

Development and Validation of the Functional Assessment of Cancer Therapy–Kidney Symptom Index (FKSI)

David Cella, PhD, Susan Yount, PhD, Hongyan Du, MS, Rahul Dhanda, PhD, Kathleen Gondek, PhD, Katie Langefeld, Jacquelyn George, William P. Bro, Celeste Kelly, and Ronald Bukowski, MD

Renal cell carcinoma is the third most common genitourinary malignancy.¹ In the United States each year, an estimated 32,000 people are diagnosed with kidney cancer, and 12,000 succumb to it²; twice as many men than women suffer from the disease. Early diagnosis and treatment of kidney cancer can result in cure. However, early-stage disease typically is asymptomatic, so this cancer often is not detected until it has metastasized or grown sufficiently to displace other organs.

Kidney cancer may present with an array of signs and symptoms. Symptoms associated with advanced disease include flank pain, hematuria, weight loss, fever, and paraneoplastic syndromes³; patients who present with symptoms tend to have a poor prognosis.^{3,4}

Surgery is the preferred treatment for resectable renal cancer, and survival rates depend on tumor stage and grade. Single-agent and combined cytotoxic chemotherapies have elicited response rates $\leq 10\%$ and do not play a major role in the treatment of primary or metastatic renal cancer.⁵ Currently, interferon is the standard treatment for kidney cancer, although newer targeted agents (eg, sorafenib [Nexavar] and sunitinib [Sutent]) have

From the Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare, Evanston, Illinois; Northwestern University Medical School, Evanston, Illinois; Bayer Healthcare Pharmaceuticals, West Haven, Connecticut; Kidney Cancer Association, Evanston, Illinois; and the Cleveland Clinic, Cleveland, Ohio.

Manuscript submitted July 18, 2005; accepted January 6, 2006.

Supported by a research grant from Bayer Pharmaceuticals.

Correspondence to: David Cella, PhD, Center on Outcomes, Research and Education (CORE), Evanston Northwestern Healthcare, 1001 University Place, Suite 100, Evanston, IL 60201; telephone: (847) 570-7370; fax: (847) 570-8033; e-mail: d-cella@northwestern.edu

J Support Oncol 2006;4:191–199

© 2006 Elsevier Inc. All rights reserved.

Abstract The Functional Assessment of Cancer Therapy (FACT)–Kidney Symptom Index (FKSI) was developed and validated to enhance treatment decision-making, practice guidelines, symptom management, and treatment efficacy for kidney cancer patients. Thirty-four symptoms related to the disease were identified and tested. An equal weighting of patient and clinician ratings of the relative importance of each of these items led to production of a 15-item index (FKSI–15) and a 10-item abbreviated option (FKSI–10). To assess psychometric properties, patients completed the FKSI, Functional Assessment of Cancer Therapy–General (FACT–G), Eastern Cooperative Oncology Group–Performance Status Rating (ECOG–PSR), and a Global Rating of Change Scale (GRCS). Patient responses to the FKSI were analyzed for internal consistency, test-retest reliability, convergent and discriminant validity, and responsiveness to change in clinical status. The FKSI–10 showed high internal consistency; correlations between both FKSI–10 and the physical and functional well being domains of the FACT–G were high. The FKSI–10 differentiated patients grouped by ECOG–PSR (all $P < 0.001$) and discriminated patients based on their GRCS rating. The minimally important difference (MID) range estimate for the FKSI–10 was 2–4 points; the psychometric properties of the FKSI–15 were very similar (MID range, 3–5 points). Thus, the FKSI–15 and FKSI–10 are reliable and valid symptom indices for evaluating kidney cancer patients.

shown promise in improving clinical outcomes.⁶

The various sites of renal metastases and the range of treatment options available suggest that many potential symptoms may affect a patient's health-related quality of life (HRQL). For people with advanced disease, who have a reduced life expectancy and no cure presently, relief of physical symptoms and maintenance of function are the primary objectives of medical intervention.^{7–9} Thus, the ability to assess symptom benefit during clinical trial evaluation becomes important.^{10–12}

Several multidimensional HRQL measures are available to assess patient outcomes in drug trials and in clinical practice,^{13,14} but a targeted kidney cancer symptom index that could provide clinically appropriate and precise evaluation of symp-

tomatology and increased sensitivity to the impact of drug therapy is needed. Such a targeted symptom index for kidney cancer would enhance treatment decision-making, practice guidelines, symptom management, and treatment efficacy (symptom benefit response).

Toward this goal, we developed the Functional Assessment of Cancer Therapy–Kidney Symptom Index (FKSI), a brief set of items selected specifically to assess symptoms in patients with advanced kidney cancer.

Methods

STUDY OVERVIEW AND DESIGN

The FKSI was developed and validated in two phases: scale development (item generation and item reduction) and scale validation. At the start of scale development, a list of symptoms related generally to cancer and specifically to kidney cancer was generated from a review of literature, our Functional Assessment of Chronic Illness Therapy (FACIT) item bank, results from the Functional Assessment of Cancer Therapy–General (FACT–G),¹³ and consultations with kidney cancer experts. All FACIT questionnaires were previously developed using input from patients with the target condition. The new list of symptoms that potentially related to kidney cancer then was prepared and presented both to patients with the disease and to clinicians with expertise in treating kidney cancer; these participants then highlighted the symptoms most important for evaluation in kidney cancer patients. In addition, other symptoms and concerns mentioned with more than chance frequency were retained as the FKSI.

PARTICIPANTS AND MEASURES

Inclusion criteria for patients participating in the scale development and validation phases follow: 1) diagnosis of advanced or recurrent kidney cancer (defined as cancer that has spread outside the kidney); 2) age \geq 18 years; 3) ability to speak, read, and write English; and 4) no current diagnosis of psychosis or dementia.

