]	0 [0.0]
	Post-operative (n [%])	12 [6.5]	12 [6.5]	0 [0]	4 [2.2]	28 [15.4]	8 [4.4]

Significant correlations were observed between all outcome measures at baseline and for change scores (**Table 4**). The strongest correlations were observed between the two Lysholm versions (r=0.98 for baseline and change scores) and between the two Lysholm versions and the IKDC score (r=0.71 to r=0.76) whereas the weakest associations were observed between the Lysholm/IKDC scores and the Cincinnati SA score (r = 0.39 to r=0.51).

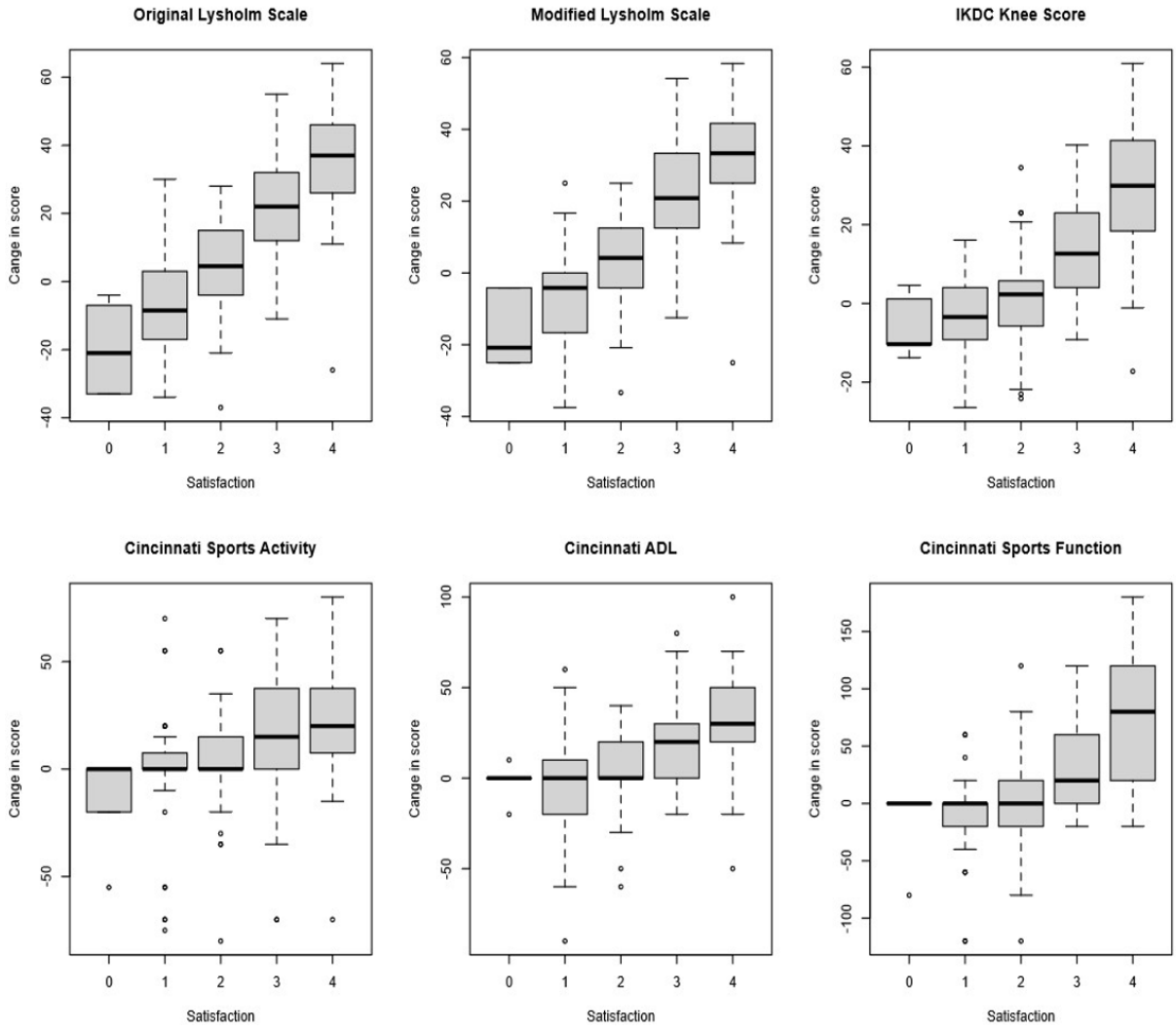
Table 4. Construct validity of the Lysholm, Modified Lysholm, IKDC and Cincinnati outcome measures.

