Johnson, J. L., Irrgang, J. J., Risberg, M. A., Snyder-Mackler, L. (2020). Comparing the Responsiveness of the Global Rating Scale With Legacy Knee Outcome Scores: A Delaware-Oslo Cohort Study. *American Journal of Sports Medicine, 48*(8), 1953-1960.

Can We Stop Asking So Many Questions? Comparing the Responsiveness of the Global Rating Scale to Legacy Knee Outcome Scores: a Delaware-Oslo Cohort Study

Abstract

Background The selection of patient reported outcome measures (PROMs) is essential for obtaining meaningful information to manage a patient, determine a plan of care, and make clinical decisions, however, the process for selection of PROMs for clinical care is difficult, with the need to balance these multiple factors. Variation makes it difficult to compare data across providers and studies.

Hypothesis/Purpose The objective of this study was to determine the responsiveness of four PROMs via effect size and the presence of a ceiling effect in the five years after ACL reconstruction. We hypothesized that the Global Rating Scale (GRS) would have similar effect size and ceiling effect to the commonly used legacy PROMs.

Study Design Secondary analysis of a cohort study.

Methods Of the 300 participants, 218 chose to have an ACLR, completed post-operative progressive criterion-based rehabilitation early after surgery, and were followed for five years. We collected the GRS, the Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS), the International Knee Documentation Committee-Subjective Knee Form (IKDC-SKF), and the Knee injury and Osteoarthritis Outcome Score (KOOS) at pre- and post-training, and at 6, 12, 24, and 60 months after ACLR.

Results The IKDC-SKF had the largest effect sizes and lowest ceiling effects. The GRS had a similar size and change in both effect size and ceiling effect compared to the longer PROMs. The GRS and IKDC had a correlation of 0.72, and the GRS had a MDC of 2.9/4.8.

Conclusion The GRS responds similarly to the IKDC-SKF, KOS-ADLS, and KOOS and is responsive to patient change. The ease of use and patient-specific nature of the question means that it may be appropriate to use the GRS in clinical care as a consistent measure throughout the course of rehabilitation.

Key Terms patient reported outcome measures, rehabilitation, ACL, knee

What is known about the subject: There are more than 20 PROM for the knee alone, many are reliable and valid. However, the lack of consensus among providers makes the selection of an appropriate outcome measure increasingly difficult.

What this study adds to existing knowledge: The Global Rating Scale was found to have similar effect sizes and presence of ceiling effects as commonly used legacy PROMs.

Introduction

While originally designed for use in research, patient-reported outcome measures (PROMs) are now also used by healthcare providers to assess the effects of clinical care and are designed to measure the patient's perspective on their symptoms, function, and health-related quality of life.[2] The selection of PROMs is essential for obtaining meaningful information to manage a patient, determine a plan of care[5] and make clinical decisions,[35] however, the process for selection of PROMs for clinical care is not easy. The relevance to the patient, psychometric properties, including reliability, validity, and responsiveness, as well as provider and patient burden are all important factors to consider in selection of a PROM.

A 2010 systematic review by Wang et al. identified and evaluated 24 separate PROMs for the knee and recommended different measures depending on diagnosis; while the International Knee Documentation Committee-Subjective Knee Form (IKDC-SKF) was the most generalizable, no measure was applicable across the spectrum of diagnoses or patient group.[36] Even within a diagnosis such as anterior cruciate ligament (ACL) injury and reconstruction, there are multiple recommended outcome measures.[12,19–21] While having multiple PROMs allows clinicians and researchers flexibility in selecting the appropriate measure for their patient, this variation makes it difficult to compare data across providers and studies.[12]

While the number of questions or length of time to complete PROMs that patients will accept varies,[25] longer PROMs place more strain on respondents and administrators and may lead to errors in responding to items and missing data,[11] as well as lower response rates.[25] Additionally, there

may be higher burden if multiple surveys with similar concepts are administered at the same time,[25] as is often the case in research studies and at initial clinical visits.

Previous work on other single-item scales has found they are valid in those with knee injuries.[28,39] The single assessment numeric evaluation (SANE) asks respondents: On a scale from zero to 100, how would you rate your knee today (with 100 being normal)? In comparison, the Global Rating Scale (GRS) asks respondents: Rate your current knee function from 0% to 100%, with 100% equaling preinjury function.