Scale construction sample. The survey included responses from 34 kidney cancer patients who attended an annual meeting of the Kidney Cancer Association (KCA) and 14 expert physicians with experience treating advanced kidney cancer. Patients ranged in age from 37–69 years, with a mean (SD) of 55 (7.6) years; nearly two thirds (65%) were male. All participants had undergone a prior nephrectomy. Some had experienced either radiation therapy (15%) or chemotherapy (47%); 35% currently were receiving some type of treatment.

Most clinicians had expertise in medical oncology (71%); among specialties mentioned were surgical oncology/urology (21%) and pharmacology (7%). They had an average of 14 years' experience, treating from 50–3,000 patients.

FKSI validation sample. Patients in the instrument validation phase consisted of 141 patient members of the KCA. The organization sent information about the study via postal mailing to approximately 7,500 of its constituents and via elec-

tronic mailing to approximately 8,000; the KCA also posted information in its monthly electronic newsletter. In all, 318 individuals called to inquire about participating in the study; contact was made with 309 of these individuals. Eligible and interested patients ($n = 157$) were mailed a consent form to be signed and returned along with study materials, and 143 interviews were scheduled; a total of 14 patients were not enrolled due to illness or death ($n = 2$), declining to participate ($n = 6$), failure to return consent forms ($n = 4$), or invalidation of consent forms ($n = 2$). Ultimately, 141 kidney cancer patients participated in the validation phase of this study (one consenting patient died, and one could not be reached).

Compliance with the study assessments was very good: 100% at baseline (Time 1), 96% at 3–7 days post baseline (Time 2), and 93% at 2–3 months post baseline (Time 3). Twelve patients went off study due to death ($n = 4$), inability to contact ($n = 6$), illness ($n = 1$), and administrative error ($n = 1$). Complete data from all three assessments were available from 126 patients (89% of enrolled sample).

Participants in the validation phase had a mean age of 60 (SD 9.8; range, 37–82) years; they mostly were male (67%) and predominately were Caucasian (99%). Eighty-two percent had an Eastern Cooperative Oncology Group–Performance Status Rating (ECOG–PSR) of 0–1 at baseline, with about one third of patients reporting either a history of treatment with chemotherapy (35%) or radiation therapy (32%) and most currently not receiving any form of treatment (66%).

FKSI scale development survey. The symptom index was constructed using a multistep method drawn from Cella et al¹⁵ that included item generation, item review/reduction, and scale construction. Candidate items were generated based on a review of symptoms or concerns related to the FACT–G, two independent literature reviews, two independent reviews of the FACIT bank of HRQL items ($n = 400$ items), and consultations with three kidney cancer experts.

The 34 scale-development patients and 14 expert clinicians completed a brief survey to collect information on sociodemographic characteristics, treatment, and symptoms. In a brief questionnaire assessing treatment response for kidney cancer, clinicians were asked to assign 100 points across nine domains of symptoms or concerns generated from the literature reviews (hematuria, pain, appetite and weight loss, fatigue, fever, urinary symptoms, pulmonary symptoms, distress or worry, or “other” [write-in]) based on the relative weight each domain should receive.

In an item-rating survey derived from literature review, expert consultation, and review of the FACIT item bank, both patients and clinicians were asked to select, from 34 symptoms/concerns associated with advanced kidney cancer, “no more than 10 that you would consider to be the most important symptoms or concerns to monitor when assessing the value of drug therapy for advanced kidney cancer.” Of the 10 symptoms nominated, each respondent then was asked to select up to 5 as “the very most important symptoms or concerns to monitor when assessing the value of drug treatment for advanced kidney cancer.” Additional write-in nominations also were solicited.

FKSI validation study. Patients participating in the validation phase of the study also completed the FKSI (see www.supportiveoncology.net/journal/0404.html to download a copy of the FKSI); the FACT-G, a widely known instrument with established reliability, validity, and responsiveness to change¹³; the ECOG-PSR, a single-item survey recording a patients' degree of participation in typical activities without a need for rest¹⁶; and the Global Rating of Change Scale (GRCS), a series of questions concerning the rate of change in patients' health status over a specified period.¹⁷ For each domain, patients specified whether they were feeling worse, about the same, or better. If worse or better, patients then quantified the recognized change using a rating scale ranging from +7 (very great deal better) to -7 (very great deal worse).

PROCEDURE

A convenience sample of patients attending an annual meeting of the KCA was enrolled in the study's scale construction component. After giving verbal consent, the advanced kidney cancer patients and healthcare providers completed their respective sociodemographic forms and the scale development survey; the patients also completed a clinical form.

Item review/reduction. Based on responses to the scale-development survey, the frequency of patient and expert endorsement of each symptom/concern was tabulated. The items most commonly endorsed by the combined sample were retained in the final symptom index; items were retained depending on their probability of chance endorsement as 1 of the top 10 symptoms/concerns (29%). Additionally, a multilingual translation team reviewed the items to be retained for understandability and the ability to be translated meaningfully into commonly spoken languages.

Scale construction. When the pool of items was reduced, the original panel of expert providers received the candidate questions for final review, comment, and revision. The final 15-item scale (FKSI-15) was evaluated to ensure that it met standards of a sixth-grade reading level or below; an even shorter 10-item version (FKSI-10) comprised the first 10 items on the FKSI-15. Both versions were analyzed for reliability and validity.

FKSI VALIDATION STUDY

After eligibility screening, respondents were scheduled to take part in a baseline telephone interview and were mailed a consent form and a copy of the first assessment measures. After providing consent, these patients completed the Time 1 questionnaire battery via computer-assisted telephone interview (CATI). Patients attending the KCA conference completed the first assessment during the conference.

Three to 7 days following the baseline assessment, all patients completed only the Time-2 FKSI via CATI. To examine the FKSI's responsiveness to change over time, patients completed the same battery of questionnaires 2-3 months following the baseline administration via CATI (Time 3).

