



Published in final edited form as:

Support Care Cancer. 2018 June ; 26(6): 1889–1895. doi:10.1007/s00520-017-4027-7.

Fatigue among patients with renal cell carcinoma receiving adjuvant sunitinib or sorafenib: Patient-reported outcomes of ECOG-ACRIN E2805 trial

Fengmin Zhao, MHS, PhD¹, David Cella, PhD², Judith Manola, MS¹, Robert S. DiPaola, MD³, Lynne I. Wagner, PhD⁴, and Naomi S. B. Haas, MD⁵

¹Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston MA, USA

²Department of Medical Social Sciences, Northwestern University, Chicago IL

³Rutgers Cancer Institute of New Jersey, New Brunswick USA

⁴Social Sciences & Health Policy Comprehensive Cancer Center, Wake Forest School of Medicine, Winston-Salem, NC, USA

⁵Abramson Cancer Center, University of Pennsylvania, Philadelphia, PA, USA

Abstract

Background—E2805 was a phase III trial to test whether adjuvant sunitinib or sorafenib could improve disease-free survival compared to placebo in patients with renal cell carcinoma. Patient-reported outcomes (PRO), focusing on fatigue, were evaluated as a secondary endpoint.

Patients and Methods—A total of 463 patients participated in the PRO study. Fatigue was measured by FACIT Fatigue Scale and PROMIS Fatigue SF1 measure at baseline, week 10 and week 22. The primary endpoint was change in fatigue score from baseline to week 22, measured by the FACIT Fatigue Scale. Secondly, the psychometric properties of PROMIS Fatigue SF1 were assessed in relation to the FACIT Fatigue Scale.

Results—Fatigue got significantly worse on all arms after two cycles of treatment, and especially so in patients on sunitinib (−9.6 vs. −5.6 on sorafenib vs. −4.7 on placebo). Fatigue remained stable during week 10 and week 22. Overall, the mean score change between baseline and week 22 was −7.9 ($p<0.001$) on sunitinib, −6.4 ($p<0.001$) on sorafenib and −5.6 ($p<0.001$) on placebo arm. The difference in score change was not statistically significant between the two experimental arms and the placebo arm (difference=−2.34 [$p=0.110$] and −0.87 [$p=0.535$] for sunitinib vs. placebo

Corresponding author: Dr. Fengmin Zhao, Department of Biostatistics and Computational Biology, Dana Farber Cancer Institute, Boston, MA 02215, Tel: 617-5829669, fezhao@jimmy.harvard.edu.

Conflicts of Interest: All authors state that they have no conflicts of interest.

AUTHOR CONTRIBUTIONS

FZ did the statistical analysis and wrote the first draft of the manuscript. DC and LW designed the quality of life proposal. JM and NBH designed the parent trial as the study statistician and study chair, respectively. RSD was the ECOG-ACRIN Genitourinary committee chair at the time. All authors interpreted the data, reviewed and edited the manuscript drafts, and approved the final version of the manuscript for submission.

and sorafenib vs. placebo). PROMIS Fatigue SF1 had good internal consistency reliability and construct and criterion validity, and was highly correlated with the FACIT Fatigue Scale score.

Conclusions—Fatigue got worse during study period, especially in patients on sunitinib. The PROMIS Fatigue SF1 was highly correlated with FACIT Fatigue and produced similar results.

Keywords

fatigue; patient reported outcome; renal cell carcinoma; adjuvant therapy; sunitinib; sorafenib

INTRODUCTION

Fatigue is one of the most common symptoms among patients with renal cell carcinoma (RCC)¹. Before 2006, cytokine-based therapies had been the standard of care for advanced RCC², but are associated with high rates of treatment-related fatigue³. In the past few years, several new target therapies have been developed and significantly improved progression-free and overall survival in patients with advanced RCC⁴⁻⁷. With the success in the metastatic setting, these targeted therapies, especially tyrosine kinase inhibitors (TKIs) targeting the vascular endothelial growth factor and its related receptor (such as sorafenib^{4,8} and sunitinib^{5,8,9}), also have been or are being tested in the adjuvant setting. Fatigue remains a major toxicity with these targeted therapies^{4,5,8-10} and can affect patient willingness to stay on therapy.

