Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

CrossMark

CLINICAL RESEARCH

What is the Responsiveness and Respondent Burden of the New Knee Society Score?

Rajesh N. Maniar MS Orth, MCh Orth, Parul R. Maniar MS, FRCO,
Debashish Chanda MS, Ortho, Dnyaneshwar Gajbhare MBBS, MD,
Toral Chouhan MBA

Abstract

Background Although the new Knee Society score (NKSS) has been validated by a task force, a longitudinal study of the same cohort of patients to evaluate the score's responsiveness and respondent burden has not been reported, to our knowledge.

Questions/Purposes We analyzed the NKSS for (1) responsiveness; (2) respondent burden; and (3) convergent validity in 148 patients studied longitudinally during more than 1 year.

Methods During an 8-month period, 165 patients underwent TKA by the same surgeon at our institution, of whom 148 (90%) completed this study; the others were excluded because of distance to travel or loss to followup at the specified time. The NKSS, WOMAC, and SF-12 were completed by each patient 1 day before surgery and at 3 and 12 months postoperatively. At the same times, the original KSS (OKSS) which is designed as an observer's assessment, was completed by the same orthopaedic fellow for all patients. Responsiveness of the NKSS was assessed by determining effect size, standardized response mean (SRM), and ceiling and floor effects. Respondent burden was assessed through time to completion recorded in minutes and ease of completion which was measured objectively on a Likert scale of 1 to 5 by the patients. Convergent validity was assessed by correlating the NKSS with the WOMAC, SF-12, and OKSS (current, widely used scales) by Pearson's correlation coefficient.

Results Effect size was largest (2.83 and 3.38) and SRM was highest (2.29 and 2.68) for the NKSS at 3 and 12 months respectively, indicating the NKSS to be the most-responsive score followed by the OKSS, WOMAC, and

R. N. Maniar (✉)
Department of Orthopaedic Surgery, Lilavati Hospital & Research Centre, A-791, Bandra Reclamation, Bandra (W), Mumbai 400050, India
e-mail: drmaniar@jointspeciality.com

P. R. Maniar
Department of Refractive Surgery, New Vision Laser Centres, Mumbai, India

D. Chanda
Department of Orthopaedic Surgery, W Pratiksha Hospital, Gurgaon, India

D. Gajbhare
Department of Community Medicine, T N Medical College, Mumbai, India

T. Chouhan
Research Unit, Nook Clinic, Mumbai, India

SF-12. The NKSS exhibited no ceiling and floor effects. The NKSS took a longer time to complete (5.49 ± 3.56 minutes) compared with the WOMAC (4.64 ± 3.19 minutes) and SF-12 (4.35 ± 3.27 minutes). The mean difference in time taken for the NKSS versus the WOMAC was 0.85 minutes (95% CI, 0.54–1.17 minutes; p < 0.001) and the mean difference for the NKSS versus the SF-12 was 1.14 minutes (95% CI, 0.76–1.15 minutes; p < 0.001). Its ease of completion generally was comparable to that of the WOMAC and SF-12. Convergent validity showed a strong correlation (r > 0.6; p < 0.001) of the NKSS with the WOMAC at all times and moderate to strong correlation (r = 0.4–0.6; p < 0.001) with the SF-12 and OKSS at the first two assessments, which became strong (r > 0.6; p < 0.001) at 12 months.

Conclusions The NKSS exhibited greater responsiveness than the WOMAC, SF-12, and OKSS scales and showed no ceiling effect, indicating adequate potential for recording future improvement. The NKSS also showed reliable convergent validity when correlated with these other scores. However, it posed a greater respondent burden in terms of time to completion.

Clinical Relevance As independent nondevelopers of the NKSS, we found it to be a responsive tool for assessment of TKA outcomes. We have confirmed that the NKSS can be used interchangeably for this purpose with the WOMAC scale and that it correlates positively with other established scales of the SF-12 and OKSS. Further study of the short-form version will establish whether it also can be used effectively while reducing the respondent burden.

Introduction

The outcome of TKA is evaluated by numerous scoring systems, many of which have evolved with time [4, 8, 9, 14–16]. With increasing demands of younger, more-active patients, the fulfillment of patient expectations is increasingly important for patient satisfaction. Thus, in addition to objective clinical assessment, the patient's own evaluation of function and satisfaction through patient-reported outcome measures is now recognized to be a fundamental component of many scoring systems.

The new Knee Society Score (NKSS) was introduced in 2012 and combines an objective, clinician-derived component and a patient-reported outcome component for complete assessment of functional activities of diverse lifestyles and patient satisfaction [15]. The usefulness of any scoring system typically is determined by its validity (ability to measure what it is designed to measure) and its responsiveness (ability to detect changes that may occur during a period of time). The NKSS was developed and

validated by a Knee Society task force. Noble et al. [15] developed a comprehensive survey of activities, based on which a prototype knee scoring system was administered to 497 patients. Objective and subjective data were analyzed and compared with the Knee Injury and Osteoarthritis Score and SF-12 scores for validation. Statistical analysis confirmed the internal consistency, construct and convergent validity, and reliability of the new score. Based on this analysis, minor modifications led to the NKSS. Their study population was culturally homogenous, which was a limitation of their work and they suggested studying other populations outside the United States and Canada [15]. Further research to determine the responsiveness of the NKSS in measuring changes in response to TKA, and longitudinal followup of the same cohort of patients for evaluation of responsiveness were suggested [15].