	Lysholm Scale	IKDC	CKRS[a]		
			SA	ADL	SF
Lysholm Scale – correlation with other outcome measure, r [95%CI]	0.98 [0.97-0.98][a] 0.98 [0.97-0.99][b]	0.71 [0.63-0.77][a] 0.74 [0.66-0.80][b]	0.49 0.38-0.60][a] 0.39 [0.26-0.51][b]	0.63 [0.54-0.71][a] 0.45 [0.33-0.56][b]	0.45 [0.33-0.56][a] 0.59 [0.49-0.68][b]
Modified Lysholm Score - correlation with other outcome measure, r [95%CI]	0.98 [0.97-0.98][a] 0.98 [0.97-0.99][b]	0.72 [0.65-0.79][a] 0.76 [0.69-0.82][b]	0.51 0.40-0.61][a] 0.42 [0.29-0.53][b]	0.69 [0.61-0.76][a] 0.65 [0.55-0.73][b]	0.47 [0.35-0.58][a] 0.61 [0.51-0.69][b]
IKDC Knee Score - correlation with other outcome measure, r [95%CI]	-	-	0.46 0.33-0.57][a] 0.43 [0.30-0.54][b]	0.67 [0.58-0.74][a] 0.57 [0.46-0.66][b]	0.61 [0.51-0.69][a] 0.73 [0.65-0.79][b]

Values are Pearson correlation coefficients (r); [a] Correlation between pre-operative scores; [b] Correlation between change scores (post-operative minus pre-operative).

The Spearman correlation coefficient between change in score and anchor value was more 0.3 for all six scales (minimal value for Cincinnati SA, rho=0.44, 95%CI 0.31-0.55). Regarding the three responsiveness hypotheses, the change in both Lysholm scores and IKDC Knee score showed a clear gradient of scores versus satisfaction level, whereas such a gradient was less pronounced for the three Cincinnati score domains (**Figure 1**). Second, the changes in the original and modified Lysholm scores correlated strongly with the change in IKDC score (**Table 4**). Finally, the changes in the original and modified Lysholm scores and IKDC scores correlated stronger with the Cincinnati ADL score than the Cincinnati Sports Activity score (**Table 4**).

Figure 1. Box-plot of change in score values from baseline to 1 year versus satisfaction level for the Original Lysholm Scale, Modified Lysholm Scale, IKDC and three CKRS domains