E2805 is one of the adjuvant trials comparing sunitinib and sorafenib versus placebo in patients with RCC⁸. Fatigue was one of the most common adverse events (others were hypertension, hand foot syndrome and rash) leading to dose reduction in E2805. Consequently the starting dosage was reduced for the two experimental drugs to manage a higher than desired dropout rate due to treatment intolerance. In clinical trials, including E2805, fatigue has been graded using the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE), which is dependent on the functional assessment of patients (i.e., interference in the activities of daily living) by investigators. Physicians often underreport treatment-related symptoms, and patients' self-evaluation of toxicity are recommended in clinical trials^{11,12}. Hence, at the time when the protocol was amended to reduce the starting dosage for the two experimental drugs, patient-reported outcome (PRO), focusing on fatigue, was added to E2805 as a secondary endpoint to fully capture the fatigue toxicity profile of the two TKIs using patients' own report. The inclusion of a placebo arm in the primary trial offers a unique opportunity to quantify the extent to which fatigue can be attributed to therapy, while controlling for multi-factorial factors that commonly contribute to cancer-related fatigue. Levels of fatigue were assessed using the FACIT Fatigue Scale and the PROMIS Fatigue SF1 measure. The FACIT Fatigue Scale measures the experience of fatigue and its impact on daily functioning, over a one week period. There are 13 items in the FACIT Fatigue Scale, scored according to manualized instructions to produce a single total score. PROMIS Fatigue SF1 scale is a 7-item short form of the 95-item PROMIS fatigue item bank (notably, this 95 item bank includes the FACIT Fatigue Scale within it). It is a more recently-developed measure based upon item response theory (specifically the graded response model) and used the pattern of item responses for scoring, rather than the more traditional summation such as that used in the FACIT-Fatigue. The intention going into

this study was to use both questionnaires in order to compare their scores and relative responsiveness to change. The PROMIS Fatigue scale is shorter and more patient friendly; thus, if the two scores are highly correlated, and can produce similar results and same conclusion, it could support the more widespread use of the PROMIS Fatigue scale.

E2805 did not demonstrate superiority of sorafenib or sunitinib relative to placebo with regard to survival benefit among patients with high-risk RCC. However, the inclusion of patient-reported fatigue among a cohort of trial participants offers a unique opportunity to understand the fatigue toxicity profile of two commonly used TKIs, from patient's perspective, while controlling for expectancy and other confounding factors through the inclusion of a placebo arm. Here we report the PRO results of E2805.

METHODS

Study Design

E2805 (ClinicalTrials.gov identifier: NCT00326898), a double-blind, placebo-controlled, randomized phase III trial, was designed to evaluate possible improvement in disease-free survival in locally advanced RCC patients randomly assigned to adjuvant sunitinib (Arm A) or sorafenib (Arm B) versus placebo (Arm C) after radical or partial nephrectomy.⁸ Secondary endpoints included overall survival, toxicity and PRO. The primary purpose of the PRO study was to determine if there were differences in patient-reported fatigue, assessed by the validated FACIT Fatigue Scale, between RCC patients randomized to either of the experimental arms compared to those randomized to placebo. Secondly, the study aimed to evaluate the psychometric properties of the PROMIS[®] Fatigue SF1 measure, and to compare its performance with that of the FACIT Fatigue Scale.

Patients

Patients were eligible for the study if they had histologically or cytological confirmed RCC, undergone a full surgical resection by open or laparoscopic technique with negative margins; an Eastern Cooperative Oncology Group performance status (ECOG PS) of 0–1 and adequate cardiac, renal and hepatic function. Key exclusion criteria included history of distant metastases, prior anti-cancer therapy in either adjuvant or neoadjuvant setting, or serious intercurrent illness.⁸ The protocol was approved by the Institutional Review Boards at each registering institution. All patients provided written informed consent.

Treatment

Initially patients were assigned to 50mg sunitinib x 4 weeks with 2 weeks off, 800mg sorafenib x 6 weeks, or placebo, but the starting dosage was reduced to manage a higher than desired dropout rate due to treatment intolerability after 1450 patients (target accrual goal of 1923 patients) had been enrolled. It was at this point that the PRO study was added to the trial. During the PRO part of the study, patients took sunitinib orally at 37.5mg per day for 4 weeks followed by rest for 2 weeks (1 cycle=6 weeks) or sorafenib orally at 400mg once per day for 6 weeks or equivalent doses of placebo. After cycle 1 or 2, patients who had tolerable grade 2 side effects at worst were escalated to full doses of sunitinib 50mg or sorafenib 400mg orally twice daily in the same schedules. Patients who experienced grade

2 toxicities continued at the reduced starting dose level. Patients with grade 3 or 4 side effects had their dose modified in accordance with the specified dose modification guideline in the protocol. Patients on the placebo arm took the placebos in the same manner in terms of schedule, dose and dose modification. Adverse events were graded using CTCAE, v4.0. In all arms, patients were treated for up to 9 cycles.