We therefore followed a large cohort of patients undergoing TKA, longitudinally, during more than 1 year to assess (1) responsiveness, in terms of effect size, standardized response mean (SRM), and ceiling and floor effect; (2) respondent burden in terms of time to completion and ease of completion; and (3) convergent validity by correlating the NKSS with the current, established WOMAC, SF-12, and the original KSS (OKSS) scores [1, 2, 8, 20].

Patients and Methods

One hundred sixty-five patients with the diagnosis of degenerative osteoarthritis of the knee who had been scheduled to undergo primary TKA by the senior author (RNM) between September 1, 2014, and April 30, 2015, were considered for our prospective study. Seven patients from remote regions who were unable to followup per protocol were excluded from the study. At 3 months ± 15 days, six patients were lost to followup. At 12 months ± 1 month, another four patients who did not return for followup during the specified period also were excluded. The final cohort of 148 patients (90%) was analyzed (Fig. 1).

The age of this final cohort ranged from 49 to 86 years (average, 68 years); 32 were men and 116 were women. Their BMI ranged from 19 to 49 kg/m^2 (average, 30 kg/m^2). All patients underwent surgery by the senior author (RNM), using computer navigation for alignment and balance (Kolibri™ navigation system; BRAINLAB AG, Munich, Germany) through a midline skin incision under tourniquet control. The posterior-stabilized PFC® Sigma® or Attune® implant (DePuy Orthopaedics Inc, Warsaw, IN, USA) was used with cemented fixation for all knees. Periarticular injection analgesia and patient-controlled analgesia pumps were used in all patients. A standard
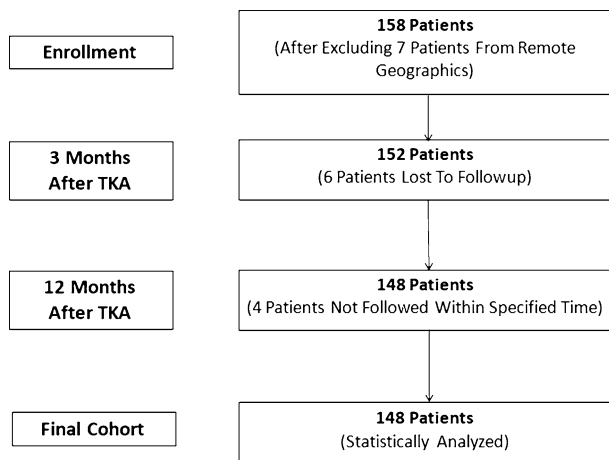
Fig. 1 The study flowchart is shown.

rehabilitation protocol was used; patients were mobilized on the evening of the day of surgery and typically discharged from the hospital on postoperative Day 4.

All 148 patients were administered the NKSS, WOMAC, and SF-12 scoring forms preoperatively after admission (ie, 1 day before surgery) and postoperatively at 3 months ± 15 days and 12 months ± 1 month. At the same times, the OKSS forms were completed for each patient by the same joint replacement fellow (DC) for all patients. The forms were evaluated for responsiveness, respondent burden, and convergent validity as described below.

Score Elements

The NKSS score was obtained from a self-administered questionnaire, which combines general demographic information, an observer-assessed objective knee score (NKSS-OKS) ranging from 0 to 100 points, and a patient-reported subjective knee score (NKSS-SKS) ranging from 0 to 155 points, giving a total maximum score of 255 points. The NKSS-OKS includes a pain component of 25 points. The NKSS-SKS includes a function component (FS) ranging from 0 to 100 points and a patient expectation and satisfaction component ranging from 0 to 55 points. The expectation component of the expectation and satisfaction subscale ranges from 0 to 15 points and the satisfaction component from 0 to 40 points. A higher score represents better knee status. The total score of the NKSS and the subscores (NKSS-SKS and NKSS-OKS) were evaluated for responsiveness. The respondent burden was measured in whole for the total scores. For convergent validity, the total score of the NKSS and its function, expectation and satisfaction, and pain components were correlated with the corresponding components of the other scores.

The WOMAC score [1, 2] (a disease-specific measure of pain, stiffness, and function) was obtained from a self-administered questionnaire, the maximum score being 96. The WOMAC includes subscores for pain (range, 0–20), stiffness (range, 0–8), and function (range, 0–68). A higher WOMAC score represents poorer status and a lower score represents better status.

The SF-12 health survey [20] (a generic health status measure) is also a self-administered questionnaire. It combines a physical component score and a mental component score, each ranging from 0 to 100, giving a maximum score of 200 points. A higher score indicates better health status.

The OKSS [8] consists of two scores, a knee score (OKSS-KS) and a function score (OKSS-FS), each ranging from 0 to 100 points, yielding a maximum score of 200 points. A higher score indicates better knee status.

The responsiveness was assessed by estimation of the effect size and SRM and ceiling and floor effects. Effect size is defined as the mean score change divided by the SD of the preoperative score [10]. Effect size greater than 0.8 is considered large [3], and the higher the value, the better the responsiveness. SRM is defined as the mean score change divided by the SD of the change in score [12]. A SRM greater than 1 is considered large, and the higher the value, the better the responsiveness. Effect sizes and SRMs of all scores and the NKSS-OKS and NKSS-SKS were calculated. Ceiling and floor effects are determined by the proportion of patients at the ceiling (ie, the best score) and at the floor (ie, the worst score) of each scale. We also calculated the proportion of patients nearing the ceiling or floor scores (ie, within 10% of the best and worst scores).