The critical properties of outcome measures are patient-relevancy, user-friendliness, reliability, validity, and responsiveness to clinical change.[10] Kirshner & Guyatt describe three applications of patient-reported measures; discriminative, used to distinguish between individuals on an underlying dimension, predictive, used as a screening or diagnostic instrument, and evaluative, used to quantitate treatment benefit and must be able to detect a clinically meaningful difference over time.[16] This work falls into the last category, assessing the ability to detect real changes in functional impairments and quality of life. Within the context of assessing change, a valid measure must be able to detect a clinically-important change[30] and responsiveness is an instrument's ability to detect real changes in the construct that it is intended to measure.[30] Construct validity may be accessed with correlations other measures of the construct.[14,32] Responsiveness may be measured with effect size,[31] a standardized measure of change in a group,[15] (mean change score in baseline standard deviation units; with pre-training as baseline)[15] and the presence of a ceiling effect (percentage of patients with the maximum score).[34] A measure that has a ceiling effect is unable to identify a further improvement in a high functioning individual.[34]

In a survey of health care providers regarding markers of a successful outcome after ACL injury and reconstruction, Lynch et al. found a consensus (≥80%) that PROMs are an important marker of success, but they did not find a consensus for the preferred use of any individual measure.[18] The highest consensus PROM was for the Global Rating Scale (GRS) with 45% of respondents reporting it as an important measure. The Knee Outcome Survey-Activities of Daily Living Scale (KOS-ADLS), the IKDC-SKF, and the Knee injury and Osteoarthritis Outcome Score (KOOS) were rated as important by 41%, 38%, and 37% of respondents respectively.[18] This lack of consensus among providers makes the selection of an appropriate outcome measure increasingly difficult.

Regardless, having one simple measure that can be used effectively throughout the rehabilitation timeline could improve plan of care decision-making as well as monitor progress across the continuum of care. If GRS has similar effect size and ceiling effect to longer measures, its use in isolation could decrease burden on patients and providers.

The objective of this study was to assess the validity of the GRS compared to the IKDC, determine the minimal detectable change (MDC) for the GRS, and to determine the responsiveness of four PROMs as identified by Lynch et al[18] via effect size and the presence of a ceiling effect in a prospective cohort of ACL reconstructed followed for 5 years. We hypothesized that the GRS would have similar effect size and ceiling effect to the most commonly use legacy PROMs.

Methods

This is a secondary analysis of an ongoing prospective observational study.[8,17] The study was approved by the ethical/human subjects committees at the Regional Ethics Committee for South-Eastern Norway and the University of Delaware and all patients provided written informed consent, or parental consent with written assent if under 18 years old at enrollment. Participants were enrolled at both centers between 2007 and 2012. The Delaware-Oslo ACL Cohort Study is supported by grant R37 HD037985 from the National Institutes of Health

Participants

At enrollment, all participants had a complete unilateral ACL rupture confirmed by 3-mm or greater difference in anterior tibial excursion with instrumented arthrometry[4] (KT1000; MEDmetric Corporation, San Diego, CO) within the previous seven months. Patients were athletes 13 to 55 years of age and regular participants in cutting and pivoting activities for at least 50 hours per year before their injury. Exclusion criteria included a repairable meniscus, symptomatic grade III injury to other knee ligaments, or greater than 1-cm^2 full-thickness articular cartilage lesion.

Study design

Following study enrollment, participants completed ten pre-operative rehabilitation sessions. Of the 300 participants, 218 chose to have an ACLR, completed post-operative progressive criterion-based rehabilitation early after surgery, and were followed for five years. We collected the GRS, KOS-ADLS, IKDC-SKF, and KOOS at pre- and post-training, and at 6, 12, 24, and 60 months after ACLR.

Global Rating Score

The GRS is a single item designed to assess current knee functional performance; patients are asked to rate their current knee function from 0% to 100%, with 100% equaling preinjury function.

Knee Outcome Survey-Activities of Daily Living Scale

The KOS-ADLS consists of 14 questions and was designed to determine symptoms and functional limitations in usual daily activities caused by various knee pathologies.[12] The KOS contains activities such as walking, stair climbing, and kneeling, and symptoms rated on their impact on these activities. It has questions related to recreational or sporting activities. Scores range between 0-100%, a greater symptoms and lower level of function resulting in a lower score.[12]

International Knee Documentation Committee-Subjective Knee Form

The IKDC-SKF has 18 items and was designed to detect improvement or deterioration in symptoms, function, and sports activities in a variety of knee conditions.[13] It is reliable and valid for use in ligament and meniscal injuries, articular cartilage lesions, arthritis, and patellofemoral populations.[13] Scores range between 0-100%, with higher scores representing lower levels of symptoms and higher levels of function and sports activity.[13]

Knee injury and Osteoarthritis Outcome Score

The KOOS has 42 questions arranged in five subscales: Symptoms, Pain, Activities of Daily Living (KOOS-ADL), Sport and Recreation function (Sport/Rec), and knee-related Quality of Life (KOOS-QOL) with scores ranging between 0-100%, with 100% equaling no difficulties.[26]

Statistics

We used Excel (Microsoft, Seattle WA) to calculate effect sizes as: (mean score at time-point minus the mean score at baseline, divided by baseline standard deviation; pre-training as baseline)[15] and the presence of a ceiling effect as a percentage of patients with the maximum score. Effect size and ceiling effects cut-offs were set *a priori:* 0.5 for medium effect size and 0.8 for large effect size,[3] and ≥15% of participants having a maximum score for the presence of a ceiling effect.[23] We correlated with GRS with the IKDC-SKF with a Pearson correlation. We calculated the MDC for the GRS as the values associated with 20%[27] and 33%[41] of the standard deviation of the measure at baseline. Power for the clinical trial was calculated *a priori*.