Analysis plan. Patient responses to the FKSI were ana-

lyzed for internal consistency, stability (test-retest reliability), convergent and discriminant validity, and responsiveness to change in clinical status. In addition, the minimally important difference (MID), defined as the smallest score difference or change that is meaningful to patients and their healthcare providers, was estimated.

Internal consistency analyses were performed at each assessment for the FKSI and the FACT-G and its subscale scores using Cronbach's α coefficients. Test-retest reliability was evaluated for FKSI at Time 1 and Time 2. Cross-sectional analyses were conducted on FKSI, FACT-G, and subscale scores at baseline to determine their ability to differentiate groups according to ECOG-PSR. To examine responsiveness to change, longitudinal data were used to calculate change scores (Time 1 to Time 3) for each patient. Patients were categorized as "worse," "same," or "better" according to their responses to the GRCS.

Both distribution- and anchor-based methods were used to identify MIDs for the FKSI. Distribution-based methods included 0.3 SD, 0.5 SD, and one standard error of measurement (SEM). According to previous research, the MID usually lies between these two values and, often, near 0.3 SD.¹⁸⁻²¹ We also used 1.0 SEM as a criterion for estimating magnitude of change. Often in the range of 0.3-0.50 SD in HRQL assessments, 1.0 SEM has been suggested as a reasonable estimate of a likely MID.²²

The SEM was computed as follows:

$$\sigma_x \sqrt{1 - r_x}$$

where r_x = the reliability of the measure and σ_x = the SD of the measure at baseline.

Anchor-based methods used the cross-sectional and longitudinal analyses described above (eg, GRCS, ECOG-PSR). For the cross-sectional analyses, effect sizes were calculated by dividing the mean difference in FKSI scores between adjacent categories of the PSR anchor by the standard deviation of the relevant score. For example:

$$\text{Effect size} = \frac{\text{mean difference in subscale score between ECOG-PSR 0 vs 1}}{\text{standard deviation of the score}}$$

Effect sizes for the longitudinal analyses were calculated by dividing the mean change score for each category of change in the GRCS anchor (worse, same, better) by the standard deviation of the baseline score.

Results

Table 1 presents the ratings and percent endorsement of patients, physicians, and the combined sample for retained items for the final FKSI scale. Of the 16 items endorsed by patients per the criteria, we retained the top-scoring eight items and added six items ranked highly from the provider list of priority symptoms (ie., weight loss, treatment side effect bother, bone pain, shortness of breath, cough, fever bother); these items were not included in the top items from the patient list. Finally, one item (hematuria) was added to this list of 14 items,

Table 1

Patient, Provider, and Combined Sample Item Rankings and Rates of Endorsement

ITEM	PATIENTS (n = 34)		PROVIDERS (n = 14)		POOLED (n = 48)	
	% ENDORSED TOP 10	RANKING	% ENDORSED TOP 10	RANKING	% ENDORSED TOP 10	RANKING
I am able to enjoy life	64.7	1	28.6	13	31.9	5
I worry that my condition will get worse	55.9	2	21.4	19	25.4	10
I feel fatigued	47.1	3	28.6	13	23.1	13
I have a lack of energy (fatigue)	44.1	4	85.7	2	59.0	1
I am able to work (includes work at home)	44.1	4	28.6	13	20.2	17
I have pain	41.2	6	78.6	4	51.1	2
I am sleeping well	38.2	7	7.1	25	13.9	23
I have a good appetite	38.2	7	50.0	5	27.9	7
I feel sad	35.3	9	14.3	20	16.0	21
I worry about dying	35.3	9	14.3	20	17.4	20
I am enjoying the things I usually do for fun	35.3	9	14.3	20	14.5	22
I have been short of breath	35.3	9	42.9	7	30.3	6
I am bothered by side effects of treatment	35.3	9	85.7	2	42.9	4
I am content with the quality of my life right now	32.4	14	28.6	13	20.2	17
I have nausea	29.4	15	28.6	13	20.2	17
I am forced to spend time in bed	29.4	15	35.7	10	23.7	12
Because of my physical condition, I have trouble meeting the needs of my family	23.5	17	42.9	7	22.9	14
I feel ill	23.5	17	28.6	13	21.6	16
I am losing hope in the fight against my illness	23.5	17	35.7	10	25.2	11
I have bone pain	20.6	20	42.9	7	25.8	9
I feel nervous	20.6	20	0	29	5.9	28
I am interested in sex	20.6	20	0	29	1.5	30
I have been coughing	20.6	20	35.7	10	22.3	15
I am bothered by constipation	17.7	24	7.1	25	6.5	26
I feel weak all over	14.7	25	7.1	25	6.5	26
I urinate more frequently than usual	14.7	25	0	29	1.5	30
I have muscle pain	11.8	27	0	29	1.5	30
I am losing weight	11.8	27	92.9	1	50.8	3
I am bothered by fevers	11.8	27	50.0	5	26.5	8
I have pain in my back	8.8	30	14.3	20	8.6	24
I am satisfied with my sex life	8.8	30	0	29	1.5	30
I have trouble controlling my urine	8.8	30	7.1	25	3.6	29
I have swelling in my stomach area	8.8	30	0	29	1.5	30
I am bothered by blood in my urine	5.9	34	14.3	20	8.6	24

Chance probability of endorsement among top 10 = 29.4%; percentages exceeding chance are italicized.

although it was not heavily endorsed by either patients or providers; instead, expert review before validation testing emphasized that it was relatively rare but still an important clinical and quality-of-life issue to patients. The uncommon nature of hematuria may be a reason that it was not highly endorsed in the survey; thus, it was allowed into the final (FKSI-15) version of the instrument.