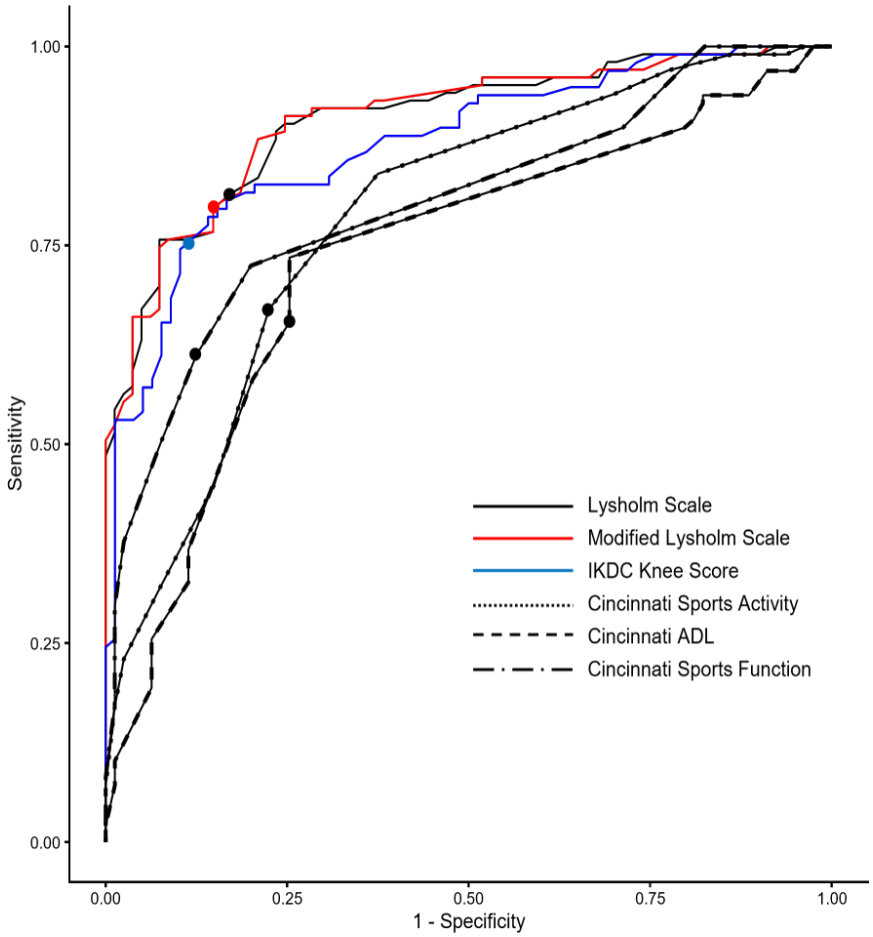
Using a distribution-based approach, the MCID was estimated at 10.4 points for the Modified Lysholm scale, 11.5 points for the original Lysholm scale, 9.5 points for the IKDC, and 15.3, 14.1 and 28.3 points for the Cincinnati SA, ADL and SF domains based on the half-SD definition (**Table 5**). Corresponding estimates using the 0.2XSD definition were 4.6, 4.2, 3.8, 6.1, 5.6 and 11.3 points, respectively. ROC curves were generated using the change in score between baseline and follow-up

assessments to determine the MCID. Greater areas under the curve indicate improved discrimination between patients reporting clinically meaningful improvement and those reporting no meaningful improvement. All scales but the Cincinnati SA scale demonstrated excellent discrimination, with areas under the ROC curve above 0.75 (**Table 5 and Figure 2**). Anchor-based MCID estimates were 12.0 and 13.1 for the modified and original Lysholm scale, 11.0 for the IKDC and 15, 20 and 40 points for the Cincinnati SA, ADL and SF domains (**Table 5**).

Table 5. Distribution and Anchor based MCID and Area Under the ROC Curve (AUC) for the Lysholm, Modified Lysholm, IKDC and three CKRS domains.

Outcome	Domain	Distribution-based MCID [95%CI]	Anchor-based MCID [95%CI]	AUC [95% CI]
Original Lysholm Scale		11.5 [10.4-12.8]	14.0 [11.5-16.4]	0.91 [0.86-0.95]
Modified Lysholm Scale		10.4 [9.5-11.6]	13.1 [10.8-15.3]	0.91 [0.86-0.95]
IKDC		9.5 [8.6-10.6]	11.0 [8.9-13.1]	0.88 [0.83-0.93]
CKRS	**SA**	15.3 [13.8-17.0]	15 [10-20]	0.73 [0.66-0.80]
	ADL	14.1 [12.8-15.7]	20 [10-20]	0.79 [0.72-0.85]
	SF	28.3 [25.6-31.5]	40 [40-60]	0.80 [0.74-0.86]

Figure 2. ROC curves for the Original Lysholm Scale, Modified Lysholm Scale, IKDC and three CKRS domains



Discussion

The present study demonstrates that both the original and Modified Lysholm Knee Scales are responsive outcome measures for evaluating patients undergoing surgical treatment for symptomatic cartilage defects of the knee. In line with previous validation studies, both instruments showed strong responsiveness and correlated highly with established outcome measures, namely the IKDC and Cincinnati ADL domain. The Modified Lysholm, developed following Rasch analysis to improve measurement properties, performed similarly to the original version while maintaining the advantage of a unidimensional measurement tool.