Respondent burden was assessed in terms of time to completion and ease of completion. Time to completion was recorded in minutes, with seconds rounded up to the nearest minute. The means ± SD of time taken were calculated. Paired comparisons of the NKSS versus the WOMAC and NKSS versus the SF-12 were done by determining the mean differences. Ease of completion was ranked by the patients on a Likert scale, 1 being easy and 5 being most difficult.

Convergent validity was assessed by Pearson's correlation coefficient test. Correlation was assessed between the total scores and between the corresponding pain, function, and satisfaction components. If the correlation was found to be statistically significant ($p < 0.05$), the strength of the correlation was determined by the correlation coefficient (r) value. An r less than 0.2 is considered a clinically irrelevant correlation, 0.2 to 0.4 is considered a weak correlation; 0.4 to 0.6 a moderate correlation; and greater than 0.6 is considered strong correlation [18]. We also correlated the change of each score from the preoperative value to the 3- and 12-month postoperative values by Pearson's correlation test and determined the correlation coefficient r. Scatterplots were charted to look at the

correlation between the NKSS with all other scores preoperatively and at 3 and 12 months postoperatively.

Descriptive statistics of the raw data for all 148 patients were tabulated during the period up to 12 months for the NKSS, WOMAC, SF-12and OKSS (Table 1).

For statistical analyses, we used MedCalc version 16 (MedCalc Software bvba, Ostend, Belgium) to determine the effect size and standardized response mean. Other analyses were done with the PSPP Release 0.10.2 (Free Software Foundation Inc, Boston, MA, USA). Paired comparisons were done using the chi-square test. A probability less than 0.05 was considered significant. Correlation between various scales was done using Pearson's coefficient correlation.

Results

Responsiveness

The NKSS was responsive; that is, it showed good ability to detect changes in the patient's status with time, and the potential to detect further changes beyond the study period. All scores showed good responsiveness at 3 months and 12

months (effect size > 0.8; SRM > 1) except for the SF-12. The NKSS was most responsive: at 3 months its effect size was 2.83 (95% CI, 2.38–3.27) and its SRM was 2.29 (95% CI, 1.93–2.62), and at 12 months its effect size increased to 3.38 (95% CI, 2.86–3.88) and its SRM increased to 2.68 (95% CI, 2.25–3.11). The NKSS was followed in responsiveness by the OKSS, WOMAC, and SF-12 in descending order, but none of their effect sizes or SRMs were greater than 2.0 (Table 2). Of the subscores, the NKSS-OKS also showed good responsiveness at 3 months with an effect size of 3.98 (95% CI, 3.07–4.78) and SRM of 3.65 (95% CI, 2.76–4.40), which increased further at 12 months. The NKSS-SKS at 3 months showed a comparatively lesser effect size of 1.32 (95% CI, 1.05–1.60) and SRM of 0.96 (95% CI, 0.76–1.14), which increased further at 12 months (Table 2). The NKSS showed no ceiling effect preoperatively, at 3 months or at 12 months (Table 3). No patient reached the maximum NKSS score at any time, but 3% and 8% of patients reached 100% of the OKSS score at 3 and 12 months, respectively, as did 0.7% and 7% of patients with the WOMAC score at those same times. Furthermore, the NKSS showed only two patients (1.4%) at 3 months and 9 patients (6%) at 12 months reaching 90% or more of the maximum NKSS value, whereas the OKSS and

Table 1. Raw data statistics

Scores	Time	Number	Mean	SD	Median	Q1	Q3	IQR	Range	
									Minimum	Maximum
WOMAC	Preoperative	148	40.80	17.32	41.00	26.00	53.00	27.00	11.00	88.00
	3 months	148	20.76	12.40	20.00	11.00	31.00	20.00	0.00	55.00
	12 months	148	16.11	13.13	14.00	6.00	24.00	18.00	0.00	81.00
OKSS	Preoperative	148	123.76	25.53	130.00	110.25	140.00	29.75	9.00	160.00
	3 months	148	166.14	15.20	166.50	159.00	180.00	21.00	125.00	200.00
	12 months	148	171.76	18.75	170.00	160.00	180.00	20.00	90.00	200.00
NKSS	Preoperative	148	103.68	24.61	103.50	89.00	120.00	31.00	39.00	189.00
	3 months	148	173.41	21.32	173.00	159.00	186.00	27.00	110.00	239.00
	12 months	148	186.99	23.50	189.00	171.25	200.00	28.75	105.00	252.00
SF-12	Preoperative	148	85.73	9.63	80.13	80.13	92.50	12.38	60.10	108.80
	3 months	148	91.09	12.03	84.55	84.55	99.70	15.15	57.00	117.40
	12 months	148	96.49	12.41	87.83	87.83	106.00	18.18	56.60	117.40
WOMAC	Change at 3 months	148	−20.03	20.13	−18.50	−34.75	−6.00	28.75	−81.00	34.00
OKSS		148	42.39	25.97	40.00	25.25	54.75	29.50	−12.00	151.00
NKSS		148	69.74	30.42	71.00	54.25	86.75	32.50	−20.00	162.00
SF-12		148	5.36	13.14	−2.18	−2.18	13.83	16.00	−24.90	38.20
WOMAC	Change at 12 months	148	−24.69	21.31	−24.50	−39.00	−9.00	30.00	−75.00	45.00
OKSS		148	48.01	25.16	45.00	30.00	64.00	34.00	−20.00	132.00
NKSS		148	83.20	31.19	84.00	63.25	105.75	42.50	−23.00	187.00
SF-12		148	10.76	14.01	0.90	0.90	21.10	20.20	−34.30	43.90

Q1 = first quartile; Q3 = third quartile; IQR = interquartile range; OKSS = original Knee Society Score; NKSS = New Knee Society Score.

