

Results: Figures and Tables**Abbreviations:**

PROM: Patient-reported Outcome Measure
CAT: Computer adaptive test
PROMIS: Patient Reported Outcomes Measurement Information System
FOTO Knee FS: Focus On Therapeutic Outcomes Knee Functional Status
OA: Osteoarthritis
PF: Physical Function
PI: Pain Interference
PB: Pain behavior
D: Depression
M: Mobility
HAQ-II: Health Assessment Questionnaire Disability Index
MLQOL: Multi-ligament Quality of Life;
WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index
KOOS: Knee injury and Osteoarthritis Outcome Score
IKDC: International Knee Documentation Committee
MARS: Marx Activity Rating Scale
EQ-5D-5L: European Quality of life-5 Dimension-5 level
SF-36: Short Form 36
GROC: Global Rating Of Change
NPS: Numeric Pain Scale
FTSST: Five Time Sit to Stand Test
TUG: Timed Up and Go
IPAQ: International Physical Activity Questionnaire
SANE: Single Assessment Numeric Evaluation
MLKI: Multi-Ligament Knee Injury
BEMD: Bilateral Extensor Mechanism Disruption
BTP: Bicondylar Tibial Plateau
CR: cruciate retaining
ICC: intraclass correlation coefficient
TKA: Total Knee Arthroplasty



PRISMA Flow Diagram

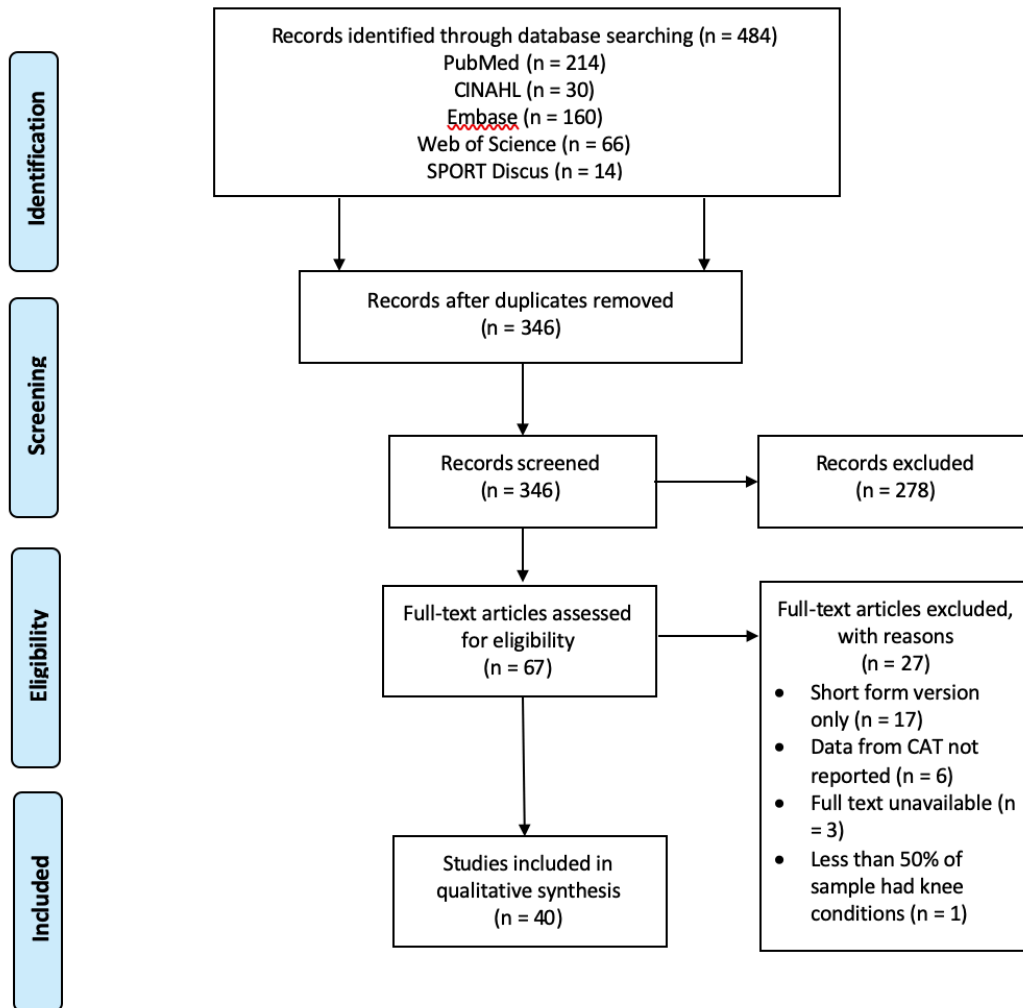


Figure 1: Flow diagram detailing the literature review and study selection process.

Table 1: CATs

CAT	Domains	Version(s)	Total # of Items in Item bank	Population(s)	Purpose(s) (studies (n))			Total # of studies	Level(s) of Evidence (range)
					Development	Testing	Outcome Assessment		
PROMIS	Physical Function (PF)	v1.0-1.1	124	Arthroscopic partial knee meniscectomy; meniscal injury; primary or revision ACLR; ACL injury; Knee OA scheduled for TKA; TKA; Knee OA; Knee pain; Knee arthroscopy; Elective knee surgery; patellofemoral malalignment and chondral disease; BTP fracture; EMD; MLKI; Knee impairments; Knee pain	0	16	23	30	Ia-IV
		v1.2	121						
		v2.0	160						
	Pain Interference (PI)	v1.1-1.2	40	Arthroscopic partial knee meniscectomy; primary ACLR; Knee OA undergoing TKA; Knee arthroscopy; Elective Knee Surgery; BTP fracture; EMD; MLKI	0	5	14	15	I Ib-IV
	Depression (D)	v1.0	28	Arthroscopic partial knee meniscectomy; Elective Knee surgery; knee arthroscopy; EMD; BTP fracture; patellofemoral malalignment and chondral disease; primary ACLR	0	2	9	9	I Ib-IV
	Mobility	v2.0	15	Multi-ligament Knee Injury (MLKI)	0	1	0	1	I Ib
	Pain Behavior	v1.0	20	Advanced Knee OA indicated for TKA	0	0	1	1	I Ib
	Social Satisfaction	v2.0	44	Elective Knee Surgery; Knee Arthroscopy	0	0	3	3	I Ib-IV
Anxiety	v1.0	29	0		0	4	4	I Ib-IV	
OA-CAT	Disability	NA	33	Knee and Hip Osteoarthritis; Knee OA undergoing TKA	1	1	1	2	NA
	OA-Function-CAT Functional Difficulty	NA	125		1	2	2	3	NA-I Ib
	OA-Function-CAT Functional Pain	NA	125		1	1	1	2	NA
FOTO	Knee Functional Status (FS)	NA	18	Patients with knee Impairments presenting to outpatient physical therapy clinics	1	6	0	7	I Ib

Table 2: Population characteristics by Diagnosis or Procedure.

Author, year	Sample size (n)	Age Mean (SD)	Sex (% female)	BMI (kg/m ²) Mean (SD)	Population	Primary Intervention	Primary application of CAT	Length of follow-up Mean (SD)
Meniscus								
Bernholt et al., 2018 ¹	75	53 (9.5)	37; 49% female	30.6 (6.1)	Patients undergoing partial knee meniscectomy	Arthroscopic partial meniscectomy (medial and/or lateral)	To determine if PROMIS PF-CAT and PI-CAT can determine early variations in post-operative outcomes.	6.3 weeks (range 4-12)
Hancock et al., 2017 ²	107	37.7 (16)	37; 34% female	29.6 (6.3)	Patients scheduled to undergo surgery for meniscal injury	NA	To evaluate the PROMIS PF-CAT correlations to other PROs that measure physical function, floor and ceiling effects, and test burden	NA
Anterior Cruciate Ligament (ACL)								
Scott et al., 2018 ³	139	24 (9.3) Range 14-53	67; 48.2% female	26.3 (4.7)	Patients undergoing primary ACLR or revision	Primary (n = 116) or revision (n = 23) ACLR with Hamstring autograft (n = 82), BTB autograft (n = 49), and tibialis anterior autograft (n = 8)	To compare the responsiveness and construct validity of PROMIS PF-CAT to established PRO's (SF-36, WOMAC, MARS, EQ-5D-5L).	Preop-2 years
Chen et al., 2018 ⁴	233	26.5 (12.5)	114; 49% female	NR	Patients undergoing primary ACLR	Primary ACLR	To determine if PROMIS scores could detect change in patients undergoing ACLR.	20.1 (12.5) days preop to 137 (27.2) days post op
Jevotovsky et al., 2019 ⁵	84	36.7 (12.82)	41; 48.8% female	NR	Patients undergoing primary ACLR and patients undergoing lumbar discectomy (DSC)	Primary ACLR	To compare the relative value of lumbar DSC by comparing the change in PROMIS PF, PI, and Depression CAT scores to	Range 40 days preop to 133 days post op

							a cohort of patients under ACLR	
Papuga et al., 2014 ⁶	106	29.5 (13.3)	57; 53.7% female	NR	Patients undergoing primary ACLR with BTB graft	ACLR with BTB graft	To compare PROMIS PF-CAT to GAITRite and IKDC to assess administration time, correlation to IKDC, and ability to detect those with a poor outcome.	Preop-52 weeks
Hancock et al., 2019 ⁷	100	26.4 (9.2)	45; 45% female	27.2 (6.2)	Patients indicated for operative management of ACL injury	NA	To evaluate the PROMIS PF-CAT correlations to other PROs that measure physical function, floor and ceiling effects, and test burden	NA
Total Knee Arthroplasty (TKA)								
Kagan et al., 2018 ⁸	91	63 (range 44-83)	42; 46% female	29.4 (range, 18.4-41.5)	Patients with OA scheduled to undergo primary unilateral TKA	Primary unilateral TKA	To characterize normal recovery patterns of unilateral TKA using PROMIS PF-CAT and PI-CAT.	Preop-1 year
Edelstein et al., 2018 ⁹	50	64-67 (7-8)	33; 66% female	32.8-34.2 (5.8)	Patients with advanced OA of the knee indicated for TKA	TKA with medial stabilized or posterior stabilized designs	To assess differences in outcomes between two TKA designs using a battery of tests including PROMIS PF, PI, PB-CATs, IKDC, Oxford, and VR-12.	Preop-24 months
Christensen et al., 2017 ¹⁰	303	64 (9)	197; 65% female	32 (7)	Patients with knee OA undergoing primary knee TKA with bicruciate retaining (BCR) or cruciate retaining (CR) implant designs	TKA with BCR or CR implant design	To determine differences in patient-reported physical function between two implant designs	18 months (range 2-32)
Pelt et al., 2019 ¹¹	141	64 (range 39-83)	85; 60% females	30.3 (range 18.8-44.3)	Patients who underwent bicruciate-retaining TKAs	Bicruciate-retaining TKA	To compare 3-year survivorship of novel implants and patient-reported outcomes (PROMIS PF CAT, PROMIS physical and mental short forms, and numeric pain scores).	3 year (range 0.34-4.9)