Assessment Schedule and Measurement of Fatigue

FACIT Fatigue Scale and the PROMIS Fatigue SF1 measure were administered as one PRO questionnaire at baseline (within 2 weeks prior to the start of protocol therapy), week 10 (Day 28 of cycle 2) and week 22 (Day 28 of cycle 4) after the start of protocol therapy, regardless of whether the patient had disease relapse or discontinued protocol therapy for any reason. The two follow-up assessment time points coincided with the completion of sunitinib in these cycles. If the assessment time point did not coincide with a scheduled clinic visit, patients were given a blank PRO questionnaire form and asked to complete it at home.

Per manual instructions, a prorated score for FACIT-Fatigue was calculated for patients as long as patients responded to more than 50% of the items. The theoretical range of the prorated scale score is 0–52 with higher score indicating less fatigue. A score of less than 30 indicates severe fatigue. The PROMIS User Manual Version 1.1 was used to guide scoring of the PROMIS Fatigue SF1 measure. Briefly, the summed raw score was calculated first and then mapped to trait scores (standardized T scores) in one-to-one correspondence using the score conversion table provided in the manual. T score distributions are standardized such that a 50 represents the average (mean) for the US general population, and the standard deviation around that mean is 10 points. The summed raw score for PROMIS Fatigue SF1 measure ranges 7–35, and the range of the corresponding standardized T scores is 29.4–83.2. Higher score indicates more severe fatigue. The Conversion table for PROMIS measure applies only when all question items on the form have been answered. Therefore, patients with missing items (<2% in the study) had missing value for the standardized T score and were excluded from corresponding analysis.

Statistical Consideration

Since the impact of full-dose TKI inhibition on fatigue was the primary question of interest, FACIT-Fatigue score change from baseline to the week 22 time point was the primary endpoint of the PRO study. The two experimental arms were compared to the placebo arm separately, and the overall type I error rate of two-sided 0.05 for the PRO study was equally split between the two comparisons. With 282 patients (94 patients in each arm) for the primary endpoint, the trial had about 80% power to detect a medium effect size (i.e., minimally important difference/standard deviation) of 0.45 using two-sample t test with a two-sided significance level of 0.025. As stated previously, after the starting dose was amended all subsequent patients were eligible to participate in the PRO study to achieve the above sample size, assuming about 75% participation rate of the PRO study and 80% of attrition rate at the week 22 assessment.

Changes in the fatigue scores between baseline and follow-up visits were examined using paired t test and compared between treatment arms using two-sample t tests. At each time point, fatigue scores were compared between the treatment arms using Wilcoxon rank-sum test. All patients with fatigue data at the specific time point were included in the analysis. Multivariable linear mixed effects models with random intercept (repeated measures within single patients with unstructured covariance matrices) were fit using the maximum likelihood method to estimate the time profile of fatigue and to assess the treatment difference in fatigue over time, after adjusting other covariates (age, sex, ECOG PS, fatigue prior to surgery, fatigue after surgery, disease discovery, history of cardiovascular disease), assuming that any missing data were missing at random. Time points were coded as categorical variable in the mixed effect models. Patients with at least one assessment were included in these models.

The psychometric properties of the PROMIS Fatigue SF1 measure were assessed. Cronbach coefficient α was computed to estimate the internal-consistency reliability. Known-group validity and exploratory factor analysis were conducted to check construct validity. The area under the receiver operating characteristics (ROC) curve was used to examine the criterion validity, and severity of fatigue (severe vs. non-severe fatigue) defined by FACIT Fatigue Scale score was used as the dependent variable in the logistic regression model. Spearman's and Pearson correlation coefficients were used to explore the association between scores on the FACIT Fatigue Scale and PROMIS Fatigue SF1 measure.

All tests were two-sided, $p < 0.025$ was considered statistically significant for the two primary endpoints. No adjustment was made for multiplicity and $p < 0.05$ was used for statistical significance for other endpoints. STATA 13.1 was used to conduct all the analyses.

RESULTS

Patients

A total of 1943 patients were enrolled to E2805 between April 24, 2006 and September 1, 2010⁸, and 463 patients participated in the PRO study and reported their fatigue level at one or more time points. Of them, 418 (90.3%) patients reported data about fatigue at baseline; 45 patients had missing data at baseline but reported some data about fatigue at week 10 and/or week 22 assessments. Taking all three visits together, the most common reason for missing fatigue data was patient not given form by staff ($n=90$), and very few patients had missing data due to patient illness ($n=7$) or death ($n=2$). This report was based on data as of April 14, 2015.