The scale items were reviewed to ensure that items were allocated to kidney cancer-relevant domains proportionate to weightings provided by clinicians. A symptom with a median score between 1 and 10 was allocated one question, whereas a symptom with a median score greater than 10 received two or more questions. Thus, the domains of pain, appetite/weight loss, fatigue, and pulmonary symptoms were allocated two

items, and the domains of hematuria, fever, urinary symptoms, and distress/worry each were represented with one question.

Reliability. Descriptive statistics for the FACT-G total score and subscales and FKSI scales are presented in Table 2. All FKSI and FACT-G total and subscale scores showed high internal consistency at the baseline assessment (Cronbach's α range, 0.80–0.92), Time 2 assessment (FKSI only; Cronbach's α range, 0.81–0.84), and Time 3 assessment (Cronbach's α range, 0.76–0.93). All Cronbach's α values were > 0.70, which is a common minimum standard for internal consistency reliability. Both the 10-item and 15-item FKSI scores demonstrated good internal consistency (FKSI-15: Cronbach's α range, 0.84–0.88; FKSI-10: Cronbach's α range, 0.81–0.85), suggesting that they can be used as independent measures of symptoms and functioning.

Table 2
Descriptive Statistics for the FACT-G and FKSI Scales

FACT SCALE/SUBSCALE (NO. OF ITEMS)	MEAN			SD			OBSERVED SCORE RANGE			CRONBACH'S α		
	TIME 1 (n = 141)	TIME 2 (n = 135)	TIME 3 (n = 131)	TIME 1 (n = 141)	TIME 2 (n = 135)	TIME 3 (n = 131)	TIME 1 (n = 141)	TIME 2 (n = 135)	TIME 3 (n = 131)	TIME 1 (n = 141)	TIME 2 (n = 135)	TIME 3 (n = 131)
PWB (7)	22.50	–	22.05	5.35	–	5.93	4–28	–	2–28	0.86	–	0.88
SWB (7/6) ^a	23.08	–	22.74	4.70	–	4.55	6–28	–	9–28	0.80/0.82	–	0.76/0.80
EWB (6)	17.61	–	17.81	4.63	–	4.44	4–24	–	2–24	0.82	–	0.82
FWB (7)	20.11	–	19.76	6.13	–	6.43	4–28	–	1–28	0.87	–	0.89
FACT-G (27/26) ^b	83.30	–	82.52	16.08	–	17.13	39–107	–	35–106	0.92/0.91	–	0.93/0.93
FKSI-10	29.73	30.04	29.63	6.76	6.30	7.20	10–40	10–40	9–40	0.82	0.81	0.85
FKSI-15	46.21	46.49	45.92	9.01	8.30	9.89	18–60	19–60	14–60	0.85	0.84	0.88

^aCronbach's α s in parentheses are for 6 items without item, "I am satisfied with my sex life."

^bCronbach's α s in parentheses are for 26 items without item, "I am satisfied with my sex life."

Abbreviations: FACT-G= Functional Assessment of Cancer Therapy-General; FKSI = Functional Assessment of Cancer Therapy-Kidney Symptom Index; PWB=physical well-being; SWB=social/family well-being; EWB=emotional well-being; FWB=functional well-being.

Table 3
Baseline FACT-G and FKSI Scale Scores by ECOG-PSR

SCALE	ECOG-PSR	n	MEAN (SD)	COMMON SD	GROUP COMPARISONS			
					GROUP	MEAN DIFFERENCE	EFFECT SIZE ^a	P VALUE ^b
PWB	0	65	26.15 (2.03)	5.35	0-1	4.59	0.86	< 0.0001
	1	50	21.56 (3.7)		1-2	6.76	1.26	
	2	25	14.8 (5.29)		0-2	11.35	2.12	
SWB	0	65	24.65 (3.68)	4.70	0-1	2.36	0.50	0.0006 ^c
	1	50	22.29 (4.94)		1-2	1.43	0.30	
	2	25	20.86 (5.27)		0-2	3.79	0.81	
EWB	0	65	18.48 (4.8)	4.63	0-1	1.24	0.27	0.0653
	1	50	17.24 (4.52)		1-2	1.20	0.26	
	2	25	16.04 (4.14)		0-2	2.44	0.53	
FWB	0	65	23.71 (4.04)	6.13	0-1	5.25	0.86	< 0.0001
	1	50	18.46 (5.76)		1-2	4.70	0.77	
	2	25	13.76 (4.78)		0-2	9.95	1.62	
FACT-G	0	65	92.99 (10.84)	16.08	0-1	13.44	0.84	< 0.0001
	1	50	79.55 (14.76)		1-2	14.09	0.88	
	2	25	65.46 (11.62)		0-2	27.53	1.71	
FKSI-10	0	65	34.33 (4.23)	6.76	0-1	6.59	0.97	< 0.0001
	1	50	27.74 (4.93)		1-2	5.94	0.88	
	2	25	21.80 (6.16)		0-2	12.53	1.85	
FKSI-15	0	65	52.53 (5.19)	9.01	0-1	9.17	1.02	< 0.0001
	1	50	43.36 (6.93)		1-2	7.92	0.88	
	2	25	35.44 (7.58)		0-2	17.09	1.90	

^aCalculated for specified row comparisons.

^bAll P values reflect results of three-group ANOVA comparing PSR 0 vs 1 vs 2.

^cSWB differentiated group PSR = 0 from other groups but did not differentiate group PSR = 1 from PSR = 2 (Tukey's post hoc).

Abbreviations: ECOG-PSR = Eastern Cooperative Oncology Group Performance Status Rating; FACT-G= Functional Assessment of Cancer Therapy-General; FKSI = Functional Assessment of Cancer Therapy-Kidney Symptom Index; PWB = physical well-being; SWB = social/family well-being; EWB = emotional well-being; FWB = functional well-being; 0 = normal activity; 1 = symptoms but fully ambulatory; 2 = requires some daytime rest (< 50% of day).