Table 2. Effect size and standardized response mean

Responsiveness	NKSS (CI)	WOMAC (CI)	SF-12 (CI)	OKSS (CI)	NKSS OKS (CI)	NKSS SKS (CI)
Effect size						
At 3 months	2.83 (2.38–3.27)	−1.16 (−1.36 to −0.97)	0.56 (0.35–0.79)	1.66 (1.41–1.88)	3.98 (3.07–4.78)	1.32 (1.05–1.60)
At 12 months	3.38 (2.86–3.88)	−1.43 (−1.64 to −1.23)	1.12 (0.84–1.40)	1.88 (1.54–2.14)	4.26 (3.35–5.13)	1.95 (1.67–2.29)
Standardized response mean						
At 3 months	2.29 (1.93–2.62)	−1.0 (−1.19 to −0.81)	0.41 (0.25–0.59)	1.63 (1.41–1.86)	3.65 (2.76–4.40)	0.96 (0.76–1.14)
At 12 months	2.68 (2.25–3.11)	−1.16 (−1.36 to −0.95)	0.77 (0.57–0.97)	1.91 (1.66–2.17)	3.98 (2.97–4.92)	1.34 (1.12–1.55)

OKSS = original Knee Society Score; NKSS = New Knee Society Score; SKS = subjective knee score; OKS = objective knee score.

Table 3. Ceiling and floor effect

Ceiling effect	Scores at 100% of maximum score			Scores at 90% of maximum and above		
Time	NKSS	WOMAC	OKSS	NKSS	WOMAC	OKSS
Preoperatively	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
3 months	0 (0%)	1 (0.68%)	4 (2.70%)	2 (1.35%)	30 (20.27%)	40 (27.03%)
12 months	0 (0%)	10 (6.76%)	12 (8.11%)	9 (6.08%)	47 (31.76%)	63 (42.57%)
Floor effect	Scores at 0% (floor)			Scores within 10% of floor		
Time	NKSS	WOMAC	OKSS	NKSS	WOMAC	OKSS
Preoperatively	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.70%)
3 months	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
12 months	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

NKSS = New Knee Society Score; OKSS = Original Knee Society Score.

WOMAC had at least 20% of patients at 3 months and at least 31% of patients at 12 months reaching more than 90% of the score's maximum value. No patient with any of the three scales (NKSS, OKSS, and WOMAC) had the least-possible score preoperatively or at 3 and 12 months, indicating no floor effect (Table 3). Further observation of the percentage of patients in the lowest 10% of the scores' values also showed no patients with the NKSS, one patient (0.7%) with the OKSS, and no patients with the WOMAC in this range.

Respondent Burden

The NKSS took longer for patients to complete than did the other outcomes tools. The mean time to completion for the NKSS was 5.49 ± 3.56 minutes, followed by the WOMAC which was 4.64 ± 3.19 minutes, followed by the SF-12 which was 4.35 ± 3.27 minutes. The mean difference in time taken for the NKSS versus the WOMAC was 0.85 minutes (95% CI, 0.54–1.17 minutes; $p < 0.001$) and the mean difference for the NKSS versus the SF-12 was 1.14

minutes (95% CI, 0.76–1.15 minutes; $p < 0.001$) (Table 4). In terms of respondent burden as measured by ease of completion on a Likert scale, with the numbers available there were no differences among the three scores. Mean difference for the NKSS versus WOMAC was 0.14 (95% CI, −0.01 to −0.28; $p = 0.061$) and for the NKSS versus SF-12 was 0.05 (95% CI, −0.11 to −0.21; $p = 0.565$) (Table 4).

Convergent Validity

We observed strong convergent validity of the NKSS with the WOMAC at all assessment points and moderate convergent validity with the SF-12 and OKSS at the first two assessment points, which became strong at 12 months. The NKSS and WOMAC scores therefore can be used interchangeably. Specifically, the NKSS correlated preoperatively and postoperatively at 3 and 12 months with the WOMAC, SF-12, and OKSS scores. The correlation was strong ($r > 0.6$: $p < 0.001$) with the WOMAC at all times. The strength of the correlation with the SF-12 and

Table 4. Time and ease of completion

Score n = 148	Time to completion			Ease of completion		
	Mean ± SD (95% CI)	Q1	Q3	Mean ± SD	Q1	Q3
NKSS	5.49 ± 3.56 (4.91–6.07)	3.00	5.00	1.84 ± 1.00	1.00	2.00
WOMAC	4.64 ± 3.19 (4.12–5.16)	3.00	5.00	1.70 ± 0.91	1.00	2.00
SF-12	4.35 ± 3.27 (3.83–4.89)	2.00	5.00	1.79 ± 0.97	1.00	2.00
Pairwise comparison	Mean difference (95% CI)		p value	Mean difference (95% CI)		p value
NKSS vs WOMAC	0.85 (0.54–1.17)		< 0.001	0.14 (−0.01 to 0.28)		0.061
NKSS vs SF-12	1.14 (0.76–1.51)		< 0.001	0.05 (−0.11 to 0.21)		0.565

Q1 = first quartile; Q3= third quartile; ease of completion measured on Likert scale (1–5, 1 = easy, 5 = most difficult).\; NKSS = new Knee Society Score.