To our knowledge, this is the first study to compare responsiveness and MCID estimates for the original Lysholm, Modified Lysholm, IKDC and three domains of the CKRS in a cohort of patients undergoing a range of cartilage repair procedures. Both Lysholm scales demonstrated large effect sizes and moderate to high responsiveness. These findings are consistent with previous studies which demonstrated favourable psychometric performance of the Lysholm and Modified Lysholm in patients with cartilage defects and strong associations with established knee-specific outcome measures [20,21]. The effect sizes in the present study are at the lower end of those reported previously following microfracture or ACI [8,20]. However, differences in follow-up duration, patient characteristics including activity levels and treatment types may explain these findings [37,38]. The higher effect sizes were observed in studies with 2 years or even longer follow-up [8,20]. In the present study, responsiveness was assessed at a median of 12.6 months following surgery, a time point that may precede maximal recovery in some patients. Evidence from the ACTIVE trial demonstrated continued improvement following cartilage repair

procedures beyond the first post-operative year, with sustained clinical benefits observed at both intermediate (5 years) and long-term (10 years) follow-up [26,27]. Consequently, the responsiveness of the outcome measures evaluated in the present study may differ if assessed at later time points when recovery and adaptation to treatment are more complete.

Among the four PROMs evaluated, the IKDC demonstrated the greatest responsiveness in terms of effect size, whereas the Modified Lysholm scale had the highest standardised response mean. The first finding is consistent with previous reports in mixed knee pathology populations and cartilage repair cohorts, which have identified the IKDC as one of the most responsive knee-specific outcome measures following cartilage restoration procedures [39,40]. The IKDC has also been recommended by the International Cartilage Regeneration & Joint Preservation Society (ICRS) for use in studies of articular cartilage repair because of its strong reliability, validity and responsiveness [8]. The second finding addresses the concern about the responsiveness of the modified Lysholm scale [8]. The three CKRS domains demonstrated the lowest responsiveness. The Cincinnati score was originally developed for use in athletic populations, and differences in their score composition, resulted in strong floor effects for the sports function component among the current population and may explain the lower responsiveness observed in the present study [25]. Nevertheless, all scales except the Cincinnati SA domain exceeded the accepted thresholds for responsiveness and showed a clear improvement following treatment, supporting their utility as outcome measures in patients undergoing cartilage repair procedures.

An important finding of this study is the absence of floor effects in all scores but the Cincinnati SF scale where they were found at baseline and 1-year postoperatively, and the low ceiling effects across all PROMs except the Cincinnati ADL scale 1-year postoperatively. At one-year follow-up, ceiling effects of all scales bar the Cincinnati SF scale remained below the 15% threshold proposed by Terwee et al. [10] for evaluating the measurement properties of health status questionnaires, indicating that none of the other instruments demonstrated substantial score saturation despite the clinical improvements observed following treatment. Comparison of the four measures that had ceiling or floor effects below 15% revealed only modest differences in ceiling effects, suggesting that each retained the ability to discriminate effectively between patients reporting favourable outcomes. These findings are particularly relevant in cartilage repair populations, where post-operative improvements may be expected and ceiling effects have the potential to limit the detection of further gains. The low floor and ceiling effects observed in the present study suggest that all the two Lysholm scales, the IKDC score and the Cincinnati ADL domain remain appropriate outcome measures for cartilage repair populations and that score saturation is unlikely to have considerably influenced the assessment of post-operative improvement [10,41].

As expected, the strongest correlations were observed between the original and Modified Lysholm scales, reflecting the fact that both scores are derived from the same patient responses. Importantly, both versions also demonstrated strong correlations with the IKDC and CKRS, supporting their construct validity. The slightly weaker correlations observed with the CKRS score may reflect differences in the

conceptual focus of the instrument, which was originally designed for more athletically active populations.