Kohring et al., 2018 ¹²	271 total (157 TKA)	62 (range 24-84)	158; 58% female	30 (range 15-85)	Patients undergoing THA or TKA classified as depressed, non-depressed, or medically treated depressed	TJA of the hip or knee (n = 157)	To determine if there were differences in patient-reported physical function between non-depressed, depressed, and medically treated depressed patients undergoing TJA.	Preop to Minimum of 1-year follow-up
Kohring et al., 2018 ¹³	540 total (300 TKA)	64 (range 28-95)	303; 56% female	29 (range 16-50)	Patients undergoing THA or TKA.	TJA of the hip or the knee (n = 300)	To evaluate the correlation between patient care satisfaction (Press Ganey) surveys and PROMIS PF-CAT, PROMIS global-10 mental and physical short forms in patients undergoing TJA.	Preop to minimum 1 year follow-up
McDonough et al. 2016 ¹⁴	66 (knee arthroplasty)	61 (9.3)	42; 63.6% female)	NR	Patients with hip or knee OA undergoing joint replacement or osteotomy	TKA, THA, periacetabular osteotomy	To evaluate the sensitivity to change of OA-CAT in patients after TKA, THA, or periacetabular osteotomy.	2 weeks pre-op to 3 months post-op
Knee Osteoarthritis (OA)								
Givens et al., 2018 ¹⁵	65	62.6 (8.9)	40; 61.5% female	32.6 (5.6)	Patients with end-stage OA (Kellgren-Lawrence grade 4) indicated for TKA	NA	To assess whether TUG, BMI, NPRS, and smoking status were predictors or PROMIS PF-CAT scores	NA
Chang et al., 2018 ¹⁶	104	65.6 (7.4)	85; 81.7% female	NR	Patients with knee OA and knee pain	6-week exercise program	To assess the degree of correlation and sensitivity to change of two CATs before and after an exercise program	Baseline to post-exercise (6 weeks)
Jette et al., 2009 ¹⁷	323 (182 knee-only)	61.8 (15.2)	209; 64.7% female	NR	Patients with a diagnosis of knee or hip OA	NA	To develop and evaluate a disability CAT for osteoarthritis research using IRT and CAT methodologies	NA
Jette et al., 2009 ¹⁸	328 (186 knee-only)	61.8 (15.1)	212; 64.5% female	NR	Patients with a diagnosis of knee or hip OA	NA	To develop and evaluate a functional pain and difficulty CATs using IRT and CAT methodologies	NA
Open and Arthroscopic Knee Surgeries								

Kenney et al., 2019 ¹⁹	76	48.9 (11.1)	36; 47.4% female	30.5 (6.6)	Patients undergoing knee arthroscopy for meniscal tear, chondromalacia, loose body, or synovial hyperplasia.	Knee arthroscopy	To compare the responsiveness of PROMIS PF-CAT and PI-CAT to disease-specific measures (IKDC, KOOS).	2 weeks post op-12 months
Wojahn et al., 2018 ²⁰	221	46.2 (range 14-76)	107; 48.4% female	29 (range 17.2-70)	Patients undergoing knee arthroscopy	Knee arthroscopy (partial meniscectomy, loose body removal, debridement, chondroplasty)	To determine if preoperative PROMIS scores (anxiety, depression, pain interference, and physician function) were predictive of higher opioid consumption after knee arthroscopy.	Preop to 6 weeks post-op
Miles et al., 2019 ²¹	412	39.4 (16.2)	181; 44% female	29.3 (5.7)	Patients undergoing elective knee surgery	Various arthroscopic and open procedures	To determine the correlation between PROMIS PF-CAT and IKDC in patients undergoing various elective knee surgeries.	NA
Nadarajah et al., 2019 ²²							To determine the correlation between PROMIS PI-CAT and NPS ratings in patients undergoing various elective knee surgeries	
Meredith et al., 2019 ²³	383	40.2 (13.6) opioid users; 39 (16.7) non-users	213; 44.6% female	30 (6.7) opioid users; 29.1 (5.6) Non-users	Patients undergoing elective knee surgery.	Various arthroscopic and open procedures	To determine differences in preoperative function, pain, and health status between patients who opioid users and non-opioid users using a battery of tests: PROMIS CAT domains: physical function, pain interference, depression, anxiety, social satisfactions, and fatigue; IKDC, subjective knee evaluation form, MODEMS, Tegner, IPAQ, MARS, knee pain, and body pain	NA
Stevens et al., 2019 ²⁴	399	39.2 (16.2)	176; 44% female	29.3 (6)	Patients undergoing knee surgery	Various arthroscopic and open procedures	To determine the relationship between preoperative expectations using: PROMIS	NA

							CAT domains: physical function, pain interference, depression, anxiety, social satisfactions, and fatigue; IKDC, subjective knee evaluation form, MODEMS, Tegner, IPAQ, MARS, knee pain, and body pain	
Patellofemoral Malalignment and Chondral Disease								
Carender et al., 2019 ²⁵	37	33.4 (11.1)	29; 73% female	31.6 (7.8)	Patients with patellofemoral malalignment and chondral disease scheduled to undergo surgery.	NA	To evaluate the PROMIS PF-CAT in a population of patients with patellofemoral malalignment and chondral disease relative to established patient-reported outcome (PRO) instruments.	NA
Bicondylar Tibial Plateau (BTP) Fractures								
Cavallero et al., 2018 ²⁶	56	49-51	22; 39.2% female	30-31	Patients undergoing surgical treatment for bicondylar tibial plateau fractures.	Operative fixation of bilateral tibial plateau fractures with locking or non-locking implants	PROMIS PF-CAT and PI-CAT used as secondary measure of outcomes in 2 surgical groups.	24.3 months post op
Virkus et al., 2018 ²⁷	52	48, one-stage fixation; 51, two-stage fixation	21; 40.1% female	30, one-stage fixation; 31 two-stage fixations	Patients with bicondylar tibial plateau (BTP) fractures treated with open reduction internal fixation	One-stage or two-stage fixation of BTP fractures.	To determine functional outcomes of 2 procedures for fixation of BTP fractures using PROMIS PF-CAT and PROMIS PI-CAT	Preop-21.8 months (range 6-41)
Extensor Mechanism Disruption								
Seifert et al., 2019 ²⁸	235	52.9-54.5	NR	31.9-36.7	Patients with unilateral or bilateral knee extensor mechanism disruption	Quadricep or patellar tendon rupture repair	To determine differences in outcomes of physical function, depression, and pain interference in two groups of patients with extensor mechanism disruption	268 days postoperatively
Multi-ligament Knee Injury								

Trasolini et al., 2019 ²⁹	30	36.3 (13.5)	4; 13% female	30.1 (8.4)	Patients with multi-ligament knee injury and minimum 1-year follow-up of surgical reconstruction of at least one ligament.	NA	To assess the correlation of PROMIS PF, PI, and mobility CAT domains correlation with the MLQOL, Tegner, and Lysholm.	NA
Knee Impairments								
Hart et al., 2005 ³⁰	1772 total (949 knee-only)	48 (17)	1134; 64% female	NR	Patients with hip, knee, or foot/ankle impairments receiving rehabilitation in outpatient clinics	NA	To develop a CAT to assess lower extremity function using items from the LEFS using IRT and CAT methodologies.	NA
Hart et al., 2008 ³¹	21,896	50 (17)	12,042; 55% female	NR	Patients with knee impairments receiving rehabilitation in outpatient clinics from January 2005-June 2007	NA	To assess the practicality of using a CAT in routine clinical practice and assess psychometric performance	NA
Wang et al., 2009 ³²							To determine meaningful interpretations of outcomes data from FOTO knee CAT.	
Hart et al., 2009 ³³	28,320 (14,160 U.S.; 14,160 Israel)	50.8 (16.7) U.S.; 48.8 (16.7) Israel	17,275; 61% female	NR	Patients with knee impairments receiving rehabilitation services at outpatient clinics in the U.S. and Israel from 2005-2008	NA	To determine the impact of differential item functioning in knee CAT items in English and Hebrew speaking patients receiving outpatient rehabilitation services.	NA
Wang et al., 2011 ³⁴	6,651	53 (17)	3,725; 56% female	NR	Patients with knee impairments receiving outpatient physical therapy from January 2005-October 2009.	NA	To determine if minimal clinically important change is dependent on patient's admission knee CAT scores and demographics.	NA
Deutscher et al., 2010 ³⁵	9134 (4567 Hebrew; 4567 Russian)	48.5 (16.7) Hebrew; 58.1 (13.9) Russian	2,366; 51.8% female Hebrew; 3,060; 67% female Russian	NR	Patients who spoke Hebrew or Russian with knee impairments from outpatient physical therapy clinics in Israel from 2005-2008	NA	To determine if differential item function is a possible cause for observed differences in outcomes between 2 distinct cultural groups in a polyglot society.	NA
Schmitt et al., 2014 ³⁶	8,955	51.7 (18.7)	5463;	NR	Patients with knee impairments receiving	NA	To determine if length of recall period affected the	0-180 days