Results

Of the 300 total participants, 218 (100 women) completed training and chose to have an ACLR (106 hamstring grafts, 42 patellar tendon grafts, and 62 allografts); demographics at enrollment are in Table 1. Results shown are for all available data (Table 2); there was no change when the analysis used only those participants with a complete data set (n=114). We did not begin collecting KOOS responses from study participants until midway through enrollment, thus the analysis of the KOOS subscales included only the 69 participants with baseline KOOS scores.

	Mean (range)
Age (years)	25.0 (13-52)
Height (cm)	174.5 (148-195)
Weight (kg)	75.7 (43.3-139)
Body Mass Index (BMI) (kg/m^2)	24.7 (18.6-40.2)
Time from injury to enrollment (weeks)	8.3 (1-38)

Table 1: Demographics of participants at enrollment

Effect Sizes

Effect sizes for all PROMs peaked at the 24-month time-point, with minimal differences observed between 12m, 24m, and 60m (Figure 1, Table 2). The IKDC-SKF had the largest effect sizes at all time-points (Figure 1). The GRS had a similar effect size and change in effect size compared to the longer PROMs.

Ceiling Effect

The KOOS-QOL subscale did not demonstrate a ceiling effect at any time-point and the IKDC-SKF only showed a ceiling effect at the 24m time-point (Figure 2, Table 2). All other measures had a ceiling effect at at least one time-point (most at 12m, 24m, and 60m) while the KOOS-ADL subscale had a ceiling effect at all time-points, and the KOOS-Pain subscale had ceiling effects at all but pre-training. The GRS had a similar effect size and change in ceiling effect compared to the longer PROMs.

Validity

The Pearson correlation coefficient between the GRS and the IKDC-SKF was 0.72 ($p<0.001$) when data were pooled across all time-points.

Minimally Important Change

The MDC for the GRS was 2.9 when using an 20%[27] and 30%[41] of the standard deviation of the scores, which correspond to a small and small to medium effect.