The stability of the FKSI over time was high, with an intraclass correlation (ICC) of 0.90 for FKSI-15 and 0.86 for FKSI-10; the range of possible values for the ICC was 0–1.0. Thus, the two-symptom indices show high test-retest reliability between baseline and 3–7 days post baseline.

Convergent validity. The associations between FKSI, FACT-G, and subscale scores were evaluated using Spearman correlations. Because they capture primarily physical symptoms of disease, FKSI scores were expected to be more highly correlated with physical domains of the FACT-G (physical well-being [PWB]

and functional well-being [FWB]) than with psychosocial (emotional well-being [EWB] and social/family well-being [SWB]) scores. Indeed, this was the case, with very high correlations (all $P < 0.0001$) found between FKSI and PWB scores (r range, 0.87–0.89; $P < 0.0001$), FWB scores (r range, 0.80–0.85; $P < 0.0001$), and FACT-G scores (r range, 0.85–0.88; $P < 0.0001$). Correlations with EWB and SWB were moderately high for the FKSI (r range, 0.37–0.65; $P < 0.0001$).

The FKSI-15 contains three PWB items, three FWB items, and one EWB item; the magnitude of these correlations

Table 4

Change in FACT–G and FKSI Scores by GRCS

GRCS DOMAIN	CHANGE IN FACT–G AND FKSI SCALES						
	ΔPWB PWB ^a	ΔSWB SWB	ΔEWB EWB ^a	ΔFWB FWB ^a	ΔFACT–G OVERALL HRQL	ΔFKSI–10 SYMPTOM ^a	ΔFKSI–15 SYMPTOM ^a
Worse, n ^b	16	2	8	15	6	13	13
Mean	–5.44	–4.50	–3.25	–4.20	–9.11	–4.31	–6.00
SD	6.51	2.12	3.20	5.78	8.59	5.06	6.82
Same, n ^b	100	117	107	104	106	108	108
Mean	–0.40	–0.41	0.04	–0.44	–1.52	–0.42	–0.66
SD	2.82	3.21	3.20	3.46	7.62	4.37	5.66
Better, n ^b	15	11	16	12	18	10	10
Mean	1.40	–1.85	1.00	–0.25	–3.31	2.20	3.40
SD	3.11	1.97	1.97	2.90	12.89	4.54	6.57
F	17.77	2.66	5.29	6.77	2.43	6.58	7.74
(df)	(2,128)	(2,127)	(2,128)	(2,128)	(2,127)	(2,128)	(2,128)
P	< 0.0001	0.0739	0.0062	0.0016	0.0918	0.0019	0.0007

^aDifference is between “worse” vs “better” and “same,” but no difference between “same” and “better.”

^bWorse: <1 on GRCS; same: –1 to 1 on GRCS; better: >1 on GRCS.

Abbreviations: FACT–G = Functional Assessment of Cancer Therapy–General; FKSI = Functional Assessment of Cancer Therapy–Kidney Symptom Index; GRCS = Global Rating of Change Scale; PWB = physical well-being; SWB = social/family well-being; EWB = emotional well-being; FWB = functional well-being; HRQL = health-related quality of life.

was inflated by the redundancy of items between the scales. However, when redundancy was removed, the FKSI–15 (or FKSI–10) remained highly correlated with abbreviated PWB (r range, 0.81–0.84), FWB (r range, 0.77–0.79), and EWB (r range, 0.53–0.66). FKSI–10 correlations were 0.82 and 0.81 for the reduced PWB, 0.75 for FWB (at Time 2 and Time 3), and from 0.37–0.66 for EWB and SWB.

Discriminant (known-groups) validity. The ECOG–PSR was trichotomized into PSR = 0, PSR = 1, and PSR > 1 (Table 3). For the FACT–G total score and subscales and the FKSI–10 and FKSI–15, all scores across PSR groups were in the appropriate direction; patients with the lowest PSR had the highest FACT–G and FKSI scores (ie, greater well-being and symptom status), and those with a higher PSR had FACT and FKSI

scores reflecting poorer well being and symptom status.

The FKSI–10, FKSI–15, and all FACT–G subscale and aggregated scores differentiated patients grouped by PSR (all P < 0.001); however, in the case of EWB, the discrimination between groups was marginally significant (Table 3). Following Cohen’s guidelines for effect sizes,²³ those for adjacent PSR groups were small to moderate for SWB (0.30–0.50) and EWB (0.26–0.27) and large for the FACT–G total score, PWB, and FWB (all P < 0.0001). Effect sizes of adjacent PSR groups also were large for the two FKSI scales (0.88–1.02; P < 0.0001).

Responsiveness to clinical change. Changes in FKSI and FACT–G aggregate and subscales scores were calculated by subtracting patients’ baseline scores from the Time 3 scores. Patients’ GRCS scores were categorized as “worse,” “same,” or “better” by collapsing GRCS domain scores that were rated on a 15-point scale (–7 to +7). Sample size restrictions at the extremes of change demanded that all gradations be collapsed into one “changed” category for “worse” and another for “better.” The GRCS scores were categorized as “worse” if they were rated ≤ 1, “same” if rated –1 to +1, and “better” if rated ≥ 1.

For most scales, changes in scores were in the anticipated direction, such that patients who had a worse rating on the GRCS had worsening scores on the relevant FACT scale, patients who had an improved GRCS rating tended to have improvements in the relevant scale, and patients who remained the same on the GRCS tended to have changes in scores on most subscales that were between those of the two other groups (Table 4).