			61% female		outpatient physical therapy from April 10, 2010 to December 30, 2011.		validity of GROC in determining true change over time by examining the correlations of GROC to Knee CAT scores.	
Robins et al., 2017 ³⁷	450 patients with knee complaints	36 (range 12-83)	185; 41% female	27 (6)	Patients who presented to a sports medicine clinic for knee or shoulder complaints.	NA	To evaluate the convergent validity, precision, and completion times for PROMIS PF-CAT compared to IKDC.	NA
Lizzio et al., 2019 ³⁸	581 (n = 220 knee)	46 (19)	291; 50% female	NR	Patients presenting to an orthopedic surgeon with various musculoskeletal diagnoses including 220 unspecified knee diagnoses.	NA	To determine if PROMIS CATs can be feasibly incorporated into daily practice.	NA
Kortlever et al., 2019 ³⁹	88	56 (13)	46; 52% female	33 (8.3)	Patients with unilateral or bilateral knee pain, atraumatic or traumatic in nature	NA	To determine the correlations between PROMIS PF-CAT and KOOS, JR.	NA
Hung et al., 2018 ⁴⁰	2226 total knee and hip	61.2 (12.8)	1220; 58.4% female	NR	Patients with knee or hip impairments	NA	To establish MCID of PROMIS PF-CAT, HOOS, JR, and KOOS JR	>180 days

Table 3: Summary of the measurement properties of all CATs used in patients with knee conditions.

Measure		Content validity	Internal Consistency	Construct Validity	Criterion validity	Reproducibility		Responsiveness	Floor or ceiling effects	Interpretability
						Agreement	Reliability			
PROMIS PF	PF-CAT v1.0	0	0	+ (i) ¹⁶	0	0	0	0	0	0
	PF-CAT v1.2	0	0	+ (c ³ , d ²⁵)	0	+ (b) ⁴⁰	0	+ (b ⁴⁰ , c ^{4,6}) ? (c) ³	+ (c ² , d ²⁵)	+ (a ^{8,12} , b ⁴⁰ , c ⁴ , d ²⁵ , e ¹⁹)
	PF-CAT v2.0	0	0	+ (d) ²⁵	0	0	0	0	+ (d) ²⁵	+ (d) ²⁵
	PF-CAT	0	0	+ (b ^{37,39} , c ⁶ , e ^{19,21} , g ⁷ , h ²)	+ (b ³⁷ , c ⁶ , e ^{19,21})	0	0	+ (c ⁶ , e ¹⁹)	+ (b ^{37,39} , e ²¹ , f ²⁹ , g ⁷ , h ²)	+ (e ¹⁹)
PROMIS PI	PI-CAT v1.1	0	0	0	0	0	0	+ (c) ⁴	0	+ (a ⁷ , c ⁴)
	PI-CAT v1.2	0	0	+ (e) ²²	+ (e) ²²	0	0	0	+ (e) ²²	0
	PI-CAT	0	0	+ (e) ⁹	0	0	0	+ (e) ⁹	+ (f) ²⁹	+ (e) ⁹
PROMIS D-CAT v1.0		0	0	0	0	0	0	+ (c) ⁴	0	+ (c) ⁴
PROMIS M-CAT		0	0	0	0	0	0	0	+ (f) ²⁹	0
OA-CAT disability		+ (i) ¹⁷	+ (a) ¹⁴	+ (i) ¹⁷	0	+ (i) ¹⁷	+ (i) ¹⁷	+ (a) ¹⁴	+ (i, a) ¹⁷	0
OA-CAT FD		+ (i) ¹⁸	+ (a) ¹⁴	? (i) ¹⁸	0	+ (i) ¹⁸	+ (i) ¹⁸	+ (a) ¹⁴	+ (i, a) ^{14,18}	0
OA-CAT FP		+ (i) ¹⁸	+ (a) ¹⁴	? (i) ¹⁸	0	+ (i) ¹⁸	+ (i) ¹⁸	+ (a) ¹⁴	+ (i, a) ^{14,18}	0
FOTO Knee FS		+ (b) ³⁰	+ (b) ³⁰	+ (b) ^{30,31,33,35}	0	+ (b) ³²	+ (b) ³⁰	+ (b) ^{31,34}	+ (b) ^{30,31}	+ (b) ^{31,32,34}
<p>+ meets quality criteria described by Terwee et al.⁴¹ ? doubtful design or method 0 no information found</p> <p>(a) TKA, (b) Knee impairments and pathology (c) ACLR (d) Patellofemoral malalignment and chondral disease (e) Patients undergoing knee surgery (f) Multi-ligament knee injuries (g) ACL injury indicated for surgery (h) Meniscal injury indicated for surgery (i) Patients with OA and knee pain</p>										

Table 4: Summary of MCIDs

CAT	Domain	Author & year	Patient population	Method	MCID		
PROMIS	PROMIS PF-CAT v1.2	Hung et al., 2018	Hip or knee pathology	Anchor (rating of improvement similar to GROC) – ROC curve	8.18 Range (1.97-8.18)		
				Anchor (rating of improvement similar to GROC) – mean change	7.9		
				Distribution (1/2 SD)	3.68-4.35 (range)		
				Distribution (1/3 SD)	2.45-2.9 (range)		
				Distribution (MDC 90%)	11.64-13.89 (range)		
				Distribution (MDC 95%)	14.14-16.5 (range)		
				Distribution (MDC 99%)	18.47-21.55 (range)		
		Chen et al., 2018	ACLR	Distribution	4.6		
		Kagan et al., 2018	TKA	Distribution	3.34		
		Kohring et al., 2018			3.7		
		PROMIS PF-CAT	Kenney et al., 2019	Knee arthroscopy	Anchor (IKDC)	3.3 Moderate CID: 5	
		PROMIS PI-CAT v1.1	Kagan et al., 2018	TKA	Distribution	-4.43	
			Chen et al., 2018	ACLR	Distribution	-4	
		PROMIS PI-CAT	Kenney et al., 2019	Knee arthroscopy	Anchor (IKDC)	-3.2 Moderate CID: -5.8	
	PROMIS D-CAT v1.0	Chen et al., 2018	ACLR	Distribution	-4.9		
FOTO	FOTO Knee FS	Hart et al., 2008	Knee impairments	Anchor (GROC)	9		
		Wang et al., 2009		Distribution - 95% CI for score estimate	6		
				Distribution - Percentile rank	Percentile	Intake FS score	Discharge FS score
					25 th	33	51
					50 th	42	61
		70 th		51	74		
		Anchor (GROC)		Entire sample: 9 >13, 9, 6, or 4 FS change scores represented MCH for intake scores of 0-34, 34-45, 45-54, and 54-100, respectively			
		Distribution - Functional Staging		Stage 1: 0-19 physiologic exercise only ambulator Stage 2: 20-29 limited household ambulator Stage 3: 30-37 independent household ambulator Stage 4: 38-47 limited community ambulator Stage 5: 48-62 independent community ambulator Stage 6: >62 Active community ambulator			
		Wang et al., 2011		Anchor (GROC)	Entire sample: 12 FS scores: 0-33: 14 FS units; 33-42: 12 FS units; 42-51: 5 FS units; >51: 5 FS units		

Table 5: Summary of the correlations of CATs to other outcome measures

Measure	PROMIS PF-CAT	PROMIS PI-CAT	PROMIS M-CAT	FOTO knee FS	OA-CAT disability
IKDC	Strong (4)	Strong (2)			
Lysholm	Strong (1)		Strong (1)		
Tegner ARS	Moderate (1)	Strong (1)	Moderate (1)		
SF-36 GH	Moderate (1); Low (2)				
SF-36 PF	Strong (4)				
SF-36 pain	Strong (2)				
KOOS, JR	Strong (1)				
KOOS pain	Strong (3)				
KOOS symptoms	Strong (2); Moderate (1)				
KOOS QoL	Strong (3); Moderate (1)				
KOOS sport	Strong (4)				
KOOS ADL	Strong (4)				
WOMAC function	Strong (1)				
WOMAC pain	Strong (1)				
WOMAC stiffness	Moderate (1)				
MARS	Low-Moderate (1); Low (1)	Moderate (1)			
EQ-5D-5L	Strong (4)				
Press Ganey	Low (1)				
MLQOL-AL	Strong (1)	Strong (1)	Strong (1)		
MLQOL-PI	Moderate (1)	Moderate (1)	Moderate (1)		
MLQOL-EI		Strong (1)			
MLQOL-SI		Strong (1)			
SANE	Strong (1)				
IPAQ		Moderate (1)			
NPS Knee pain		Strong (1)			
NPS Body pain		Moderate (1)			
GROC				Low-strong (1)	
HAQ-II					Strong (1)
OA-CAT	Strong (1)				
FTSST	Low-moderate (1)				
TUG	Moderate (1)				