Table 5. Correlation of scores

Correlation	Preoperative		3 months		12 months	
	r (CI)	p value	r (CI)	p value	r (CI)	p value
NKSS vs WOMAC	−0.61 (−0.70 to −0.49)	< 0.001	−0.70 (−0.77 to −0.60)	< 0.001	−068 (−0.76 to −0.58)	< 0.001
NKSS vs SF-12	0.52 (0.39–0.63)	< 0.001	0.56 (0.43–0.66)	< 0.001	0.67 (0.57–0.75)	< 0.001
NKSS vs OKSS	0.64 (0.54–0.73)	< 0.001	0.48 (0.34–0.59)	< 0.001	0.62 (0.51–0.71)	< 0.001
Pain subscores						
NKSS vs WOMAC	−0.61 (−0.70 to −0.50)	< 0.001	−0.60 (−0.70 to −0.49)	< 0.001	−0.57 (−0.67 to −0.45)	< 0.001
NKSS vs OKSS	0.37 (0.22–0.50)	< 0.001	0.39 (0.24–0.51)	< 0.001	0.57 (0.46–0.67)	< 0.001
Function subscores						
NKSS vs WOMAC	−0.55 (−0.66 to −0.43)	< 0.001	−0.43 (−0.55 to −0.29)	< 0.001	−0.54 (−0.64 to −0.41)	< 0.001
NKSS vs SF-12 PCS	0.12 (−0.04 to 0.27)	0.16	0.42 (0.28–0.54)	< 0.001	0.53 (0.40–0.63)	< 0.001
NKSS vs OKSS FS	0.50 (0.37–0.61)	< 0.001	0.41 (0.26–0.53)	< 0.001	0.59 (0.47–0.68)	< 0.001
Satisfaction subscore						
NKSS vs SF-12 MCS	0.30 (0.14–0.44)	< 0.001	0.35 (0.20–0.48)	< 0.001	0.23 (0.07–0.38)	0.005

r = correlation coefficient; OKSS = original Knee Society Score; NKSS = new Knee Society Score; FS = function score; PCS = physical component score; MCS = mental component score.

OKSS was moderate (r = 0.4–0.6; p < 0.001) at the first two assessments but became strong (r > 0.6; p < 0.001) at 12 months. Likewise, the NKSS subscores for pain and function generally correlated well with the subscores of the WOMAC, SF-12, and OKSS, the correlation being moderate (r = 0.4–0.6; p < 0.001) at each assessment except for the one observation of weak preoperative correlation (r = 0.12; p = 0.16) of the NKSS function component with the SF-12 physical component. In addition, the NKSS satisfaction component was weakly correlated with the SF-12 mental component (all r ≤ 0.3; p < 0.001) at all times (Table 5).

The correlation between the changes of scores for the NKSS with the WOMAC, SF-12, and OKSS at 3 months and 12 months (Table 6) was strong at all times (all r values > 0.6; p < 0.001, except for the correlation with the OKSS at 12 months, which was moderate (r = 0.49; p < 0.001).

The scatterplot for preoperative comparison (Fig. 2) showed a positive correlation among all scores which was maintained at 3 months (Fig. 3) and 12 months (Fig. 4) postoperatively.

Discussion

The outcome of TKA is best assessed by combining the observer's objective evaluation with the patient's evaluation of functional and satisfaction components. The NKSS, introduced in 2012, is one such tool and we have been using it in our population since April, 2014. The usefulness of any scoring system is determined by its validity and responsiveness. The validation study for the NKSS was done by the Knee Society Task Force in a culturally homogenous population at multiple sites in the United

Table 6. Correlation of change of scores

Change of score	3 months		12 months	
	r (95% CI)	p value	r (95% CI)	p value
NKSS vs WOMAC	−0.67 (−0.74 to −0.57)	< 0.001	−0.67 (−0.74 to −0.56)	< 0.001
NKSS vs SF-12	0.52 (0.39–0.62)	< 0.001	0.53 (0.41–0.64)	< 0.001
NKSS vs OKSS	0.54 (0.42–0.64)	< 0.001	0.49 (0.35–0.60)	< 0.001

NKSS = new Knee Society Score; OKSS = original Knee Society Score; r = correlation coefficient.