Estimating MID. Distribution-based and anchor-based methods were used to estimate MID for the FKSI (Tables 5 and 6). Table 5 displays the distribution-based estimates of MID at baseline and at Time 3, as well as the change from baseline to Time 3. The full range of distribution-based MID

Table 5

Distribution-based Estimates of MID for FKSI

	n	0.3 SD	0.5 SD	CRITERION SEM ^a	RANGE OF MID
FKSI–10					
Baseline	141	2.25	3.38	2.55	2–4
Time 3	131	2.40	3.60	2.71	
Baseline to Time 3 change	131	1.55	2.32	n/a	
FKSI–15					
Baseline	141	3.00	4.50	2.92	2–5
Time 3	131	3.30	4.95	3.20	
Baseline to Time 3 change	131	2.05	3.07	n/a	

^aSEM = $\sigma_x (1-r_x)^{1/2}$

Abbreviations: MID = minimally important differences; FKSI = Functional Assessment of Cancer Therapy–Kidney Symptom Index; SEM = standard error of measurement.

Definitions: σ_x = the standard deviation of the scale / aggregate score; r_x = the reliability (internal consistency) of the scale / aggregate score.

Table 6
Anchor-based (GRCS) Estimates of MIDs for FKSI

SCALE	GRCS ^a	n	MEAN (SD)	COMMON SD	GROUP	GROUP COMPARISONS		P VALUE ^b
						MEAN DIFF	ES	
FKSI-10 change ^c	Worse	13	-4.31 (5.06)	4.64	Worse vs Same	-3.89	-0.84	0.0019
	Same	108	-0.42 (4.37)		Same vs Better	-2.62	-0.56	
	Better	10	2.20 (4.54)		Worse vs Same	-5.34	-0.87	
FKSI-15 change ^c	Worse	13	-6.00 (6.82)	6.14	Worse vs Same	-5.34	-0.87	0.0007
	Same	108	-0.66 (5.66)		Same vs Better	-4.06	-0.66	
	Better	10	3.40 (6.57)		Same vs Better	-4.06	-0.66	

^a“Worse” = GRCS ≤ 1; “same” = GRCS -1 to +1; “better” = GRCS > 1.

^b Difference is between GRCS “worse” vs “same” and “better;” no difference between “same” and “better.”

^c FKSI score change: score at Time 3 minus score at Time 1 (baseline).

Abbreviations: GRCS = Global Rating of Change Scale; FKSI = Functional Assessment of Cancer Therapy–Kidney Symptom Index; MIDs = minimally important differences; ES = effect size, calculated for adjacent categories.

was 1.55–3.60 points for the FKSI-10 and 2.05–4.95 points for the FKSI-15. The effect sizes for the cross-sectional anchor-based MID estimates using PSR for FKSI-10 and FKSI-15 were in the large range,²³ and the effect sizes for the longitudinal anchor-based MID estimates using GRCS were in the moderate-to-large range. Change scores associated with effect sizes > 0.50 (moderate) exceeded what would be considered minimal. In reconciling the rather large effect sizes of the anchor-based comparisons with the smaller distribution-based estimates, it is

reasonable to suggest 2–4 as the MID range for the FKSI-10 and 3–5 as the MID range for the FKSI-15.

Proportions of patients changing and associated effect sizes. The availability of baseline and 2–3 month change data and MID range estimates for the FACT subscales and FKSI scale enabled several additional threshold analyses. For each MID value, the proportion of cases that would have been classified as changed depending upon the MID applied was determined. The range of possible MIDs for each scale (2–3 points for each

Table 7
Patient Proportions and Magnitude of Change Across Possible MIDs: Baseline to Time 3

SCALE	BETTER				SAME				WORSE			
	%	\bar{X}_Δ	SD _Δ	ES	%	\bar{X}_Δ	SD _Δ	ES	%	\bar{X}_Δ	SD _Δ	ES
FKSI-10												
2 pt MID	24.82	4.51	2.59	1.74	36.17	-0.12	0.82	-0.15	39.01	-5.13	3.94	-1.30
3 pt MID	19.86	5.14	2.53	2.03	46.81	-0.13	1.20	-0.11	33.33	-5.81	4.04	-1.44
4 pt MID	12.06	6.53	2.37	2.76	63.83	-0.16	1.86	-0.09	24.11	-7.33	4.32	-1.70
FKSI-15												
3 pt MID	23.40	5.79	3.75	1.54	39.72	-0.08	1.39	-0.06	36.88	-7.19	5.34	-1.35
4 pt MID	17.73	6.68	3.91	1.71	51.06	-0.06	1.88	-0.03	31.21	-8.18	5.50	-1.49
5 pt MID	10.64	8.47	4.21	2.01	63.83	0.04	2.46	0.02	25.53	-9.46	5.71	-1.66
PWB												
2 pt MID	19.15	3.52	1.74	2.02	46.81	-0.16	0.72	-0.22	34.04	-5.00	4.29	-1.17
3 pt MID	12.06	4.41	1.62	2.72	61.70	-0.14	1.17	-0.12	26.24	-6.22	4.56	-1.36
SWB												
2 pt MID	19.15	3.99	1.83	2.18	40.43	-0.31	0.76	-0.41	40.43	-3.63	1.96	-1.85
3 pt MID	13.48	4.81	1.56	3.08	59.57	-0.48	1.26	-0.38	26.95	-4.75	1.89	-2.51
EWB												
2 pt MID	28.37	3.38	2.01	1.68	40.43	-0.09	0.74	-0.12	31.21	-4.00	1.97	-2.03
3 pt MID	14.18	4.75	2.07	2.29	62.41	0.15	1.33	0.11	23.40	-4.96	1.69	-2.93
FWB												
2 pt MID	23.40	3.18	1.42	2.24	34.04	0.08	0.77	0.10	42.55	-4.42	3.72	-1.19
3 pt MID	12.77	4.17	1.25	3.34	58.87	-0.07	1.43	-0.05	28.37	-6.03	4.08	-1.48
FACT-G												
5 pt MID	17.73	8.32	4.70	1.77	40.43	0.39	2.38	0.16	41.84	-10.54	7.02	-1.50
6 pt MID	14.18	9.12	4.96	1.84	51.06	-0.03	3.16	-0.01	34.75	-12.00	7.22	-1.66

Abbreviations: MIDs = minimally important differences; FKSI = Functional Assessment of Cancer Therapy–Kidney Symptom Index; ES = effect size; PWB = physical well-being; SWB = social/family well-being; EWB = emotional well-being; FWB = functional well-being; FACT-G = Functional Assessment of Cancer Therapy–General.