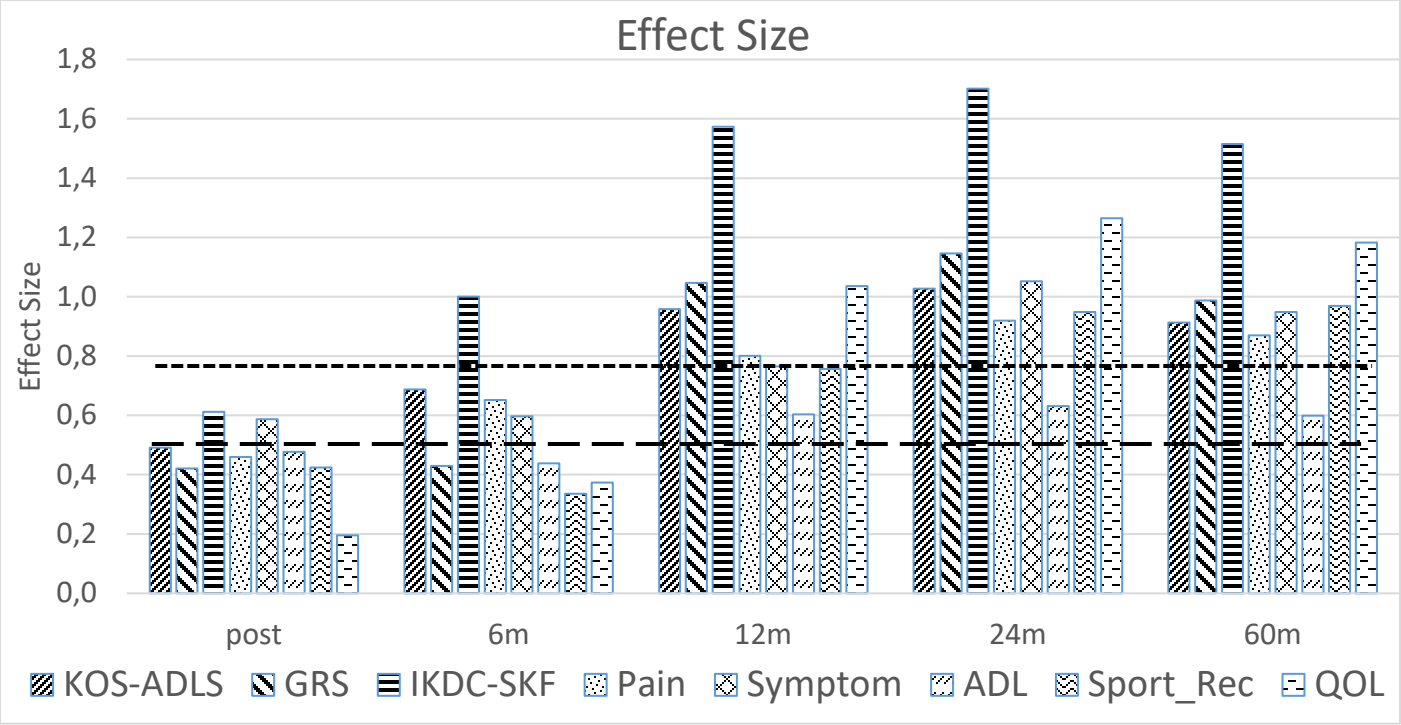

Figure 1 Effect sizes. Medium effect size ≥0.5 (dashed line), large effect size ≥0.8 (dotted line).

Post=post-training time-point

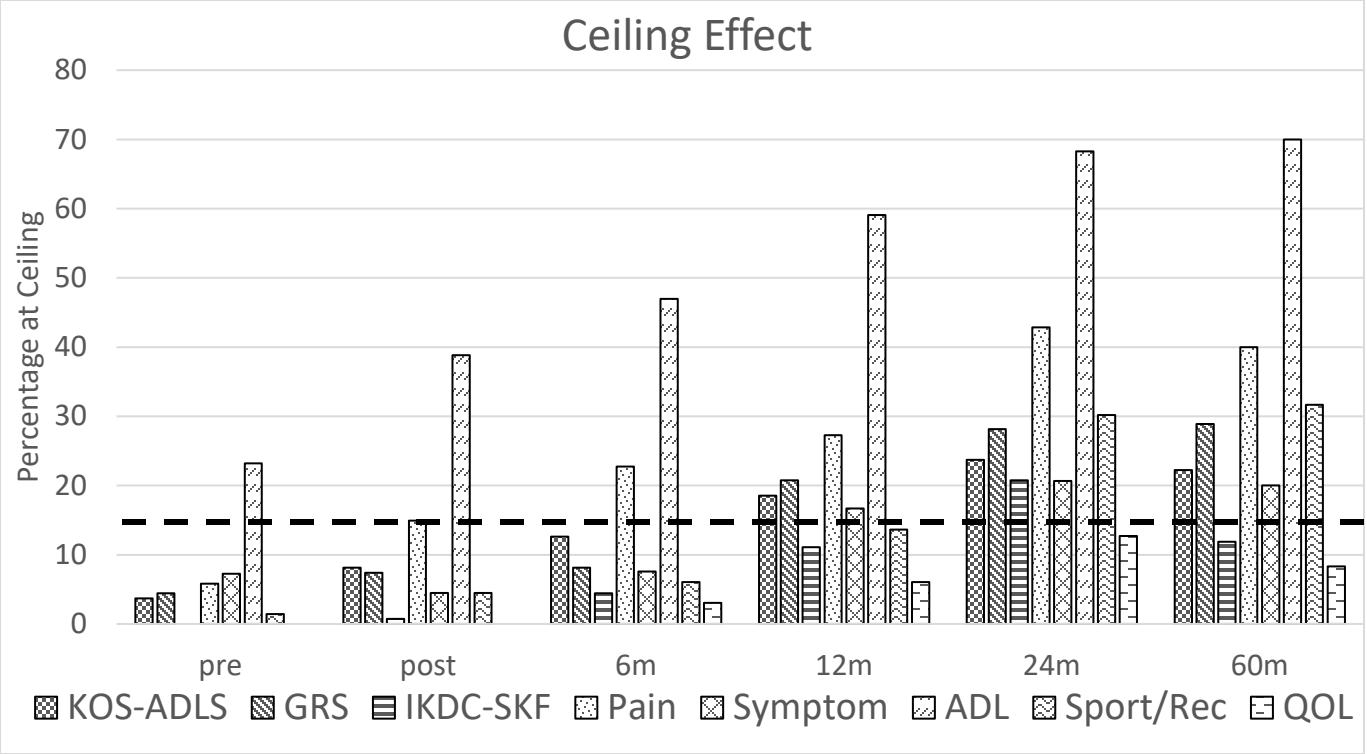

Figure 2 Ceiling effects. The dashed line is the 15% cutoff for ceiling effect.