The MCID represents the smallest change in score perceived as important by patients and provides a valuable clinical context for interpreting PROM results [14,16]. As no universally accepted gold standard exists for calculating the MCID, we applied both anchor-based and distribution-based approaches, consistent with contemporary recommendations for determining clinically meaningful change in PROMs [13,15]. An anchor-based MCID value for the IKDC after ACI (12.6 [42]) and a distribution-based MCID for the Lysholm scale after ACI (4.2 and 10.5 using 0.2SD and 0.5SD respectively [42]) and an anchor-based MCID for the same after mixed cartilage repair procedures (13.0 [43]) have been published. However, no study has reported the two MCID estimates in patients with articular cartilage defects, and none have been reported for the Modified Lysholm scale. Moreover, the reported anchor-based estimates use the raw ROC curve to determine the optimal cut points. Cut points determined in this way are highly sample-specific and overestimate their discriminative ability [35]. Robust methods were used in the present study, and its results should therefore be transferable to future studies. The MCID estimates in the present study suggest that a change of approximately 5-12 points (distribution-based MCID) or 14 points (anchor-based MCID) on the original Lysholm scale or 4-10 or 13 points on the Modified Lysholm scale represents a clinically meaningful improvement from the patient's perspective. Corresponding thresholds for the IKDC are 4-10 and 11 points. These thresholds are close to those reported by others [42,43] and may help clinicians and researchers to interpret treatment outcomes, evaluate the effectiveness of the treatment, and plan future clinical studies [13,16].

Several limitations should be acknowledged. Patients were recruited as part of the ACTIVE trial, providing a convenience sample of consecutive individuals treated for symptomatic cartilage defects. Although follow-up was planned at one year, assessments were obtained over a wider time window to fit with patient availability. However, most patients were evaluated between 10 and 18 months following surgery, and this variation is unlikely to have substantially affected the psychometric analyses because the majority of patients were assessed within a relatively narrow postoperative window. Nevertheless, longer-term follow-up from the ACTIVE trial suggests that clinically meaningful improvements continue beyond the first postoperative year [26,27], and responsiveness estimates may therefore increase at later stages of recovery.

Conclusion

The original Lysholm Knee Scale, Modified Lysholm Scale, IKDC and Cincinnati ADL and SF scales were all responsive outcome measures for assessing patients with symptomatic cartilage defects of the knee. The Modified Lysholm demonstrated responsiveness, construct validity and acceptable floor and ceiling effects comparable with established outcome measures. Given its previously demonstrated fit to the Rasch model and the clinically meaningful change thresholds identified in this study, the Modified Lysholm represents a robust and clinically useful PROM for evaluating outcomes following cartilage repair surgery.

List of Abbreviations

ACI	Autologous Chondrocyte Implantation
ACTIVE	Autologous Chondrocyte Implantation Versus Alternative Treatments (trial acronym)
ADL	Activities of Daily Living
AUC	Area Under the Receiver Operating Characteristic Curve
CKRS	Cincinnati Knee Rating System
COSMIN	Consensus-based Standards for the Selection of Health Measurement Instruments
ES	Effect Size
ICRS	International Cartilage Regeneration & Joint Preservation Society
IKDC	International Knee Documentation Committee Subjective Knee Evaluation Form
MCID	Minimal Clinically Important Difference
PROMs	Patient-Reported Outcome Measures
ROC	Receiver Operating Characteristic
SA	Sports Activity
SD	Standard Deviation
SF	Sports Function
SRM	Standardised Response Mean

Declarations

Ethics approval and consent to participate

This study utilised anonymised data from the ACTIVE trial, a multicentre randomised controlled study approved by the North Staffordshire Research Ethics Committee, UK (04/Q2604/10). The ACTIVE trial was conducted in accordance with relevant ethical guidelines and regulations, including the principles of the Declaration of Helsinki. Informed consent to participate was obtained from all participants in the original ACTIVE trial. The ACTIVE ethics application form stipulates that anonymised data may be released for secondary research purposes.

Consent for publication

Not applicable. This manuscript does not contain any individual person's identifiable data in any form.

Availability of data and materials

Requests for data should be directed to the corresponding author. Participant-level data will be made available within 6 months of publication. Requests will be assessed for scientific rigour before being granted. Data will be anonymised and securely transferred. A data-sharing agreement might be required.