Fig. 2 The scatterplot shows the correlation of the new Knee Society Score (NKSS) with the original Knee Society Score (OKSS), WOMAC, and SF-12 scores preoperatively.
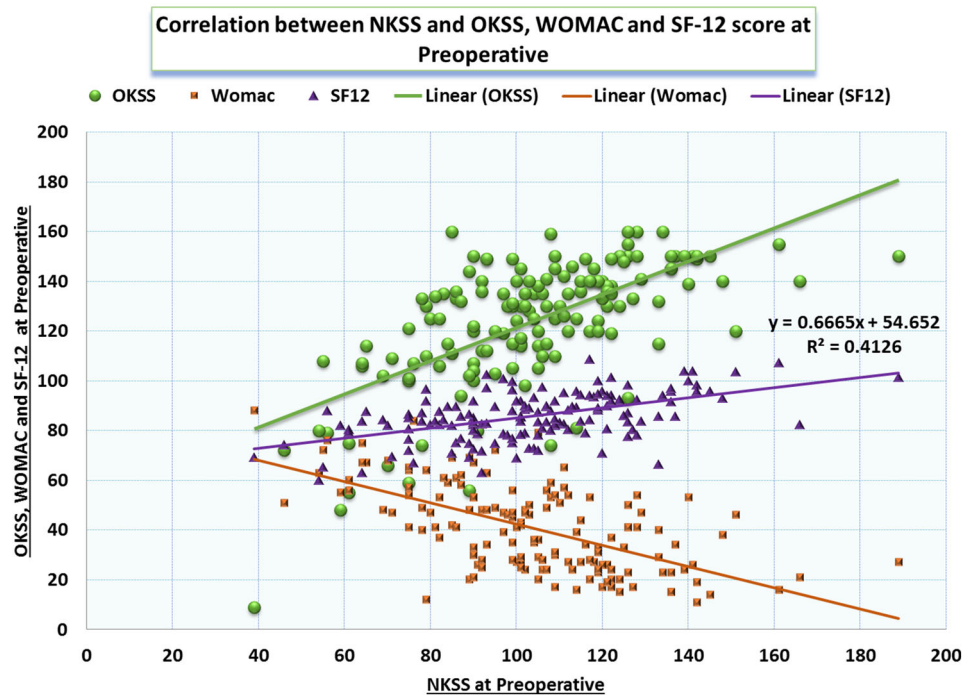
States and Canada. They suggested studying the NKSS in different populations of other countries and evaluating its responsiveness by a longitudinal followup of the same cohort of patients [15].

As independent nondevelopers, we undertook this challenge to evaluate the responsiveness and respondent burden of the NKSS in a large cohort of Indian patients followed for more than 1 year. All patients underwent surgery by the same surgeon with a posterior-stabilized implant, followed with the same postoperative protocol, and their objective assessment was done by the same orthopaedic fellow. We also correlated the NKSS with the established outcome measures of the WOMAC, SF-12, and OKSS. Ninety percent (148 patients) of our single cohort of 165 patients completed the required assessment before surgery, and at 3 months and 1 year after surgery. We

found the NKSS showed good responsiveness, increased respondent burden, and good convergent validity.

Our study had several limitations. First, the NKSS questionnaire was not adapted for our Indian population. The Dutch, French, and the Japanese have adapted and studied their respective translated version of the NKSS [5, 7, 19]. However, the Indian population that we studied was an English-speaking urban population. We also used the other scoring systems in their original form in English. However, we cannot account for any cultural differences that might be evidenced by more heterogenous populations completing the same outcome measures. Second, the study patients did not follow a fixed sequence for completing forms. The order in which the forms were completed could influence the responses. Throughout our study, the order was arbitrary and the questionnaires required each query to