Table 6: Feasibility of CATs in patients with knee conditions

CAT	Domain	Author & Year	Average # of questions (SD)	Completion Time in seconds (SD)	Fixed-length PRO	Average # of questions (SD)	Completion time in seconds (SD)
PROMIS	PF-CAT	Carender et al., 2019	5.6 (0.6)	NR			
		Scott et al., 2018	4 (1.0, 0.2, 1.9); 6.59 (3.57)	NR			
		Trasolini et al., 2019	4.53	NR			
		Hancock et al., 2017	4.4 (1.4)	NR			
		Hancock et al., 2019	4.2 (0.9)	NR			
		Papuga et al., 2014	Range, 6-12	85 (70)	IKDC	18	195 (53)
		Chang et al., 2018	4.14 (0.54)	108 (39.6)			
		Lizzio et al., 2019	4.2	66			
		Kohring et al., 2018 (press ganey)	4	NR			
		Kortlever et al., 2019	4	41 (2.2)	KOOS, Jr	7	29 (2)
		Robins et al., 2017	NR	55 (IQR 40-94)	IKDC	18	268 (IQR 204-434)
	Average		4.34	71			
	PI-CAT	Trasolini et al., 2019	5.83	NR			
		Lizzio et al, 2019	4.3	43			
Average		5.1	43				
M-CAT	Trasolini et al., 2019	5.53	NR				
D-CAT	Lizzio et al. 2019	6.4	46				
OA-CAT	Functional Difficulty	Chang et al., 2018	10.03 (2.62)	102 (108)			
		McDonough et al., 2016	5	67 (37)	WOMAC physical function	17	120 (47)
	Jette et al., 2009-2	Fixed lengths: 5, 10, 15	NR	WOMAC	24	NR	
	Functional Pain Disability	McDonough et al., 2016	5	52 (25)	WOMAC pain	5	41 (36)
Jette et al., 2009-1	Fixed lengths: 5, 10, 15	NR	HAQ-II	10	NR		
FOTO	Lower extremity function	Hart et al., 2005	6 (1.4)	NR	LEFS	18	NR
		Hart et al., 2008	6.86 (3.25)	NR			
		Average	6.43				

Appendix 1: Search Strategy

Table 1: PubMed Search Strategy

Search	Query	Items found
#10	Search (((((knee*) OR ACL*) OR anterior cruciate ligament*) OR menisc*)) AND (((computer* adapt* test*) OR "item response theory") OR PROMIS) OR "patient reported outcomes measurement information system")	160
#9	Search (((knee*) OR ACL*) OR Anterior cruciate ligament*) OR menisc*	187811
#9	Search menisc*	17533
#8	Search Anterior cruciate ligament*	20599
#7	Search ACL*	22938
#6	Search knee*	168615
#5	Search (((computer* adapt* test*) OR "item response theory") OR PROMIS) OR "patient reported outcomes measurement information system"	11050
#4	Search "patient reported outcomes measurement information system"	940
#3	Search PROMIS	1523
#2	Search "item response theory"	3068
#1	Search computer* adapt* test*	6848

Table 2: General search strategy for other databases (Embase, CINAHL, Web of Science, Sport Discus)

((knee OR acl OR cruciate) AND ligament OR mcl OR lcl OR pcl OR menis* OR patellofemoral OR tibiofemoral OR patella OR tibia OR femur) AND computer* AND adapt* AND test*

Appendix 2: Risk of Bias

Table 1: Cochrane Risk of Bias 2

Response options: Low, Some concerns, High

	Randomization Process	Deviation from Intended Interventions	Missing outcome data	Measurement of the outcome	Selection of the reported result	Overall
Edelstein, 2018	Some Concerns	Some Concerns	Low	Low	Low	Some concerns

Table 2: RoRANS

Response options: Low, Unclear, High

Author, year	Selection of participants	Confounding variables	Intervention (exposure) measurement	Blinding of outcome assessment	Incomplete Outcome data	Selective outcome reporting	Overall	Rationale
Cavellero, 2018	Low – included all patients in a given time period	Low	N/A	Unclear	High; no measure of variability in the outcome reporting	High	Unclear	Did not report preoperative PROMIS data; no measure accounting for variability in outcome reporting
Bernholt, 2018	Low – included all patients in a given time period	High – variability in time since surgery	N/A	Low	Low	Low	High	Variation of when PROMIS was delivered post-operatively could bias the results
Seifert, 2019	Low – included all patients in a given time period	Low	N/A	Low	High; no measure of variability in the outcome reporting	Low	High	Missing PRO data from UEMD could affect study outcome; no measure of variability in outcome reporting
Christensen, 2017	High – highly specific study population not representative of TKA	Low	N/A	Low	Low	High	High	Retrospective selection of participants; Did not include preoperative PROMIS data
Meredith, 2019	Unclear – included all patients but excluded those without preop data	Low	N/A	Low	Low	Low	Unclear	Unclear if excluding those without preop data could have confounded the results
Author, year	Selection of participants	Confounding variables	Intervention (exposure) measurement	Blinding of outcome assessment	Incomplete Outcome data	Selective outcome reporting	Overall	Rationale
Miles, 2019	Low – included all patients in a given time period	Low	N/A	Low	Low	Low	Low	Included all patients in a given time period;

Nadarajah, 2019	Low – included all patients in a given time period	Low	N/A	Low	Low	Low	Low	Included all patients in a given time period;
Lizzio, 2019	Low – included all patients in a given time period	Low	N/A	Low	Low	Low	Low	Included all patients in a given time period;
Givens, 2018	Low – study sample was representative of TKA population	Low	N/A	Low	Low	Low	Low	study sample was representative of TKA population
Kortlever, 2019	Low	Low	N/A	Low	Low	Low	Low	
Virkus, 2018	High – retrospective selection of injury severity	Low	Low	Low	Unclear	Low	Unclear	Retrospective selection of injury severity
Pelt, 2019	Low – patients recruited consecutively and data collected prospectively	Low	N/A	Low	Low	High – did not include preoperative PROMIS data	Low	Lack of preoperative data unlikely to affect the results of the study
Author, year	Selection of participants	Confounding variables	Intervention (exposure) measurement	Blinding of outcome assessment	Incomplete Outcome data	Selective outcome reporting	Overall	Rationale
Kagan, 2018	Low – prospective data collection	High – did not evaluate patients with very low scores separately from patients	N/A	Low	Low	Low	Low	Diverse population of patients with prospective data collection

		with very high scores						
Kohring, 2018 (depression)	High – retrospective chart review performed to identify depressed patients	High – small number of depressed patients included	N/A	Low	Low	Low	High	retrospective chart review performed to identify depressed patients; small number of depressed patients included
Kohring, 2018 (press ganey)	High – excluded patients without Press-Ganey scores	High – recall bias for Press Ganey survey administration	N/A	Low	High – only 28% of patients completed the survey	Low	High	retrospective review of secondary data; recall bias; incomplete outcome data
Wojahn, 2018	High – sample did not include patients who had been taking opioids longer than 1 month preoperatively	High – did not account for variations in surgical; did not account for variables that could affect opioid use including perioperative pain and knee OA grade.	Low	High – not blinded to participation	Low	Low	High	Selection bias, participants not blinded to participation; did not account for variations in surgical; did not account for variables that could affect opioid use including perioperative pain and knee OA grade.
Jevotovsky, 2019	High – limited participants based on select data	Low – used propensity score matching to account for baseline differences	Low – data obtained from medical records	Low	Low	Low	Low	Collection of data was part of routine practice; propensity score matching was used to account for differences in cohorts.
Hancock, 2017	Low – all participants included in	High – major confounding	N/A	Unclear	Low	Low	Unclear	Confounding variables were not controlled for in analysis; Unclear if participants were

	same time period	variables were not considered						blinded or if that would affect the results.
Hancock, 2019	Low – all participants included in same time period	High – confounding variables such as concomitant injuries not considered	N/A	Unclear	Low	Low	Unclear	Confounding variables such as concomitant injuries not considered; Unclear if participants were blinded or if that would affect the results.
Author, year	Selection of participants	Confounding variables	Intervention (exposure) measurement	Blinding of outcome assessment	Incomplete Outcome data	Selective outcome reporting	Overall	Rationale
Carender, 2019	Low – all participants included in the same time period	High – confounding variables such as concomitant injuries not considered	N/A	Unclear	Low	Low	Unclear	Confounding variables such as concomitant injuries not considered; Unclear if participants were blinded or if that would affect the results.
Kenney, 2019	Low – all participants prospectively enrolled	High – variable amounts of follow-up	N/A	Low – collected as part of routine clinical practice	High – missing follow-up data	Low – addressed all predefined hypotheses	High	Highly variable amounts of follow-up; miss data secondary to variable follow-up; no measure accounting for variability in outcome reporting.
Chen, 2018	Low – data was collected prospectively	High – length of follow-up not sufficient to capture full return to sport	N/A	Low – data collected as part of routine clinical practice	Low; included measure of variability	Low – addressed all predefined hypotheses	Low	Prospective data collection; data collected as part of routine clinical practice; accounted for variability in analysis; addressed all predefined hypotheses
Papuga, 2014	High – only included patients with BTB autograft; not	Unclear – Unclear if PT interventions were conducted at	N/A	Low – evaluated as a part of routine	Low – included measure of variability	Unclear – unclear how patients were selected to be	Unclear	Unclear if patients underwent therapy at the same or different institutions.