Competing interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received financial support for the research, authorship, and/or publication of this article from Arthritis UK [Grant Nos. 18480 and 21156] and the Institute of Orthopaedics Ltd. The ACTIVE trial was supported by the UKRI Medical Research Council [ISRCTN48911177]. Nicola Kuiper is supported by a Daphne

Jackson Fellowship co-sponsored by Orthopaedic Research UK [Grant No. 567] and the UKRI Medical Research Council.

Authors' contributions

N.D. and J.K. conducted the data analyses and drafted the first version of the article. J.K. designed the study. All authors interpreted the results, contributed important intellectual content, and constructed the final version of the article. All authors have read and approved the final article.

Acknowledgments

This study was designed largely through the efforts of Professor James Richardson (deceased, 2018). In recognition of his invaluable contributions the authors dedicate this manuscript to his memory and to his family.

The authors thank Mr Samir Mehta from the Birmingham Clinical Trials Unit, UK for collating the clinical scores used in this study.

Nicola Kuiper is a recipient of a Daphne Jackson Trust (DJT) fellowship, and she gratefully acknowledges the financial support of the DJT along with her co-sponsors Orthopaedic Research UK [567] and the UKRI Medical Research Council.

References

[1] Steadman JR, Briggs KK, Rodrigo JJ, Kocher MS, Gill TJ, Rodkey WG. Outcomes of microfracture for traumatic chondral defects of the knee: average 11-year follow-up. Arthroscopy 2003;19:477–84. https://doi.org/10.1053/jars.2003.50112.

[2] Hangody L, Füles P. Autologous osteochondral mosaicplasty for the treatment of full-thickness defects of weight-bearing joints: ten years of experimental and clinical experience. J Bone Joint Surg Am 2003;85-A Suppl 2:25–32. https://doi.org/10.2106/00004623-200300002-00004.

[3] Brittberg M, Lindahl A, Nilsson A, Ohlsson C, Isaksson O, Peterson L. Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. N Engl J Med 1994;331:889–95. https://doi.org/10.1056/NEJM199410063311401.

[4] Davies RL, Kuiper NJ. Regenerative Medicine: A Review of the Evolution of Autologous Chondrocyte Implantation (ACI) Therapy. Bioengineering (Basel) 2019;6:22. https://doi.org/10.3390/bioengineering6010022.

[5] Gross AE, Kim W, Las Heras F, Backstein D, Safir O, Pritzker KPH. Fresh osteochondral allografts for posttraumatic knee defects: long-term followup. Clin Orthop Relat Res 2008;466:1863–70. https://doi.org/10.1007/s11999-008-0282-8.

[6] Jones KJ, Cash BM. Matrix-Induced Autologous Chondrocyte Implantation With Autologous Bone Grafting for Osteochondral Lesions of the Femoral Trochlea. Arthrosc Tech 2019;8:e259–66. https://doi.org/10.1016/j.eats.2018.10.022.

[7] Wang M, Wu Y, Li G, Lin Q, Zhang W, Liu H, et al. Articular cartilage repair biomaterials: strategies and applications. Materials Today Bio 2024;24:100948. https://doi.org/10.1016/j.mtbio.2024.100948.

[8] Roos EM, Engelhart L, Ranstam J, Anderson AF, Irrgang JJ, Marx RG, et al. ICRS Recommendation Document: Patient-Reported Outcome Instruments for Use in Patients with Articular Cartilage Defects. CARTILAGE 2011;2:122–36. https://doi.org/10.1177/1947603510391084.

[9] Makhni EC, Hennekes ME, Baumhauer JF, Muh SJ, Spindler K. AOA Critical Issues: Patient-Reported Outcome Measures: Why Every Orthopaedic Practice Should Be Collecting Them. Journal of Bone and Joint Surgery 2023;105:641–8. https://doi.org/10.2106/JBJS.22.00587.

[10] Terwee CB, Bot SDM, De Boer MR, Van Der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology 2007;60:34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.

[11] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, De Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147–57. https://doi.org/10.1007/s11136-018-1798-3.

[12] Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459–68. https://doi.org/10.1016/s0895-4356(99)00206-1.