Pre=pre-training time-point. Post=post-training time-point

Discussion

In the current study, we examined the effect sizes and ceiling effects of four PROMs after an ACL injury and reconstruction. The IKDC-SKF, GRS, KOS-ADLS, and the KOOS-QOL subscale demonstrated large effect sizes at 12m, 24m, and 60m after ACL reconstruction. These values indicated that these PROMs are sensitive to detecting clinical change.[15] While the GRS had a small effect size post-training and at 6m, the overall trajectory of change is similar among all PROMs.

Gangel et al. correlated the SANE with the Lysholm knee score in 130 college-age patients after ACL reconstruction and found a good correlation (0.75, $p<0.001$).[39] Shelbourne et al. correlated the same SANE to the IKDC-SKF after ACL reconstruction and found an ICC of 0.66 and that the

precision of agreement (calculated as the mean difference ±1 standard deviation), was met for 81% of respondents and limits of agreement (calculated as the mean difference ±2 SD) met for 94%.[28] In this study, the correlation between GRS and IKDC-SKF was 0.72 (0.69-0.75), $p<0.001$, which is a moderately strong relationship that provides some evidence of construct validity for the GRS in this population.

As may be expected, the large effect sizes at the later time-points correspond to the time-points with ceiling effects. All measures except the KOOS QOL had a ceiling effect at at least one time-point (most at 12m, 24m, and 60m) while the KOOS-ADL subscale had a ceiling effect at all time-points, and the KOOS-Pain subscale had ceiling effects at all but pre-training. The GRS had a similar size and change in ceiling effect compared to the longer PROMs. However, it is worth debating if the GRS can truly have a ceiling effect. The GRS asks a respondent to rate their current knee functional performance compared to pre-injury performance, so a score of 100% indicates the full resolution of any impairments or limitations relevant to that respondent. This may mean that, regardless of the percentage of respondents with a score of 100%, the GRS does not have a ceiling effect, but a performance asymptote, or the greatest true value that can be demonstrated[37]. In a measure with a ceiling effect, a true score can only be observed if it is less than or equal to the ceiling threshold. So a patient may score 100% on the IKDC-SKF or the KOOS but still have functional limitations because not every activity or symptoms is covered on those PROM. In contrast, we would not expect there to be a score above the 100% scored on the GRS, this is an asymptote and not a ceiling, which differentiates it from the other measures.

Clinical decision making requires that health care providers choose the best tool for patient evaluation[25] and previous authors have suggested a combination of PROMs may be necessary to assess functional success.[5,18] In practice, a clinician could administer the GRS frequently, with the IKDC-SKF (or other measure as desired) given at the initial visit, then at 3 month (initiation of running[24]/return to participation)[1] 6 months (return to sport activities)[1], 9 months (return to sport[9]/return to performance)[1], and at 12 and 24 months after ACLR. Based on our estimation of the MDC of 2.8/4.9, if a participant reports GRS scores of 60, 65, 65, and 70, a therapist should continue to progress treatment as indicated by stage of healing and objective criteria. If, however, a patient reports weekly scores of 60, 65, 65, and then 40, it would be appropriate to administer other tests/measures to identify causes for an unexpected change in status. While single-item measures are reliable and valid,[28,39] their simplicity comes at the cost of detail.[29] To gain this detail, clinicians could use outcome measures related to ADLs and functional limitations (KOS-ADLS[12] or IKDC-SKF[13]), fear (Tampa Scale of Kinseophobia[40] or the ACL-Return to Sports after Injury[38]) or activity participation (Tegner[33] or Marx[22]) as appropriate for the stage of recovery and goals of rehabilitation. Additionally, no PROM alone is sufficient to base all clinical decisions upon. Quadriceps strength, performance measures, and biological healing are all important factors to consider after an ACL injury.[1,5,7,9]

Additionally, when a patient is nearing return to sport/performance, current recommendations suggest both objective and subjective criteria for clearance to return.[7,9] At the 12m time-point, the GRS had a large effect size, indicating it is responsive to a clinically meaningful change from baseline to the time of return-to-sport decision-making. Achieving a GRS of 90[9] may indicate that the patient is ready for objective criteria testing.[7,9] As previously discussed, no one PROM is

applicable to all patients at all time-points.[36] Clinicians can use the GRS for repeated, frequent assessment of change and reserve longer, more burdensome PROM for significant milestones during recovery.

Strengths and Limitations: The study population was active athletes after ACL injury and subsequent ACLR. These patients were not enrolled immediately after their ACL injury, but at a "quiet knee" stage of recovery (8.3±5 weeks, range 1-38 weeks after injury). We, therefore, do not have data from the patient's initial encounter with a health professional, nor do we have immediate post-operative data and so are unable to assess effect sizes with an early post-operative baseline. Additionally, previous research comparing this cohort to a comparable subgroup of the MOON cohort found our participants had statistically and clinically significantly higher baseline and two-year IKDC-SKF and KOOS scores[6].