	representative of full spectrum of ACL grafts	the same institution with the same therapists		clinical practice		included in the analysis		
Robins, 2017	Low – included all patients in a given time period	High – completion time of instruments could have been influenced by clinic flow; fatigue bias	N/A	Low – collected as part of routine clinical practice	High – excluded incomplete outcome data	Low – all hypotheses and outcomes were addressed	High	Confounding variables such as fatigue bias were not addressed in outcome measure administration; patients without complete data were excluded
Stevens, 2019	Low – included all patients undergoing knee surgeries at an institution	High – did not account for differences between procedures; fatigue bias	N/A	Low – blinding not likely to affect outcomes of self-report measures	Low – included data from all patients	Low – all hypotheses were addressed	High	Did not account for differences between procedures; fatigue bias
Scott, 2018	Low – included consecutive patients under ACLR or revision	Unclear – adjusted for concomitant injuries using subgroup analyses; possible fatigue bias	N/A	Low – blinding not likely to affect the outcomes of self-report measures	High – did not include all patients in the final analysis	Low-addressed all predefined outcomes	Unclear	Unclear if fatigue of completing outcome multiple assessments could bias results; did not include all patients in the final analysis
Trasolini, 2019	High – retrospective selection of patients	High – major confounding factors not considered; variable time since surgery	High – recall bias	Unclear	Low	Low – addressed all predefined outcomes	High	Retrospective selection of patients; major confounding factors not considered; recall bias; variable time since surgery
Chang, 2018	Low – included all patients that	High – major confounders	N/A	Low	High – lost patients to	Low – addressed all	High	major confounders identified but not controlled for

	were part of an RCT with knee OA	identified but not controlled for during analysis			follow-up without explanation	predefined outcomes		during analysis; patients lost to follow-up
Schmitt, 2014	Low – included all patients with knee impairments in the FOTO database	Low – controlled for confounding variables in baseline demographics	Low – controlled for variable confounders such as recall periods.	Low – data collected as part of routine clinical practice	High – patients with missing data were significantly different for multiple demographics	Low	High	Patients with missing data were different from those included

Table 3: COSMIN Risk of Bias Checklist

Response options: V = very good, A = adequate, D = doubtful, I = inadequate

	Structural validity	Internal Consistency	Measurement invariance	Reliability	Measurement error	Criterion validity	Construct Validity	Responsiveness
FOTO Knee FS								
Hart, 2005	A	V	V	D	D		A	
Hart, 2008							V	V
Hart, 2009							V	
Wang, 2009					V			
Wang, 2011								V
Deutscher, 2010							V	
OVERALL	A	V	V	D	D		A	V
OA-CAT Disability								
Jette, 2009	A	V	A	D	D		V	
McDonough, 2016		V		D			A	I
OVERALL	A	V	A	D	D		A	I
OA-CAT Functional Difficulty								
Jette, 2009	A	V	A	D	D		V	
Chang, 2018				D				D

McDonough, 2016		V		D			A	I
OVERALL	A	V	A	D	D		A	I
OA-CAT Functional Pain								
Jette, 2009	A	V	A	D	D		V	
McDonough, 2016		V		D			A	I
OVERALL	A	V	A	D	D		A	I
	Structural validity	Internal Consistency	Measurement invariance	Reliability	Measurement error	Criterion validity	Construct Validity	Responsiveness
PROMIS PF-CAT								
Hung, 2018					V			V
Hancock, 2017							V	
Scott, 2018							V	D
Papuga, 2014						V	V	
Chen, 2018								V
Hancock, 2019							V	
Carender, 2018							V	
Trasolini, 2019							I	
Kohring 2018 (press ganey)							V	
Givens, 2018							I	
Chang, 2018				D			V	D
Kenney, 2019						V	V	V
Miles, 2019						V	V	
Kortlever, 2019							V	
Robins, 2017						V	V	

	Structural validity	Internal Consistency	Measurement invariance	Reliability	Measurement error	Criterion validity	Construct Validity	Responsiveness
PROMIS PI-CAT								
Nadarajah, 2019						V	V	
Trasolini, 2019							I	
Kenney, 2019						V		V
Chen, 2018						V		V
PROMIS D-CAT								
Chen, 2018								V

Appendix 3: Study characteristics

Table 1: Study characteristics

Author, year	Study design	Level of evidence	Study type	CAT(s) used			Domains	# of domains
				PROMIS-CAT	FOTO	OA-CAT		
Bernholt 2018	Case series; retrospective	IV	Outcome assessment	x			PROMIS PF-CAT, PI-CAT, D-CAT	3
Carender 2019	Cohort study (diagnosis); prospective	IIb	Testing	x			PROMIS PF-CAT v1.2 and v2.0	1
Scott 2018	Cohort study (diagnosis); Prospective	IIb	Outcome assessment; Testing	x			PROMIS PF-CAT v1.2	1
Trasolini 2019	Cohort study; prospective	IIb	Testing	x			PROMIS PF-CAT, PROMIS PI-CAT, PROMIS M-CAT	3
Hancock 2017	Cross sectional; prospective	IIIb	Testing	x			PROMIS PF-CAT	1

Hancock 2019	Cross sectional; prospective	IIIb	Testing	x		PROMIS PF-CAT	1
Kagan 2018	Cohort study; prospective	IIb	Outcome assessment; Testing	x		PROMIS PF-CAT v 1.2, PROMIS PI-CAT v 1.1	2
Kenney 2019	Cohort study (diagnosis); prospective	IIb	Outcome assessment; Testing	x		PROMIS PF-CAT, PROMIS PI-CAT	2
Cavallero 2018	Cohort study (therapeutic); retrospective	IIIb	Outcome assessment	x		PROMIS PF-CAT; PROMIS PI-CAT	2
Chen 2018	Cohort study; retrospective	IIIb	Outcome assessment; Testing	x		PROMIS PF-CAT v 1.2, PROMIS PI-CAT v 1.1, PROMIS D-CAT v 1.0	3
Edelstein 2018	Randomized Controlled Trial; prospective	IIb	Outcome assessment	x		PROMIS PF-CAT, PROMIS PI-CAT, PROMIS PB-CAT	3
Jevotovsky 2019	Cohort study; retrospective	IIIb	Outcome assessment	x		PROMIS PF-CAT v1.2, PROMIS PI-CAT v1.1, PROMIS D-CAT v1.0	3
Papuga 2014	Cohort study; prospective	Ib	Outcome assessment; Testing	x		PROMIS PF-CAT	1
Seifert 2019	Case series; retrospective	IV	Outcome assessment	x		PROMIS PF-CAT, PROMIS PI-CAT, PROMIS D-CAT	3
Christensen 2017	Cohort study (therapeutic); retrospective	IIIb	Outcome assessment	x		PROMIS PF-CAT	1
Givens 2018	Cross sectional; prospective	IIIb	Testing	x		PROMIS PF-CAT	1
Hung 2018	Cohort study; prospective	Ia	Testing	x		PROMIS PF-CAT v1.2	1
Kohring 2018 (depression)	Cohort study; retrospective	IIb	Outcome assessment	x		PROMIS PF-CAT v1.2	1
Kohring 2018 (press ganey)	Cohort study; retrospective	IIb	Outcome assessment	x		PROMIS PF-CAT v1.2	1
Kortlever 2019	Cross-sectional; prospective	IIIb	Testing and outcome assessment	x		PROMIS PF-CAT	1
Lizzio 2019	Cross-sectional; prospective	IIIb	Outcome assessment	x		PROMIS PF-CAT v2.0, PROMIS PI-CAT v1.1, PROMIS D-CAT v1.0	3
Meredith 2019	Cross-sectional; prospective	IIIb	Outcome assessment	x		PROMIS PF-CAT, PROMIS PI-CAT, PROMIS D-CAT, PROMIS social satisfaction-CAT, PROMIS anxiety- CAT; v1.2	6
Miles 2019	Cross-sectional; prospective	IIIb	Testing; Outcome assessment	x		PROMIS PF-CAT	1
Nadarajah 2019	Cross sectional; prospective	IIIb	Testing; Outcome assessment	x		PROMIS PF-CAT, PROMIS PI-CAT, PROMIS D-CAT, PROMIS social satisfaction-CAT, PROMIS anxiety- CAT; v1.2	6

Wojahn 2018	Cohort study (prognosis); prospective	IV	Outcome assessment	x			PROMIS PF-CAT, PROMIS PI-CAT, PROMIS anxiety-CAT, PROMIS D- CAT	4
Virkus 2018	Cohort study (therapeutic); retrospective	IIIb	Outcome assessment	x			PROMIS PF-CAT, PROMIS PI-CAT	2
Pelt 2019	Case series; retrospective	IV	Outcome assessment	x			PROMIS PF-CAT v2.0	1
Robins 2017	Cross-sectional; retrospective	IIIb	Testing	x			PROMIS PF-CAT	1
Stevens 2019	Cohort study; prospective	IV	Outcome assessment	x			PROMIS PF-CAT, PROMIS PI-CAT, PROMIS D-CAT, PROMIS social satisfaction-CAT, PROMIS anxiety- CAT	6
Chang 2018	Cohort study; prospective	IIb	Testing; Outcome assessment	x		x	PROMIS PF-CAT v1.0, OA-CAT Functional Difficulty	2
Jette 2009-1	Clinical measurement study; cross sectional	NA	Development			x	OA-DISABILITY-CAT	1
Jette 2009-2	Clinical measurement study; cross sectional	NA	Development			x	OA-FUNCTION-CAT (functional difficulty, functional pain)	2
McDonough 2016	Clinical measurement study; longitudinal	NA	Testing; outcome assessment			x	OA-CAT Functional difficulty, functional pain, and disability subscales	3
Hart 2005	Cohort study; retrospective (secondary analysis)	IIb	Development		x		FOTO knee FS CAT	1
Hart 2008	Cohort study; retrospective (secondary analysis)	IIb	Testing		x		FOTO knee FS CAT	1
Hart 2009	Cohort study; longitudinal; retrospective (secondary analysis)	IIb	Testing		x		FOTO knee FS CAT	1
Wang 2009	Cohort study; prospective	IIb	Testing		x		FOTO knee FS CAT	1
Wang 2011	Cohort study; prospective	IIb	Testing		x		FOTO knee FS CAT	1
Deutscher 2010	Cohort study; prospective (secondary analysis)	IIb	Testing		x		FOTO knee FS CAT	1
Schmitt 2013	Cohort study; prospective	IIb	Testing		x		FOTO knee FS CAT	1