Fig. 3 The scatterplot shows the correlation of the new Knee Society Score (NKSS) with the original Knee Society Score (OKSS), WOMAC, and SF-12 scores at 3 months postoperatively.
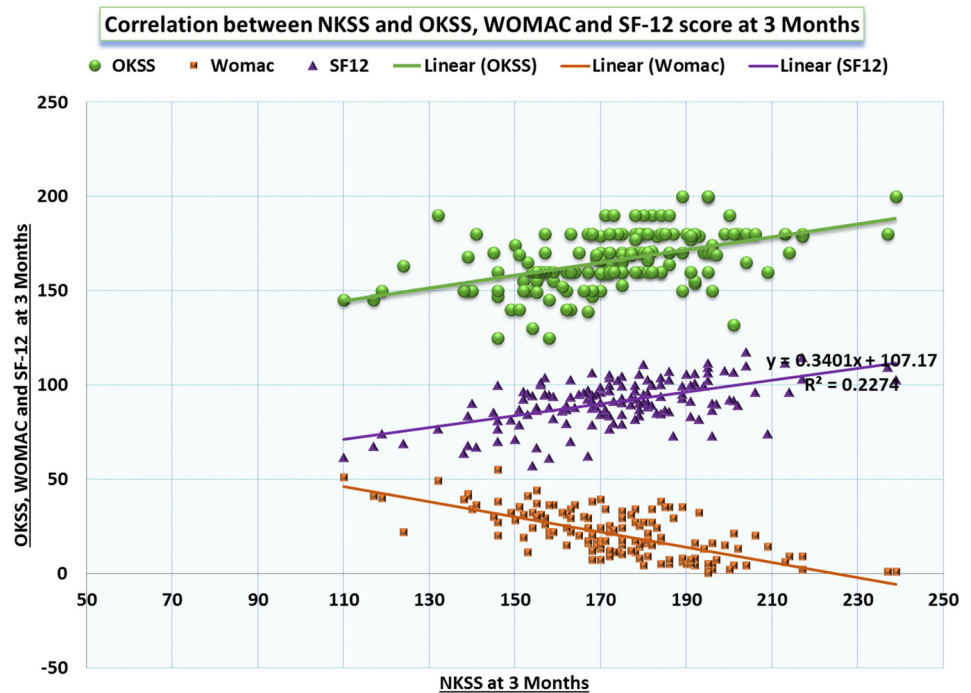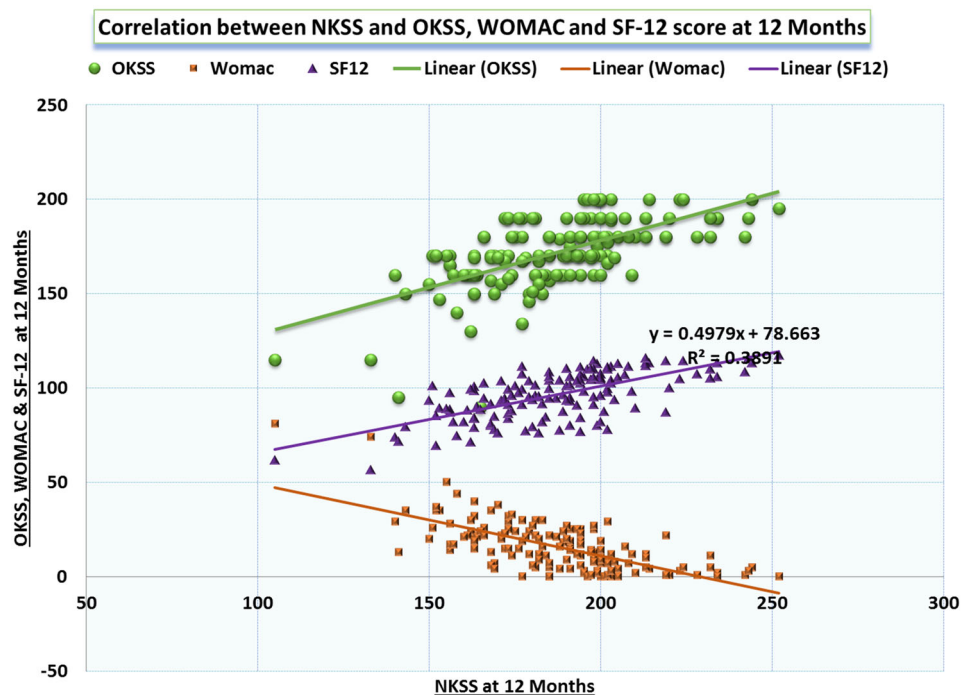

Fig. 4 The scatterplot shows correlation of the new Knee Society Score (NKSS) with the original Knee Society Score (OKSS), WOMAC, and SF-12 scores at 12 months postoperatively.

be read through before its answer could be marked, therefore we do not expect any biased responses. In addition, the observer's assessment of the OKSS was done by the same orthopaedic fellow for all patients at the same time as the patient-reported scores. We cannot be certain whether some patients may have received assistance from the fellow if they did not understand a question on one of the patient-reported outcome measures.

Responsiveness is defined as the sensitivity of an assessment technique to change with time in response to the patient's changing status. In our study, we found the NKSS to be the most responsive of the scores tested. SRM values were comparatively smaller than effect sizes, but the interpretation of results remained the same. In addition, effect sizes and SRM values were higher at 12 months compared with 3 months, indicating corresponding

improvement in the patient's condition with time. Compared with the other scores, in our study, the OKSS showed better responsiveness than the WOMAC, whereas the SF-12 showed the least responsiveness. Kreibich et al. [11] found the highest responsiveness with the WOMAC and OKSS in a comparison of six scoring systems, whereas Lingard et al. [13] reported the OKSS to be the least responsive and the WOMAC and SF-36 to be more responsive. We also measured the responsiveness of the subscores of the NKSS to ascertain the responsiveness of the subjective and objective components individually. Both subscores exhibited good responsiveness; the objective subscore (ie, the NKSS-OKS) showed greater responsiveness than the subjective subscore (ie, the NKSS-SKS). At 3 months, the NKSS-SKS was lower, but increased at 12 months, indicating increasing patient satisfaction with time. In our study, the NKSS also exhibited no ceiling effect at 3 and 12 months, indicating that it had the capacity to detect future improvement in the patient's condition. Compared with this, the WOMAC and OKSS showed a ceiling effect, which was greater at 12 months. Furthermore, there was a larger percentage of patients close to the ceiling score with the WOMAC and OKSS at 3 and 12 months. In our study, 7% of patients had reached the ceiling score with the WOMAC at 12 months and 32% patients had reached more than 90% of the maximum score. Marx et al. [14] also reported a 4% ceiling effect at 12 months, with 20% of patients achieving more than 95% of the score with the WOMAC.