We assessed validity, responsiveness, and determined the MDC of the GRS in this study. Future research should establish test-retest reliability, standard error of the measurement. Also, while none of our measures are ACL specific, we only analyzed patients with ACL rupture and subsequent ACLR. We do not know the validity, responsiveness or MDC in patients with other knee injuries/surgeries.

Also, because we did not begin using the KOOS until midway through enrollment, analysis of the KOOS was only done in the 69 participants who had baseline KOOS scores.

We were unable to directly assess response burden in our participants. The difference in the length and thus the time needed to complete each PROMs is however much shorter in the GRS than the other three measures. The simplicity of administering the GRS may decrease burden on respondents and administrators while still providing responsive and valid information regarding patient status.

Conclusions

The IKDC-SKF has the largest effects sizes while the KOOS-QOL had the smallest ceiling effects. The GRS, however, responds similarly to the IKDC-SKF, KOS-ADLS, and KOOS measures and is responsive to patient change, with evidence of construct validity and a small MDC. The ease of use and patient-specific nature of the question means that, for clinical practice, it may be appropriate to use the GRS as a frequent measure throughout the course of rehabilitation, with different measures used at the beginning of treatment and other measures used at the later stages, or specific scales based on patient's deficits or goals.

Time point	Outcome	N	Mean	Std. Deviation	Effect size	Percentage at ceiling
Pre	KOS-ADLS	216	84.5	10.3	n/a	3%
	GRS	216	77.6	14.7	n/a	6%
	IKDC-SKF	210	69.6	12.6	n/a	0%
	KOOS-Pain	69	84.0	10.8	n/a	6%
	KOOS-Symptom	69	75.8	13.5	n/a	7%
	KOOS-ADL	69	93.5	7.4	n/a	23%
	KOOS-Sport/Rec	69	66.5	18.7	n/a	1%
	KOOS-QOL	69	51.0	19.1	n/a	0%
Post	KOS-ADLS	202	89.5	8.6	0.49	8%
	GRS	202	83.3	14.3	0.39	8%
	IKDC-SKF	188	76.8	12.8	0.57	1%
	KOOS-Pain	67	89.0	8.9	0.46	15%
	KOOS-Symptom	67	83.7	12.1	0.59	4%
	KOOS-ADL	67	97.1	3.7	0.48	39%
	KOOS-Sport/Rec	67	74.4	16.1	0.42	4%
	KOOS-QOL	67	54.7	18.8	0.20	0%
6m	KOS-ADLS	198	91.4	9.6	0.67	14%
	GRS	195	84.7	15.9	0.49	9%
	IKDC-SKF	188	82.6	12.5	1.03	4%
	KOOS-Pain	66	91.0	8.5	0.65	23%
	KOOS-Symptom	66	83.8	12.2	0.60	8%
	KOOS-ADL	66	96.8	5.1	0.44	47%
	KOOS-Sport/Rec	66	72.7	20.2	0.34	6%
	KOOS-QOL	66	58.1	20.0	0.37	3%
12m	KOS-ADLS	184	93.4	8.7	0.86	20%
	GRS	184	92.1	10.5	1	21%
	IKDC-SKF	181	89.2	11.8	1.54	11%
	KOOS-Pain	66	92.6	8.7	0.80	27%
	KOOS-Symptom	66	86.1	14.7	0.77	17%
	KOOS-ADL	66	98.0	4.8	0.60	59%
	KOOS-Sport/Rec	66	80.6	20.0	0.76	14%
	KOOS-QOL	66	70.8	19.7	1.04	6%
24m	KOS-ADLS	166	94.0	8.1	0.92	28%
	GRS	166	93.8	8.2	1.11	31%
	IKDC-SKF	166	90.7	11.1	1.66	22%
	KOOS-Pain	63	93.9	9.6	0.92	43%
	KOOS-Symptom	63	90.0	11.7	1.05	21%
	KOOS-ADL	63	98.2	5.1	0.63	68%
	KOOS-Sport/Rec	63	84.2	17.7	0.95	30%

	KOOS-QOL	63	75.2	20.9	1.26	13%
60m	KOS-ADLS	169	93.3	9.2	0.86	24%
	GRS	169	91.9	12.1	0.98	30%
	IKDC-SKF	169	88.8	12.4	1.52	13%
	KOOS-Pain	60	93.4	8.4	0.87	40%
	KOOS-Symptom	60	88.6	13.0	0.95	20%
	KOOS-ADL	60	98.0	4.9	0.60	70%
	KOOS-Sport/Rec	60	84.6	18.6	0.97	32%
	KOOS-QOL	60	73.7	21.3	1.18	8%

Table 2: Means, standard deviations, and percentage at ceiling for each PROM by time point.