[13] Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported

631 outcomes. Journal of Clinical Epidemiology 2008;61:102–9.
632 https://doi.org/10.1016/j.jclinepi.2007.03.012.
633 [14] Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining
634 the minimal clinically important difference. Control Clin Trials 1989;10:407–15.
635 https://doi.org/10.1016/0197-2456(89)90005-6.
636 [15] Sedaghat AR. Understanding the Minimal Clinically Important Difference (MCID)
637 of Patient-Reported Outcome Measures. Otolaryngol--Head Neck Surg
638 2019;161:551–60. https://doi.org/10.1177/0194599819852604.
639 [16] Copay AG, Subach BR, Glassman SD, Polly DW, Schuler TC. Understanding
640 the minimum clinically important difference: a review of concepts and methods.
641 Spine J 2007;7:541–6. https://doi.org/10.1016/j.spinee.2007.01.008.
642 [17] Devji T, Carrasco-Labra A, Qasim A, Phillips M, Johnston BC, Devasenapathy
643 N, et al. Evaluating the credibility of anchor based estimates of minimal
644 important differences for patient reported outcomes: instrument development
645 and reliability study. BMJ 2020;369:m1714. https://doi.org/10.1136/bmj.m1714.
646 [18] Lysholm J, Gillquist J. Evaluation of knee ligament surgery results with special
647 emphasis on use of a scoring scale. Am J Sports Med 1982;10:150–4.
648 https://doi.org/10.1177/036354658201000306.
649 [19] Tegner Y, Lysholm J. Rating systems in the evaluation of knee ligament injuries.
650 Clin Orthop Relat Res 1985:43–9.
651 [20] Kocher MS, Steadman JR, Briggs KK, Sterett WI, Hawkins RJ. Reliability,
652 validity, and responsiveness of the Lysholm knee scale for various chondral
653 disorders of the knee. J Bone Joint Surg Am 2004;86:1139–45.
654 https://doi.org/10.2106/00004623-200406000-00004.

[21] Smith HJ, Richardson JB, Tennant A. Modification and validation of the Lysholm Knee Scale to assess articular cartilage damage. Osteoarthritis Cartilage 2009;17:53–8. https://doi.org/10.1016/j.joca.2008.05.002.

[22] Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? Arthritis Rheum 2007;57:1358–62. https://doi.org/10.1002/art.23108.

[23] Irrgang JJ, Anderson AF, Boland AL, Harner CD, Kurosaka M, Neyret P, et al. Development and validation of the international knee documentation committee subjective knee form. Am J Sports Med 2001;29:600–13. https://doi.org/10.1177/03635465010290051301.

[24] Higgins LD, Taylor MK, Park D, Ghodadra N, Marchant M, Pietrobon R, et al. Reliability and validity of the International Knee Documentation Committee (IKDC) Subjective Knee Form. Joint Bone Spine 2007;74:594–9. https://doi.org/10.1016/j.jbspin.2007.01.036.

[25] Noyes FR, Barber SD, Mooar LA. A rationale for assessing sports activity levels and limitations in knee disorders. Clin Orthop Relat Res 1989:238–49.

[26] Snow M, Middleton L, Mehta S, Roberts A, Gray R, Richardson J, et al. A Randomized Trial of Autologous Chondrocyte Implantation Versus Alternative Forms of Surgical Cartilage Management in Patients With a Failed Primary Treatment for Chondral or Osteochondral Defects in the Knee. Am J Sports Med 2023;51:367–78. https://doi.org/10.1177/03635465221141907.

[27] Snow M, Middleton L, Mehta S, Roberts A, Richardson J, Kuiper JH, et al. Ten year follow-up of a randomized trial of autologous chondrocyte implantation versus alternative forms of surgical cartilage management in the knee.