For respondent burden in terms of time taken, the NKSS took the longest time to completion. The patients had to read through 44, 24, and 14 items for the NKSS, WOMAC, and SF-12 to answer 30, 24, and seven questions, respectively. Paired comparisons showed greater respondent burden of the NKSS compared with the WOMAC and SF-12. In terms of ease of completion as graded by the patients on a Likert scale, with the numbers available there were no differences among the three scores. The NKSS and SF-12 were at par with the burden slightly higher than the WOMAC. The magnitude of the difference, however, was not beyond what could be expected by chance only. Dinjens et al. [6] investigated clinimetric parameters of the patient-reported outcome measurement part of the NKSS in 415 patients undergoing primary TKA. They used a validated Dutch-translated version of the NKSS and reported a response rate of 96% and completion rate of only 43%. The low completion rate was found to be mainly attributable to missing answers in the function subscore for advanced and discretionary activities. They recommended improvements like shortening the scale and simplifying the design to increase the disappointing completion rate. A short-form version of the NKSS subsequently has been developed and

shown to be practical, valid, reliable, and responsive for assessing the functional outcome of TKA [17].

Our correlation tests established convergent validity of the NKSS strongly with the WOMAC at all assessment points and moderately with the SF-12 and OKSS at the first two assessment points, which became strong at 12 months. Van Der Straeten et al. [19] reported that the Dutch NKSS correlated well with the Dutch WOMAC (r = 0.75; p < 0.001) and with the Dutch SF-12 (r = 0.57; p < 0.001). We also correlated the subscores of the NKSS with corresponding subscores of other scales to determine how equivalent parameters correlated. We found that the corresponding pain and function components correlated well. However, the NKSS satisfaction subscore correlated with the SF-12 mental component score weakly, a finding also reported by The Knee Society task force in their initial study [15].

As independent nondevelopers, we established adequate convergent validity of the NKSS in our diverse Indian population and conclude that it is a highly responsive scale with a limited ceiling effect, allowing evaluation of recovery after TKA beyond a year. We recommend that the short-form version be similarly evaluated to ascertain whether it can be used equally effectively while reducing the respondent burden. Although our findings are not strictly generalizable beyond the patient population evaluated here, they are in conformity with other patient populations studied previously.

References

1. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of the WOMAC: a health status instrument for measuring clinically important patient-relevant outcomes following total hip or knee arthroplasty in osteoarthritis. *J Orthop Rheumatol.* 1988;1:95–108.
2. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol.* 1988;15:1833–1840.
3. Cohen J. *Statistical Power Analysis for the Behavioral Sciences.* New York, NY: Academic Press Inc; 1977.
4. Davies AP. Rating systems for total knee replacement. *Knee.* 2002;9:261–266.
5. Debette C, Parratte S, Maucort-Boulch D, Blanc G, Pauly V, Lustig S, Servien E, Neyret P, Argenson JN. French adaptation of the new Knee Society Scoring System for total knee arthroplasty. *Orthop Traumatol Surg Res.* 2014;100:531–534.

6. Dinjens RN, Senden R, Heyligers IC, Grimm B. Clinimetric quality of the new 2011 Knee Society score: high validity, low completion rate. *Knee.* 2014;21:647–654.

7. Hamamoto Y, Ito H, Furu M, Ishikawa M, Azukizawa M, Kuriyama S, Nakamura S, Matsuda S. Cross-cultural adaptation and validation of the Japanese version of the new Knee Society Scoring System for osteoarthritic knee with total knee arthroplasty. *J Orthop Sci*. 2015;20:849–853.

8. Insall JN, Dorr LD, Scott RD, Scott WN. Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989;248:13–14.

9. Insall JN, Ranawat CS, Aglietti P, Shine J. A comparison of four models of total knee-replacement prostheses. *J Bone Joint Surg Am*. 1976;58:754–765.

10. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care.* 1989;27(3 suppl):S178-189.

11. Kreibich DN, Vaz M, Bourne RB, Rorabeck CH, Kim P, Hardie R, Kramer J, Kirkley A. What is the best way of assessing outcome after total knee replacement? *Clin Orthop Relat Res.* 1996;331:221–225.

12. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care.* 1990;28:632–642.

13. Lingard EA, Katz JN, Wright RJ, Wright EA, Sledge CB; Kinemax Outcomes Group. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. *J Bone Joint Surg Am.* 2001;83:1856–1864.

14. Marx RG, Jones EC, Atwan NC, Closkey RF, Salvati EA, Sculco TP. Measuring improvement following total hip and knee arthroplasty using patient-based measures of outcome. *J Bone Joint Surg Am.* 2005;87:1999–2005.

15. Noble PC, Scuderi GR, Brekke AC, Sikorskii A, Benjamin JB, Lonner JH, Chadha P, Daylamani DA, Scott WN, Bourne RB. Development of a new Knee Society scoring system. *Clin Orthop Relat Res.* 2012;470:20–32.

16. Roos EM, Toksvig-Larsen S. Knee injury and Osteoarthritis Outcome Score (KOOS): validation and comparison to the WOMAC in total knee replacement. *Health Qual Life Outcomes.* 2003;1:17.

17. Scuderi GR, Sikorskii A, Bourne RB, Lonner JH, Benjamin JB, Noble PC. The Knee Society Short Form reduces respondent burden in the assessment of patient-reported outcomes. *Clin Orthop Relat Res.* 2016;474:134–142.

18. Swinscow TD, Campbell MJ, *Statistics at Square One.*10th ed. London, UK: BMJ Books; 2002.

19. Van Der Straeten C1, Witvrouw E, Willems T, Bellemans J, Victor J. Translation and validation of the Dutch new Knee Society Scoring System ©. *Clin Orthop Relat Res.* 2013; 471:3565–3571.

20. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220–233.