Definitions: \bar{X}_Δ = Mean change at 2–3 months post baseline; SD_Δ = Standard deviation of change at 2–3 months post baseline; ES = $\bar{X}_\Delta / SD_\Delta$.

of the FACT–G subscales, 5–6 points for the FACT–G, 2–4 points for the FKSI–10, and 3–5 points for the FKSI–15) was selected, and the proportion of patients who were better or worse at the final assessment was computed. These data are displayed in Table 7.

The proportion of improved patients at the low end of the MID was in the 20%–25% range for each subscale (including FKSI); that same proportion was in the 10%–15% range at the high end of the MID. By comparison, the proportion of worsened patients at the low end of the MID was in the 30%–40% range and close to 25% for most of the scales when the upper bound estimate was applied. The FKSI–based estimates of decline and improvement were comparable to the established ranges in the FACT–G scales.^{18–21}

Discussion

The FKSI was derived from the FACIT item bank, a literature review, and clinician and patient input and was designed to capture the most important symptoms and concerns of people with kidney cancer. The FKSI–10 and FKSI–15 demonstrated strong measurement properties, with internal consistency reliability coefficients at all time points exceeding acceptable standards. Furthermore, the stability (test-retest reliability) of the instrument was quite high. Across a 3- to 7-day time period during which clinical status was expected not to change, patient responses to the FKSI questions also did not change, which supported the reliability of the 7-day recall period in the questionnaire.

In addition to support for instrument reliability, there was evidence of convergent validity of the FKSI, ECOG–PSR, and FACT–G subscales. At baseline, the symptom indices were significantly related to all FACT scales. The relationships were strongest with PWB, FWB, and FACT–G, as would be expected, given that the FKSI represents largely physical symptoms. The FKSI–10 and FKSI–15 successfully differentiated patients known to differ at baseline by PSR, with patients reporting better performance status also reporting better symptom status. The FKSI scales also demonstrated responsiveness to change in clinical status over time, successfully differentiating patients who reported improvement, no change, or worsening in their status across multiple FACT–G domains.

Cross-sectional and longitudinal *distribution*-based criteria suggested estimates of MIDs for the FKSI–10 and FKSI–15 of 2 and 3 points, respectively. Cross-sectional and longitudinal *anchor*-based criteria were suggestive of larger MIDs, reflecting large effects in excess of the likely MID. This offers encouraging validity for the symptom index; however, modification is needed when estimating MIDs from these data. Because the effect sizes associated with the anchor-based analyses were consistently in the moderate-to-large range, the MID estimates were based on review of anchor-based data as modified by distributional information and current standards for estimating MIDs. Keeping this in mind, we suggest 2–4 as the MID range for the FKSI–10 and 3–5 as the MID range for the FKSI–15. We recommend applying the lower end of the MID

range for group comparisons and the higher end for individual classification. Therefore, for example, if one were using the FKSI–15, a group difference over time of 3 points (about 0.3 SD) would likely reflect a meaningful treatment effect, and a 5-point change (about 0.5 SD) in an individual would provide reasonable confidence in classifying people as changed from baseline.

Post-hoc threshold analyses allowed us to determine, for each MID value, the proportion of cases that would have been classified as changed depending upon the MID applied. Based on these analyses, approximately 10%–25% of these patients improved over the 2–3 month study, whereas 25%–40% worsened. This phenomenon reflects the nature of advanced kidney cancer, which continues to have limited therapeutic options enabling patient improvement. Effect sizes of the improved and worsened cases consistently exceeded 1.0 and quite often exceeded 2.0, suggesting the potential for responsiveness to clinical change when applied in clinical trials.

Given the comparable psychometric performance of the two FKSI versions, the choice of using the FKSI–10 or FKSI–15 likely depends upon preferences for the length of the instrument and clinical interest in the items represented on each scale. Each individual user must decide whether competing considerations of content relevance, clinical interpretability, and length would suggest use of one or the other in a given application. Both appear to perform well.

Among this study's limitations was the use of convenience samples of constituents from KCA in the development and validation phases. This self-selected sample may have represented patients on the healthier end of the spectrum relative to the population of kidney cancer patients, even though all were diagnosed with advanced kidney cancer. In addition, clinical data on the patient sample in the validation phase were limited; a broader range of clinical data may have provided additional insight into both patients' health status and the implications for generalizing results to a broader population of advanced kidney cancer patients.

Patient and clinician input concerning priority symptoms and concerns for monitoring when undergoing advanced kidney cancer treatment were key strengths of this study. The FKSI expands the scope of our previous work,¹⁵ which involved expert physician and nurse input only in the development of symptom indices for nine tumor types.

Observed consistencies and differences in item endorsement between patients and physicians were informative in several respects. First, the 15 items comprising the final FKSI included *all* symptoms endorsed as having high priority by both patients and experts, and the top 8 symptoms and concerns were expressed by kidney cancer patients irrespective of expert rating. It also comprised the top six physician-endorsed concerns, so no very important area would appear to be unrepresented.