1. Ardern CL, Glasgow P, Schneiders A, et al. 2016 Consensus statement on return to sport from the First World Congress in Sports Physical Therapy, Bern. *Br J Sports Med*. 2016;50(14):853-864. doi:10.1136/bjsports-2016-096278.
2. Black N. Patient reported outcome measures could help transform healthcare. *BMJ*. 2013;346:f167. doi:10.1136/bmj.f167.
3. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.*; 1988.
4. Daniel DM, Stone M Lou, Dobson BE, Fithian DC, Rossman DJ, Kaufman KR. Fate of the ACL-injured Patient. *Am J Sports Med*. 1994;22(5):632-644. doi:10.1177/036354659402200511.
5. Eitzen I, Moksnes H, Snyder-Mackler L, Engebretsen L, Risberg MA. Functional tests should be accentuated more in the decision for ACL reconstruction. *Knee Surg Sports Traumatol Arthrosc*. 2010;18(11):1517-1525. doi:10.1007/s00167-010-1113-5.
6. Failla MJ, Logerstedt DS, Grindem H, et al. Does Extended Preoperative Rehabilitation Influence Outcomes 2 Years After ACL Reconstruction? *Am J Sports Med*. 2016;44(10):2608-2614. doi:10.1177/0363546516652594.
7. Filbay SR, Grindem H. Evidence-based recommendations for the management of anterior cruciate ligament (ACL) rupture. *Best Pract Res Clin Rheumatol*. 2019;33(1):33-47. doi:10.1016/j.berh.2019.01.018.
8. Grindem H, Eitzen I, Engebretsen L, Snyder-Mackler L, Risberg MA. Nonsurgical or Surgical Treatment of ACL Injuries: Knee Function, Sports Participation, and Knee Reinjury: The Delaware-Oslo ACL Cohort Study. *J Bone Joint Surg Am*. 2014;96(15):1233-1241. doi:10.2106/JBJS.M.01054.
9. Grindem H, Snyder-Mackler L, Moksnes H, Engebretsen L, Risberg MA. Simple decision

rules can reduce reinjury risk by 84% after ACL reconstruction: the Delaware-Oslo ACL cohort study. *Br J Sports Med*. 2016. doi:10.1136/bjsports-2016-096031.

10. H Liang M, M Jette A. Measuring functional ability in chronic arthritis. a critical review. *Arthritis Rheum*. 1981;24(1):80-86. doi:10.1002/art.1780240113.

11. Hays RDRD, Reeve BBB. Measurement and Modeling of Health-Related Quality of Life. *Int Encycl Public Heal*. January 2008:241-252. https://www.sciencedirect.com/science/article/pii/B9780123739605003361. Accessed November 27, 2018.

12. Irrgang J, Snyder-Mackler L, Wainner RS, et al. *Development of a Patient-Reported Measure of Function of the Knee*. Vol 6058.; 1998. https://insights.ovid.com/pubmed?pmid=9730122. Accessed September 5, 2018.

13. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med*. 2001;29(5):600-613. doi:10.1177/03635465010290051301.

14. Irrgang JJ, Snyder-Mackler L, Wainner RS, Fu FH, Harner CD. Development of a Patient-Reported Measure of Function of the Knee. *J Bone Jt Surg*. 1998;80(8):1132-1145. https://insights.ovid.com/crossref?an=00004623-199808000-00006. Accessed March 29, 2019.

15. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care*. 1989;27(3 Suppl):S178-89. http://www.ncbi.nlm.nih.gov/pubmed/2646488. Accessed September 6, 2018.

16. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis*. 1985;38(1):27-36. doi:10.1016/0021-9681(85)90005-0.

17. Logerstedt D, Grindem H, Lynch A, et al. Single-legged hop tests as predictors of self-reported knee function after anterior cruciate ligament reconstruction: the Delaware-Oslo ACL cohort study. *Am J Sports Med*. 2012;40(10):2348-2356. doi:10.1177/0363546512457551.

18. Lynch AD, Logerstedt DS, Grindem H, et al. Consensus criteria for defining "successful outcome" after ACL injury and reconstruction: A Delaware-Oslo ACL cohort investigation. *Br J Sports Med*. 2015;49(5):335-342. doi:10.1136/bjsports-2013-092299.

19. Lysholm J, Tegner Y. Knee injury rating scales. *Acta Orthop*. 2007;78(4):445-453. doi:10.1080/17453670710014068.

20. Marx RG. Knee rating scales. *Arthrosc J Arthrosc Relat Surg*. 2003;19(10):1103-1108. https://www.sciencedirect.com/science/article/pii/S0749806303009721?via%3Dihub. Accessed February 21, 2019.

21. Marx RG, Jones EC, Allen AA, et al. *Reliability, Validity, and Responsiveness of Four Knee Outcome Scales for Athletic Patients*.; 2001. https://insights.ovid.com/pubmed?pmid=11679594. Accessed September 5, 2018.