Appendix 4

Table 1: Psychometric properties of CATs by Pathology or Procedure

CAT	Author & year	Population	Content validity	Internal consistency	Construct validity	Criterion validity	agreement	reliability	responsiveness	Floor or ceiling effects	Interpretability
Meniscus											
PROMIS PF-CAT	Hancock et al., 2017	Meniscal injury indicated for surgery			The PROMIS PF CAT had a high correlation with the SF-36 Physical Functioning (PF) ($r = 0.82, p < 0.01$) and KOOS Sport ($r = 0.76, p < 0.01$) scores; a high-moderate correlation with the KOOS Quality-of-Life (QOL) ($r = 0.63, p < 0.01$) and EQ-5D ($r = 0.62, p < 0.01$) instruments; and a moderate correlation					No floor or ceiling effects	

					with the SF-36 Pain ($r = 0.60, p < 0.01$), KOOS Symptoms ($r = 0.57, p < 0.01$), KOOS Activities of Daily Living (ADL) ($r = 0.60, p < 0.01$), and KOOS Pain ($r = 0.60, p < 0.01$) scores.						
					Discriminatory analysis to assess unidimensionality						
					3.1% of unexplained variance						
ACL											
PROMIS PF-CAT v1.2	Scott et al., 2018	Primary ACLR or revision			Evaluated construct validity by assessing correlation to other measures. Correlated strongly with other measures of physical function and had low correlations to MARS until 2-year follow-up and moderate correlations with SF-36 GH. (n = 139)				Does not meet criteria: Used effect size: Baseline to 6 weeks: 0.12 (-0.07=0.31) SRM: 0.1 (-0.09-0.29) 6 weeks to 6 months: 1.43 (1.17 to 1.69) SRM: 1.197 (0.95-1.44)	No floor effects Ceiling effect: 2 (1.4%) at 6 months 10 (9%) at 2 years	
PROMIS PF-CAT v1.2	Chen et al., 2018	ACL reconstruction							AUC: 0.826 P<0.001		MCID Distribution: 4.6 Cutoff for achieving MCID: 42.5
PROMIS PI-CAT v1.1	Chen et al., 2018	ACL reconstruction							AUC: 0.715 P<0.001		MCID Distribution: 4 Cutoff for achieving MCID: 56.2
PROMIS D-CAT v1.0	Chen et al., 2018	ACL reconstruction							AUC: 0.836 P<0.001		MCID Distribution: 4.9

											Cutoff for achieving MCID: 44.8
PROMIS PF-CAT	Papuga et al., 2014	ACLR with BTB autograft			Convergent validity with IKDC at all time points (r = 0.8954)	Criterion validity with IKDC at all time points				At final visit: ROC AUC: 0.851 (95% CI, 0.75-0.95)	
PROMIS PF-CAT	Hancock et al., 2019	ACL injury indicated for surgery			Each patient completed the PROMIS PF CAT, EQ-5D, KOOS, Marx, and SF-36 subscales of physical function, general health and pain in random order on a computer kiosk during their preoperative clinic visit. Both generic and anatomical site-specific PRO instruments were included to evaluate for both divergent and convergent construct validity. (see table)					No floor or ceiling effects	
Total Knee Arthroplasty (TKA)											
PROMIS PF-CAT v1.2	Kagan et al., 2018	TKA									MCID: Distribution 3.34
PROMIS PI-CAT v1.1	Kagan et al., 2018	TKA									MCID Distribution: 4.43
PROMIS PF-CAT v1.2	Kohring et al., 2018 (depression)	TKA or THA									MCID: Distribution 3.7
PROMIS PF-CAT v1.2	Kohring et al., 2018 (press ganey)	TKA or THA			For random selection of visits, preop, post op, and 1 year change, there were little to no correlations between press ganey and PROMIS PF CAT R = 0.04, -0.10, -0.18, 0.2.						
OA-CAT disability	McDonough et al., 2016	Knee OA undergoing TKA			Cronbach's alpha: 0.88					Effect size: 0.81	No floor effects

									(95% CI, 0.56, 1.14) MDC (90): 7.77	>10% at any time 6.1% ceiling effect	
OA-CAT FD	McDonough et al., 2016	Knee OA undergoing TKA		Cronbach's alpha: 0.88					Effect size: 1.06 (95% CI, 0.74, 1.45) MDC (90): 6.79	No floor effects >10% at any times	
OA-CAT FP	McDonough et al., 2016	Knee OA undergoing TKA		Cronbach's alpha: 0.86					Effect size: 1.15 (95% CI 0.82-1.6) MDC (90): 6.88	No floor effects >10% at any times 6% ceiling effects	
Knee Osteoarthritis (OA)											
OA-CAT disability	Jette et al., 2009-1	Knee or hip OA	Provides detailed description of item bank and CAT development	Used confirmatory factor analyses to assess scale unidimensionality and IRT methods to calibrate the items and examine the fit of the data. No cronbach alpha provided	Construct validity to HAQ-II for 5,10, and 15-item banks (r = 0.71-74)			Includes graphical representation of SEM, however, Used simulation of CAT scores which may overestimate agreement levels	Reliability was estimated as follows: $1/1 + (\text{standard error})^2$ 95% of reliability estimates were over 0.83	No floor effect 4.02% ceiling effect	
OA-CAT FD	Jette et al., 2009-2	Knee or hip OA	Provides detailed description of item bank and CAT development					Used simulation of CAT scores which may overestimate agreement levels	$1/1 + (\text{standard error})^2$ 95% of reliability estimates were over 0.97	No floor effects 0.6% ceiling effects	
OA-CAT FP	Jette et al., 2009-2	Knee or hip OA	Provides detailed description of item bank and CAT development					Used simulation of CAT scores which may overestimate agreement levels	$1/1 + (\text{standard error})^2$ 95% of the reliability estimates were over 0.96	No floor effects 0.6% ceiling effects	

PROMIS PF-CAT	Givens et al., 2018	Patients with knee OA awaiting TKA			No hypotheses formed; moderate negative correlation with TUG					
PROMIS PF-CAT v1.0	Chang et al., 2018	Patients with knee OA and knee pain			Convergent validity with OA-CAT FD (r = 0.65-0.69) Weak correlations with FTSST: (-0.21 to -0.34)			reliability, calculated as [1-average (squared standard error/vari- ance of score)] PROMIS PF-CAT: 0.85 OA-CAT FD: 0.89-0.9	Effect size: PROMIS PF-CAT: 0.42 (95% CI, 0.24-0.63) small OA-CAT FD: 0.62 (95% CI, 0.43-0.87) Medium FTSST: 0.5	
Open and Arthroscopic Procedures										
PROMIS PF-CAT	Kenney et al., 2019	Knee arthroscopy			Convergent validity with IKDC	Criterion validity with IKDC (r = 0.76)			Based on 11.5 change in IKDC: ROC AUC 0.88 (0.04) (95% CI, 0.81-0.96) Based on 20.5 change in IKDC: Mod CID 0.89 (0.04) (95% CI, 0.81-0.96) Effect size index: 0.04, 0.54, 1.14, 1.33, and 0.86 for 2 weeks, 3 months, 6 months, 12 months, and all, respectively	MCID: Anchor (IKDC) 3.3 Moderate CID 5

PROMIS PI-CAT	Kenney et al., 2019	Knee arthroscopy			Convergent validity with IKDC				Based on 11.5 change on IKDC: ROC AUC 0.85 (0.04) (95% CI, 0.77-0.94) Based on 20.5 change in IKDC: Mod CID 0.89 (0.04) (95% CI, 0.82-0.96) Effect size index: 0.44, 0.54, 1.25, 1.27, 0.93, for 2 weeks, 3 months, 6 months, 12 months, and all, respectively	MCID: Anchor (IKDC) -3.2 Moderate CID -5.8
PROMIS PF-CAT	Miles et al., 2019	Patients undergoing knee surgery			Convergent validity with IKDC (r = 0.71)	Criterion validity with IKDC (r = 0.71)				1 (0.5%) at floor No ceiling effects
PROMIS PI-CAT v1.2	Nadarajah et al., 2019	Patients undergoing knee surgery			Convergent validity with IKDC (r = -0.73)	Criterion validity with IKDC (r = -0.73)				No floor or ceiling effects
Patellofemoral Malalignment and Chondral Disease										
PROMIS PF-CAT v1.2, 2.0	Carender et al., 2019	Patellofemoral malalignment and chondral disease			Evaluated convergent and discriminant validity using spearman correlation coefficients. High to high moderate correlations with all measures except for moderate to weak correlations with KOOS symptoms, WOMAC stiffness, and MARS					No floor or ceiling effects observed for PROMIS PF-CAT