Second, the representation of symptoms such as pain and fatigue with more than one question reflects adherence to relative weight placed on these important concerns by the expert

panel. Patients rated both fatigue items (“I feel fatigued,” “I have a lack of energy”) among their highest priority symptoms, whereas physicians included only one fatigue item; conversely, physicians included two pain items (“I have pain,” “I have bone pain”) among their priority items, whereas patients included only one pain item. These ratings were consistent with patient and physician recognition of the well-known prevalence and importance of pain and fatigue in advanced cancer,^{24,25} but the discrepancies may have reflected, in part, relative differences in available treatment options for pain versus fatigue.

Third, comparison of patient and physician priorities revealed the added weight patients assigned to symptoms or concerns of a psychosocial nature, including emotional distress, losing hope, concern about illness progressing, and HRQL concerns. The psychosocial concerns of patients diagnosed and undergoing treatment for advanced disease are critically important, yet the composition of the final FKSI reflects the original purpose of this study—to develop and validate a scale targeted to the symptom experience of kidney cancer patients.

The use of brief assessment tools to assess symptomatology

may serve the interests of patients undergoing treatment, clinical investigators, and regulatory authorities. From a clinician’s perspective, assessment of symptomatology may represent an efficient and clinically relevant means of obtaining information related to the symptom component of quality of life; it also may help to identify patients who would benefit from palliative interventions.²⁶ Systematic symptom assessment may help to clarify a treatment’s toxicity, its potential palliative benefit, or a need to make a change in the patient’s clinical management.²⁷

This study provides evidence that items pertaining to symptoms and concerns of advanced kidney cancer patients may be derived from well established, multidimensional quality of life questionnaires. The FKSI represents a brief, targeted instrument for assessing symptoms and concerns that both clinicians and patients have endorsed as the most important to monitor when evaluating treatment for advanced kidney cancer. We continue to advocate continued assessment of quality of life across a range of biopsychosocial domains—and the FKSI may provide acceptable alternatives in the face of symptom-focused clinical or research interests.

References

- Jemal A, Thomas A, Murray T, Thun M. Cancer statistics, 2002. *CA Cancer J Clin* 2002;52:23–47.
- American Cancer Society. Cancer Facts and Figures 2003. Atlanta, Ga: American Cancer Society; 2003.
- Tsui K-H, Shvarts O, Smith RB, Figlin R, de Kernion JB, Belldegrun A. Renal cell carcinoma: prognostic significance of incidentally detected tumors. *J Urol* 2000;163:426–430.
- Bostwick DG. In: Lenhard RE Jr, Osteen RT, Gansler T, eds. *The American Cancer Society’s Clinical Oncology*. Atlanta, Ga: American Cancer Society; 2001:415–419.
- Hartmann JT, Bokemeyer C. Chemotherapy for renal cell carcinoma. *Anticancer Res* 1999;19:1541–1543.
- Stadler WM. Targeted agents for the treatment of advanced renal cell carcinoma. *Cancer* 2005;104:2323–2333.
- Coons SJ, Kaplan RM. Assessing health-related quality of life: application to drug therapy. *Clin Ther* 1992;14:850–858.
- Fish LS, Lewis BE. Quality of life issues in the management of ovarian cancer. *Semin Oncol* 1999;26:532–539.
- Sutherland HJ, Lockwood GA, Boyd NF. Ratings of the importance of quality of life variables: therapeutic implications for patients with metastatic breast cancer. *J Clin Epidemiol* 1990;43:661–666.
- Kong SX, Gandhi SK. Methodologic assessments of quality of life measures in clinical trials. *Ann Pharmacother* 1997;31:830–836.
- MacKeigan LD, Pathak DS. Overview of health-related quality of life measures. *Am J Hosp Pharm* 1992;49:2236–2245.
- Johnson JR, Temple R. Food and Drug Administration requirements for approval of new anticancer drugs. *Cancer Treat Rep* 1985;69:1155–1157.
- Cella DF, Tulsky DS, Gray G, et al. The Functional Assessment of Cancer Therapy (FACT) scale: development and validation of the general measure. *J Clin Oncol* 1993;11:570–579.
- Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–376.
- Cella D, Paul D, Yount S, et al. What are the most important symptom targets when treating advanced cancer? A survey of providers in the National Comprehensive Cancer Network (NCCN). *Cancer Invest* 2003;21:526–535.
- Zubrod C, Schneiderman M, Frei E III, et al. Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. *J Chronic Dis* 1960;11:7–33.
- Jaeschke R, Singer J, Guyatt G. Measurement of health status: ascertaining the minimal clinically important difference. *Controlled Clin Trials* 1989;10:407–415.
- Yost KJ, Eton DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. *Eval Health Prof* 2005;28:172–191.
- Eton D, Cella D, Yost K, et al. A combination of distribution- and anchor-based approaches determined minimally important differences (MIDs) for four endpoints in a breast cancer scale. *J Clin Epidemiol* 2004;57:898–910.
- Cella D, Eton DT, Lai JS, Peterman AH, Merkel DE. Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. *J Pain Symptom Manage* 2002;24:547–561.
- Cella D, Zagari M, Vandoros C, Gagnon DD, Hurtz HJ, Nortier JW. Epoetin alfa treatment results in clinically significant improvements in quality of life in anemic cancer patients when referenced to the general population. *J Clin Oncol* 2002;21:366–373.
- Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 1999;37:469–478.
- Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Portenoy RK, Thaler HT, Kornblith AB, et al. Symptom prevalence, characteristics and distress in a cancer population. *Qual Life Res* 1994;3:183–189.
- Chang VT, Hwang SS, Feuerman M, Kasmis BS. Symptom and quality of life surveys of medical oncology patients at a Veterans Affairs Medical Center: a role for symptom assessment. *Cancer* 2000;88:1175–1183.
- Guyatt G, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–629.
- Ingham JM, Portenoy RK. Symptom assessment. *Hematol Oncol Clin North Am* 1996;10:21–39.