Osteoarthritis Cartilage 2026;34:1243–51.
https://doi.org/10.1016/j.joca.2026.03.112.
[28] Kazis LE, Anderson JJ, Meenan RF. Effect Sizes for Interpreting Changes in Health Status: Medical Care 1989;27:S178–89.
https://doi.org/10.1097/00005650-198903001-00015.
[29] Liang MH, Fossel AH, Larson MG. Comparisons of Five Health Status Instruments for Orthopedic Evaluation: Medical Care 1990;28:632–42.
https://doi.org/10.1097/00005650-199007000-00008.
[30] Kane, R.L. (ed.) (1997) Understanding health care outcomes research. Gaithersburg, MD: Aspen Publishers. n.d.
[31] Wright A, Hannon J, Hegedus EJ, Kavchak AE. Clinimetrics corner: a closer look at the minimal clinically important difference (MCID). Journal of Manual & Manipulative Therapy 2012;20:160–6.
https://doi.org/10.1179/2042618612Y.0000000001.
[32] Terwee CB, Peipert JD, Chapman R, Lai J-S, Terluin B, Cella D, et al. Minimal important change (MIC): a conceptual clarification and systematic review of MIC estimates of PROMIS measures. Qual Life Res 2021;30:2729–54.
https://doi.org/10.1007/s11136-021-02925-y.
[33] Turner D, Schünemann HJ, Griffith LE, Beaton DE, Griffiths AM, Critch JN, et al. The minimal detectable change cannot reliably replace the minimal important difference. J Clin Epidemiol 2010;63:28–36.
https://doi.org/10.1016/j.jclinepi.2009.01.024.
[34] Perkins NJ, Schisterman EF. The inconsistency of “optimal” cutpoints obtained using two criteria based on the receiver operating characteristic curve. Am J Epidemiol 2006;163:670–5. https://doi.org/10.1093/aje/kwj063.

[35] Thiele C, Hirschfeld G. **cutpointr** : Improved Estimation and Validation of Optimal Cutpoints in *R*. J Stat Soft 2021;98. https://doi.org/10.18637/jss.v098.i11.

[36] Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, et al. Discrimination and Calibration of Clinical Prediction Models: Users' Guides to the Medical Literature. JAMA 2017;318:1377–84. https://doi.org/10.1001/jama.2017.12126.

[37] Knutsen G, Drogset JO, Engebretsen L, Grøntvedt T, Isaksen V, Ludvigsen TC, et al. A Randomized Trial Comparing Autologous Chondrocyte Implantation with Microfracture: Findings at Five Years. The Journal of Bone & Joint Surgery 2007;89:2105–12. https://doi.org/10.2106/JBJS.G.00003.

[38] Gudas R, Gudaitė A, Pocius A, Gudienė A, Čekanauskas E, Monastyreckienė E, et al. Ten-Year Follow-up of a Prospective, Randomized Clinical Study of Mosaic Osteochondral Autologous Transplantation Versus Microfracture for the Treatment of Osteochondral Defects in the Knee Joint of Athletes. Am J Sports Med 2012;40:2499–508. https://doi.org/10.1177/0363546512458763.

[39] Howard JS, Lattermann C, Hoch JM, Mattacola CG, Medina McKeon JM. Comparing Responsiveness of Six Common Patient-Reported Outcomes to Changes Following Autologous Chondrocyte Implantation: A Systematic Review and Meta-Analysis of Prospective Studies. CARTILAGE 2013;4:97–110. https://doi.org/10.1177/1947603512470684.

[40] Abed V, Kapp S, Nichols M, Shephard L, Jacobs C, Conley C, et al. Responsiveness of Patient-Reported Outcome Measures After Large Knee Articular Cartilage Transplantation: A Systematic Review and Meta-analysis.

Am J Sports Med 2024;52:2676–82.
https://doi.org/10.1177/03635465231196156.
[41] McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293–307. https://doi.org/10.1007/BF01593882.
[42] Ogura T, Ackermann J, Barbieri Mestriner A, Merkely G, Gomoll AH. Minimal Clinically Important Differences and Substantial Clinical Benefit in Patient-Reported Outcome Measures after Autologous Chondrocyte Implantation. Cartilage 2020;11:412–22. https://doi.org/10.1177/1947603518799839.
[43] Chahal J, Lansdown DA, Davey A, Davis AM, Cole BJ. The Clinically Important Difference and Patient Acceptable Symptomatic State for Commonly Used Patient-Reported Outcomes After Knee Cartilage Repair. Am J Sports Med 2021;49:193–9. https://doi.org/10.1177/0363546520969883.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1Appendix.docx