					Sample size to small (n = 39)						
Multi-ligament Knee Injuries											
PROMIS PF-CAT	Trasolini et al., 2019	Multi-ligament knee injuries			Evaluated convergent validity of Lysholm, Tegner, MLQOL sample size too small (n = 30)					No floor or ceiling effects	
PROMIS PI-CAT	Trasolini et al., 2019	Multi-ligament knee injuries			Evaluated convergent validity of Lysholm, Tegner, MLQOL-AL r = 0.71 sample size too small (n = 30)					No floor or ceiling effects	
PROMIS M-CAT	Trasolini et al., 2019	Multi-ligament knee injuries			Evaluated convergent validity of Lysholm, Tegner, MLQOL sample size too small (n = 30)					No floor or ceiling effects	
Knee Impairments											
FOTO Knee FS development	Hart et al., 2005	Hip, knee, foot, ankle impairments	Provides detailed description of item bank and CAT development	Cronbach's alpha: 0.96	Discriminant validity: If the RP > 1 the measurement method in the numerator is superior in differentiating function compared to method in the denominator. The greater the F-value, the greater the amount of systematic variance a measurement method accounts for and, therefore, the greater its ability to discriminate groups of patients. Knee FS Acuity RP = 1 (95% CI, 0.6, 1.6) Age RP = 0.9 (0.7, 1.2)			person reliability: 0.95			
FOTO Knee FS	Hart et al., 2008	Knee impairments; receiving			Known group construct validity for age group,				FS discharge scores:	Intake only:	MCII: 9 FS change

		outpatient therapy in the United States			symptom acuity, surgical history, comorbidities, and exercise history				<p>MDC (90): 75 MDC (95): 72</p> <p>Upper CI limits of MDC90 and MDC95 values for all patients were 11.6 and 13.8, respectively</p> <p>MDC90 and MDC95 upper CI limit values for the 92% of patients with FS intake scores between 20 and 70 were 6.2 and 7.3, respectively.</p> <p>ROC AUC: 0.73 (95% CI, 0.708, 0.757) for all groupings of intake scores. SE = 0.013</p>	<p>Floor effect: 1 scored 0 (0.7%) 0-5</p> <p>Ceiling effect: 86 (0.4%) 95-100</p> <p>Discharge: floor effect: 6 (0.05%) scored between 0 and 5</p> <p>Ceiling effect: 536 (5%) scored between 95-100</p>	(67% reported FS change scores equal to or greater than MCII.
FOTO Knee FS	Hart et al., 2009	Knee impairments; receiving outpatient therapy in Israel and the United States			Differential item function (DIF) Several items had non-uniform DIF for language, but FS estimates were strongly concordant (ICC (2,1) >0.97) and differences represented <0.4% of the FS standard deviation. This was not clinically relevant DIF.						

FOTO Knee FS	Wang et al., 2009	Knee impairments; receiving outpatient therapy in the United States					Standard error of measurement calculated: 2.99 FS points		ROC analysis mentioned, but values were derived from earlier study by Hart et al.	<p>95% CI for score estimate: All patients FS point estimate: +/- 6; FS scores between 20-70: 95% CI: +/-5</p> <p>Percentile rank of FS scores: Intake 25th: FS = 33 50th FS = 42 75th FS = 42 Discharge 25th FS = 51 50th FS = 61 75th FS = 74</p> <p>Responsiveness: Statistically reliable change: 9 FS change units >13, 9, 6, or 4 FS change scores represented MCII for intake scores of 0-34, 34-45, 45-54, and 54-100, respectively.</p> <p>Functional staging: 6 stages 19 is cut score between stage 1 and 2. 29 is cut score between 2 and 3, and so on.</p>
-----------------	-------------------	---	--	--	--	--	--	--	---	--

FOTO Knee FS	Wang et al., 2011	Knee impairments; receiving outpatient therapy in the United States							ROC AUC: (95% CI) Entire sample: 0.75 (0.73-0.77) Intake FS: 0-33: 0.75 (0.72-0.79) 33-42: 0.78 (0.74-0.81) 42-51: 0.79 (0.76-0.83) >51: 0.76 (0.72-0.8)		MCIH: Anchor-based ROC cut point (MCIH): Entire sample: 12 0-33: 14 33-42: 12 42-51: 5 >51: 5
FOTO Knee FS	Deutscher et al., 2010	Knee impairments; receiving outpatient physical therapy in Israel.			DIF: Several items showed DIF, but FS estimates that accounted for DIF were equal (ICC (2,1) >0.999) and differences accounted for <0.3% of the unadjusted FS estimates. Not clinically relevant. Changes between groups likely reflect true changes.						
PROMIS PF-CAT	Robins et al., 2017	Patients who presented to a sports medicine clinic for knee complaints			Convergent validity with IKDC (r = 0.75) (95% CI 0.7-0.8)	Criterion validity with IKDC (r = 0.75)					1 (0.2%) floor effect 2 (0.4%) ceiling effect
PROMIS PF-CAT	Kortlever et al., 2019	Patients with knee pain			Convergent validity with KOOS, JR (r = 0.74)						No floor or ceiling effects
PROMIS PF-CAT v1.2	Hung et al., 2018	Hip or knee pathology					Calculated MDC at various confidence levels: 90, 95, and 99%:		Utilized ROC to determine cutoff points that maximize sensitivity		Mean change: 7.29-8.41 Median MCID Anchor: 7.9

													At 6 months: 90%: 13.89 95%: 16.5 99%: 21.55		and specificity ROC cut- off: 1.97, 3.44, 3.52, and 8.18 at 3, >3, 6, and >6-month follow-up.		Distribution 4.4 Range: 2.45- 21.55
--	--	--	--	--	--	--	--	--	--	--	--	--	---	--	--	--	--

Appendix 5: Correlations of CATs to other measures

Table 1: Correlation of PROMIS PF-CAT to other outcome measures

Domain	Author & year	Patient-reported outcome measures																							
		IKDC	Lysholm	Tegner	SF-36 GH	SF-36 PF	SF-36 pain	KOOS, Jr	KOOS Pain	KOOS Symptoms	KOOS Quality of Life	KOOS Sport	KOOS ADL	WOMAC Function	WOMAC Pain	WOMAC Stiffness	MARS	EQ-5D-5L	Press Ganey	MLQOL-AL	MLQOL-PI	SANE			
PROMIS PF-CAT	Carender et al., 2019					0.8			0.74	0.47	0.68	0.72	0.8	0.8	0.72	0.38	0.22	0.72							
	Scott et al., 2018 (baseline, 6 weeks, 6 months, 2 years)				0.16 0.34 0.32 0.43	0.75 0.8 0.77 0.72					0.52 0.58 0.58 0.61	0.69 0.32 0.57 0.67	0.7 0.63 0.64 0.57				0.01 0 0.24 0.42	0.71 0.68 0.6 0.6							
	Trasoliniet al., 2019		0.5	0.49																	-0.63	-0.48			
	Hancock et al., 2017				0.27	0.82	0.6		0.6	0.57	0.63	0.76	0.6					0.05	0.62						
	Hancock et al., 2019				0.12	0.82	0.51		0.58	0.54	0.49	0.7	0.74					0.08	0.7						
	Kenney et al., 2019	0.76																							
	Papuga et al., 2014	0.89																							

	Kohring et al., 2018 (press ganey)																			-0.18 to 0.2
	Kortlever et al., 2019						0.74													
	Miles et al., 2019	0.7																		
	Robins et al., 2017	0.75																		0.60

Table 2: PROMIS Pain Interference and Mobility domains correlation with outcome measures

CAT	Domain	Author & year	Patient-reported outcome measures																	
			IKDC	MLQOL-EI	MLQOL-SI	MLQOL-AL	MLQOL-PI	IPAQ	Tegner	Lysholm	MARS	Knee Pain	Body Pain							
PROMIS	Pain Interference	Trasolini et al., 2019		0.61	0.55	0.71	0.47													
		Kenney et al., 2019	-0.67																	
		Nadarajah et al., 2019	-0.73						-0.33	-0.5			-0.31	0.58	0.37					
	Mobility	Trasolini et al., 2019				-0.62	-0.45			0.42	0.5									

Table 3: FOTO and OA-CAT correlations with patient-reported outcome measures

CAT	Domain	Author & year	Outcome measure	
			GROC	HAQ-II
FOTO	Knee FS	Schmitt et al., 2014	Intake: -0.05-0.13 Discharge: 0.41-0.51	
OA-CAT	Disability	Jette et al., 2009-1		0.71-0.74

Table 4: PROMIS PF-CAT correlations with OA-CAT and performance-based outcome measures

CAT	Author & year	Outcome measure		
		OA-CAT	FTSST	TUG
PROMIS PF-CAT	Chang et al., 2018	0.65-0.79	-.21 to -0.34	
	Givens et al., 2018			-0.43

Bibliography

1. Bernholt D, Wright RW, Matava MJ, Brophy RH, Bogunovic L, Smith MV. Patient reported outcomes measurement information system scores are responsive to early changes in patient outcomes following arthroscopic partial meniscectomy. *Arthroscopy* 2018;34(4):1113-1117. doi:10.1016/j.arthro.2017.10.047.
2. Hancock KJ, Glass N, Anthony CA, et al. Performance of PROMIS for healthy patients undergoing meniscal surgery. *J. Bone Joint Surg. Am.* 2017;99(11):954-958. doi:10.2106/JBJS.16.00848.
3. Scott EJ, Westermann R, Glass NA, Hettrich C, Wolf BR, Bollier MJ. Performance of the PROMIS in patients after anterior cruciate ligament reconstruction. *Orthop. J. Sports Med.* 2018;6(5):2325967118774509. doi:10.1177/2325967118774509.
4. Chen RE, Papuga MO, Voloshin I, et al. Preoperative PROMIS scores predict postoperative outcomes after primary ACL reconstruction. *Orthop. J. Sports Med.* 2018;6(5):2325967118771286. doi:10.1177/2325967118771286.
5. Jevotovsky DS, Thirukumaran CP, Rubery PT. Creating value in spine surgery: using patient reported outcomes to compare the short-term impact of different orthopedic surgical procedures. *Spine J.* 2019;19(11):1850-1857. doi:10.1016/j.spinee.2019.05.595.
6. Papuga MO, Beck CA, Kates SL, Schwarz EM, Maloney MD. Validation of GAITRite and PROMIS as high-throughput physical function outcome measures following ACL reconstruction. *J. Orthop. Res.* 2014;32(6):793-801. doi:10.1002/jor.22591.
7. Hancock KJ, Glass N, Anthony CA, et al. PROMIS: a valid and efficient outcomes instrument for patients with ACL tears. *Knee Surg. Sports Traumatol. Arthrosc.* 2019;27(1):100-104. doi:10.1007/s00167-018-5034-z.
8. Kagan R, Anderson MB, Christensen JC, Peters CL, Gililland JM, Pelt CE. The Recovery Curve for the Patient-Reported Outcomes Measurement Information System Patient-Reported Physical Function and Pain Interference Computerized Adaptive