22. Marx RG, Stump TJ, Jones EC, Wickiewicz TL, Warren RF. Development and Evaluation of an Activity Rating Scale for Disorders of the Knee. *Am J Sports Med*. 2001;29(2):213-218. doi:10.1177/03635465010290021601.

23. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res*. 1995;4(4):293-307. doi:10.1007/BF01593882.

24. Rambaud AJM, Ardern CL, Thoreux P, Regnaux JP, Edouard P. Criteria for return to running after anterior cruciate ligament reconstruction: A scoping review. *Br J Sports*

Med. 2018;52(22):1437-1444. doi:10.1136/bjsports-2017-098602.

25. Rolstad S, Adler J, Rydén A. Response Burden and Questionnaire Length: Is Shorter Better? A Review and Meta-analysis. *Value Heal*. 2011;14(8):1101-1108. doi:10.1016/J.JVAL.2011.06.003.

26. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)—Development of a Self-Administered Outcome Measure. *J Orthop Sport Phys Ther*. 1998;28(2):88-96. doi:10.2519/jospt.1998.28.2.88.

27. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining Clinically Important Differences in Health Status Measures. *Pharmacoeconomics*. 1999;15(2):141-155. doi:10.2165/00019053-199915020-00003.

28. Shelbourne KD, Barnes AF, Gray T. Correlation of a Single Assessment Numeric Evaluation (SANE) Rating With Modified Cincinnati Knee Rating System and IKDC Subjective Total Scores for Patients After ACL Reconstruction or Knee Arthroscopy. *Am J Sports Med*. 2012;40(11):2487-2491. doi:10.1177/0363546512458576.

29. Sloan JA, Aaronson N, Cappelleri JC, et al. Assessing the clinical significance of single items relative to summated scores. In: *Mayo Clinic Proceedings*. Vol 77. Elsevier Ltd; 2002:479-487. doi:10.4065/77.5.479.

30. Stratford PW, Binkley JM, Riddle DL. Health Status Measures: Strategies and Analytic Methods for Assessing Change Scores. *Phys Ther*. 1996;76(10):1109-1123.

31. Streiner DL, Norman GR, Cairney J. Oxford Medicine Online Health Measurement Scales : A practical guide to their development and use (5 ed .) Chapter : Selecting the items. 2019:1-27. doi:10.1093/med/9780199685219.001.0001.

32. Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific

and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol*. 1995;48(11):1369-1378. doi:10.1016/0895-4356(95)00054-2.

33. Tegner Y, Lysholm J, Odensten M, Gillquist J. Evaluation of cruciate ligament injuries A review. *Acta Orthop Scand*. 1988;59(3):336-341. doi:10.3109/17453678809149379.

34. Uttl B. Measurement of individual differences: lessons from memory assessment in research and clinical practice. *Psychol Sci*. 2005;16(6):460-467. http://www.ncbi.nlm.nih.gov/pubmed/15943672. Accessed January 22, 2019.

35. Valier AR, Lam KC. Beyond the Basics of Clinical Outcomes Assessment: Selecting Appropriate Patient-Rated Outcomes Instruments for Patient Care. *Athl Train Educ J*. 2015;10(1):91-100. doi:10.4085/100191.

36. Wang D, Jones MH, Khair MM, Miniaci A. Patient-reported outcome measures for the knee. *J Knee Surg*. 2010;23(3):137-151. http://www.ncbi.nlm.nih.gov/pubmed/21329255. Accessed August 17, 2018.

37. Wang L, Zhang Z, McArdle JJ, Salthouse TA. Investigating Ceiling Effects in Longitudinal Data Analysis. *Multivariate Behav Res*. 2009;43(3):476-496. doi:10.1080/00273170802285941.

38. Webster KE, Feller JA, Lambros C. Development and preliminary validation of a scale to measure the psychological impact of returning to sport following anterior cruciate ligament reconstruction surgery. *Phys Ther Sport*. 2008;9(1):9-15. doi:10.1016/J.PTSP.2007.09.003.

39. Williams GN, Taylor DC, Gangel TJ, et al. Comparison of the Single Assessment Numeric Evaluation Method and the Lysholm Score. *Clin Orthop Relat Res*. 2000;373:184-192. doi:10.1097/00003086-200004000-00022.

40. Woby SR, Roach NK, Urmston M, Watson PJ. Psychometric properties of the TSK-11: A shortened version of the Tampa Scale for Kinesiophobia. *Pain*. 2005;117(1-2):137-144. doi:10.1016/j.pain.2005.05.029.

41. Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: The FACIT Experience. *Eval Heal Prof*. 2005;28(2):172-191. doi:10.1177/0163278705275340.