- Tests After Primary Total Knee Arthroplasty. *J. Arthroplasty* 2018;33(8):2471-2474. doi:10.1016/j.arth.2018.03.020.
9. Edelstein AI, Bhatt S, Wright-Chisem J, Sullivan R, Beal M, Manning DW. The Effect of Implant Design on Sagittal Plane Stability: A Randomized Trial of Medial- versus Posterior-Stabilized Total Knee Arthroplasty. *J. Knee Surg.* 2019. doi:10.1055/s-0039-1678524.
 10. Christensen JC, Brothers J, Stoddard GJ, et al. Higher Frequency of Reoperation With a New Bicruciate-retaining Total Knee Arthroplasty. *Clin. Orthop. Relat. Res.* 2017;475(1):62-69. doi:10.1007/s11999-016-4812-5.
 11. Pelt CE, Sandifer PA, Gililland JM, Anderson MB, Peters CL. Mean Three-Year Survivorship of a New Bicruciate-Retaining Total Knee Arthroplasty: Are Revisions Still Higher Than Expected? *J. Arthroplasty* 2019;34(9):1957-1962. doi:10.1016/j.arth.2019.04.030.
 12. Kohring JM, Erickson JA, Anderson MB, Gililland JM, Peters CL, Pelt CE. Treated versus untreated depression in total joint arthroplasty impacts outcomes. *J. Arthroplasty* 2018;33(7S):S81-S85. doi:10.1016/j.arth.2018.01.065.
 13. Kohring JM, Pelt CE, Anderson MB, Peters CL, Gililland JM. Press Ganey Outpatient Medical Practice Survey Scores Do Not Correlate With Patient-Reported Outcomes After Primary Joint Arthroplasty. *J. Arthroplasty* 2018;33(8):2417-2422. doi:10.1016/j.arth.2018.03.044.
 14. McDonough CM, Stoiber E, Tomek IM, et al. Sensitivity to change of a computer adaptive testing instrument for outcome measurement after hip and knee arthroplasty and periacetabular osteotomy. *J. Orthop. Sports Phys. Ther.* 2016;46(9):756-767. doi:10.2519/jospt.2016.6442.
 15. Givens DL, Eskildsen S, Taylor KE, Faldowski RA, Del Gaizo DJ. Timed Up and Go test is predictive of Patient-Reported Outcomes Measurement Information System physical function in patients awaiting total knee arthroplasty. *Arthroplasty Today* 2018;4(4):505-509. doi:10.1016/j.artd.2018.07.010.
 16. Chang F-H, Jette AM, Slavin MD, Baker K, Ni P, Keysor JJ. Detecting functional change in response to exercise in knee osteoarthritis: a comparison of two computerized adaptive tests. *BMC Musculoskelet. Disord.* 2018;19(1):29. doi:10.1186/s12891-018-1942-9.
 17. Jette AM, McDonough CM, Haley SM, et al. A computer-adaptive disability instrument for lower extremity osteoarthritis research demonstrated promising breadth, precision, and reliability. *J. Clin. Epidemiol.* 2009;62(8):807-815. doi:10.1016/j.jclinepi.2008.10.004.
 18. Jette AM, McDonough CM, Ni P, et al. A functional difficulty and functional pain instrument for hip and knee osteoarthritis. *Arthritis Res. Ther.* 2009;11(4):R107. doi:10.1186/ar2760.
 19. Kenney RJ, Houck J, Giordano BD, Baumhauer JF, Herbert M, Maloney MD. Do Patient Reported Outcome Measurement Information System (PROMIS) Scales Demonstrate Responsiveness as Well as Disease-Specific Scales in Patients Undergoing Knee Arthroscopy? *Am. J. Sports Med.* 2019;47(6):1396-1403. doi:10.1177/0363546519832546.
 20. Wojahn RD, Bogunovic L, Brophy RH, et al. Opioid consumption after knee arthroscopy. *J. Bone Joint Surg. Am.* 2018;100(19):1629-1636. doi:10.2106/JBJS.18.00049.

21. Miles M, Nadarajah V, Jauregui JJ, et al. Evaluation of the PROMIS physical function computer adaptive test in patients undergoing knee surgery. *J. Knee Surg.* 2019. doi:10.1055/s-0039-1688691.
22. Nadarajah V, Glazier E, Miller K, et al. Evaluation of preoperative pain using PROMIS pain interference in knee surgery patients. *J. Knee Surg.* 2019. doi:10.1055/s-0039-1688769.
23. Meredith SJ, Nadarajah V, Jauregui JJ, et al. Preoperative opioid use in knee surgery patients. *J. Knee Surg.* 2019;32(7):630-636. doi:10.1055/s-0038-1666868.
24. Stevens KN, Nadarajah V, Jauregui JJ, et al. Preoperative expectations of patients undergoing knee surgery. *J. Knee Surg.* 2019. doi:10.1055/s-0039-1698805.
25. Carender CN, Bollier MJ, Wolf BR, Duchman KR, An Q, Westermann RW. Preoperative performance of PROMIS in patients with patellofemoral malalignment and chondral disease. *Orthop. J. Sports Med.* 2019;7(7):2325967119855001. doi:10.1177/2325967119855001.
26. Cavallero M, Rosales R, Caballero J, Virkus WW, Kempton LB, Gaski GE. Locking plate fixation in a series of bicondylar tibial plateau fractures raises treatment costs without clinical benefit. *J Orthop Trauma* 2018;32(7):333-337. doi:10.1097/BOT.0000000000001188.
27. Virkus WW, Caballero J, Kempton LB, Cavallero M, Rosales R, Gaski GE. Costs and Complications of Single-Stage Fixation Versus 2-Stage Treatment of Select Bicondylar Tibial Plateau Fractures. *J Orthop Trauma* 2018;32(7):327-332. doi:10.1097/BOT.0000000000001167.
28. Seifert C, Vokes J, Roberts A, Gorczyca J, Judd K. Simultaneous bilateral extensor mechanism disruptions: more than double the trouble? *J. Knee Surg.* 2019. doi:10.1055/s-0039-1688779.
29. Trasolini NA, Korber S, Gipsman A, San AE, Weber AE, Hatch GFR. Performance of PROMIS Computer Adaptive Testing As Compared With Established Instruments for Multiple-Ligament Knee Injuries. *Orthop. J. Sports Med.* 2019;7(9):2325967119867419. doi:10.1177/2325967119867419.
30. Hart DL, Mioduski JE, Stratford PW. Simulated computerized adaptive tests for measuring functional status were efficient with good discriminant validity in patients with hip, knee, or foot/ankle impairments. *J. Clin. Epidemiol.* 2005;58(6):629-638. doi:10.1016/j.jclinepi.2004.12.004.
31. Hart DL, Wang Y-C, Stratford PW, Mioduski JE. Computerized adaptive test for patients with knee impairments produced valid and responsive measures of function. *J. Clin. Epidemiol.* 2008;61(11):1113-1124. doi:10.1016/j.jclinepi.2008.01.005.
32. Wang Y-C, Hart DL, Stratford PW, Mioduski JE. Clinical interpretation of computerized adaptive test-generated outcome measures in patients with knee impairments. *Arch. Phys. Med. Rehabil.* 2009;90(8):1340-1348. doi:10.1016/j.apmr.2009.02.008.
33. Hart DL, Deutscher D, Crane PK, Wang Y-C. Differential item functioning was negligible in an adaptive test of functional status for patients with knee impairments who spoke English or Hebrew. *Qual. Life Res.* 2009;18(8):1067-1083. doi:10.1007/s11136-009-9517-8.
34. Wang Y-C, Hart DL, Stratford PW, Mioduski JE. Baseline dependency of minimal clinically important improvement. *Phys.*

- Ther.* 2011;91(5):675-688. doi:10.2522/ptj.20100229.
35. Deutscher D, Hart DL, Crane PK, Dickstein R. Cross-cultural differences in knee functional status outcomes in a polyglot society represented true disparities not biased by differential item functioning. *Phys. Ther.* 2010;90(12):1730-1742. doi:10.2522/ptj.20100107.
 36. Schmitt JS, Abbott JH. Patient global ratings of change did not adequately reflect change over time: a clinical cohort study. *Phys. Ther.* 2014;94(4):534-542. doi:10.2522/ptj.20130162.
 37. Robins RJ, Anderson MB, Zhang Y, Presson AP, Burks RT, Greis PE. Convergent Validity of the Patient-Reported Outcomes Measurement Information System's Physical Function Computerized Adaptive Test for the Knee and Shoulder Injury Sports Medicine Patient Population. *Arthroscopy* 2017;33(3):608-616. doi:10.1016/j.arthro.2016.08.032.
 38. Lizzio VA, Blanchett J, Borowsky P, et al. Feasibility of PROMIS CAT Administration in the Ambulatory Sports Medicine Clinic With Respect to Cost and Patient Compliance: A Single-Surgeon Experience. *Orthop. J. Sports Med.* 2019;7(1):2325967118821875. doi:10.1177/2325967118821875.
 39. Kortlever JTP, Leyton-Mange A, Keulen MHF, et al. PROMIS Physical Function Correlates with KOOS, JR in Patients with Knee Pain. *J. Knee Surg.* 2019. doi:10.1055/s-0039-1688780.
 40. Hung M, Bounsanga J, Voss MW, Saltzman CL. Establishing minimum clinically important difference values for the Patient-Reported Outcomes Measurement Information System Physical Function, hip disability and osteoarthritis outcome score for joint reconstruction, and knee injury and osteoarthritis outcome score for joint reconstruction in orthopaedics. *World J Orthop* 2018;9(3):41-49. doi:10.5312/wjo.v9.i3.41.
 41. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J. Clin. Epidemiol.* 2007;60(1):34-42. doi:10.1016/j.jclinepi.2006.03.012.