Fatigue measured by FACIT Fatigue Scale

Of the 418 patients who reported their fatigue level at baseline, 381 (91.1%) patients reported fatigue data at week 10 assessment, and 328 (78.5%) patients reported fatigue data at week 22 assessment using FACIT Fatigue Scale. Furthermore, two patients at week 10 and 7 patients at week 22 were excluded from the corresponding analyses due to having 7 items in the FACIT Fatigue Scale unanswered. A total of 321 patients (107 on sunitinib, 95 on sorafenib, 119 on placebo) had score for FACIT-Fatigue at both baseline and week 22

assessments and were the primary analysis population for the PRO study. Treatment arms were well balanced regarding patient characteristics in the 321 patients (Table 1). The 321 patients also had similar characteristics compared to the 142 patients included in the PRO study but did not have fatigue data at both baseline and week 22, and the 1480 patients who did not participate in the PRO study (Supplemental Table S1).

As expected, patients had similar fatigue level at baseline in the three arms with average FACIT-Fatigue score of 42.5 (SD=9.0) in the overall sample (Supplemental Table S2). This is a baseline level of fatigue comparable to that of the general US population¹³. After 10 weeks of treatment, fatigue score reduced (fatigue worsened) on all three arms (Figure 1A). The mean score change between week 10 and baseline was -9.6 ($p<0.001$) on sunitinib, -5.6 ($p<0.001$) on sorafenib and -4.7 ($p<0.001$) on placebo arm. The difference in score change (4.9, $p<0.001$) between the sunitinib arm and the placebo arm reached statistical significance. During week 10 and week 22, fatigue level remained stable in all arms. Overall, the mean score change between week 22 and baseline was -7.9 ($p<0.001$) on sunitinib, -6.4 ($p<0.001$) on sorafenib and -5.6 ($p<0.001$) on placebo arm, and the difference in score change was not statistically significant between the two experimental arms and the placebo arm (Table 2). Linear mixed effect model analysis showed similar results (data not shown).

Patient-reported fatigue using FACIT-Fatigue was cross-checked with investigator-reported fatigue via CTCAE during cycle 2 (week 10) and cycle 4 (week 22). In both assessments, about a quarter of patients had FACIT Fatigue score<30 (severe fatigue) while the CTCAE fatigue grade was zero (Supplemental Figure S1). At cycle 2, 5.6% (6/107) of patients with FACIT-Fatigue score<30 had CTCAE fatigue grade of 3, the proportion was 4.0% (4/101) for cycle 4 assessment.

Fatigue measured by PROMIS Fatigue SF1

Of the 418 patients who reported their fatigue data at baseline, 411 (98.3%) of them answered all 7 questions and had PROMIS Fatigue standardized T score. At week 10 and week 22, 399 (95.4%) and 348 (83.2%) patients had PROMIS Fatigue standardized T scores, respectively. The baseline level and distribution, overall trend over time and comparison of fatigue between treatment arms showed similar results using PROMIS Fatigue standardized T score compared to that assessed via the FACIT Fatigue Scale (Supplemental Table S2, Table 2, and Figure 1B).

Psychometric properties of the PROMIS Fatigue-SF1 scale

Internal consistency (Cronbach's alpha) of the PROMIS Fatigue SF1 measure was 0.85 at baseline, 0.88 at week 10 and 0.90 at week 22 assessments. Using the FACIT-Fatigue score as the criterion, the area under the ROC curve for PROMIS Fatigue standardized T score was 0.96 at baseline, 0.94 at week 10 and 0.96 at week 22. As expected, patients with performance status (PS) of 1 had higher (worse) fatigue score than patients with PS=0, and patients with fatigue present at surgery had more severe fatigue than patients without fatigue at surgery (Supplemental Table S3). In the exploratory factor analysis, the first factor accounted for most of the variance (92.6%) in a 2-factor model, and only the first factor had

Eigenvalue>1 (3.77 vs. 0.30 for the second factor). Loadings on the first factor (0.32–0.86) were both high and similar to the loadings for a 2-factor solution (0.31–0.86). These results suggest that the PROMIS Fatigue SF1 measure is, as previously demonstrated^{14,15}, unidimensional. The PROMIS Fatigue standardized T scores and FACIT Fatigue scores were highly correlated. The Spearman's correlation coefficient between the two scores was 0.83 at baseline, 0.90 at week 10 and 0.89 at week 22 assessments. The Pearson correlation coefficient between the two scores was quite similar (0.83, 0.89 and 0.88, respectively).

DISCUSSION

We examined fatigue level in patients with unfavorable RCC using the well-validated FACIT Fatigue Scale and the more recently developed PROMIS Fatigue SF1 measure. The results show that the study patients had levels of fatigue comparable to the US general population when entering the study¹³, probably due to their clinical status at the time (disease free, locally advanced RCC with no prior radiation and systemic therapy, ECOG performance status of 0–1). With 2 cycles of treatment, fatigue got significantly worse in all patients. In patients with placebo, fatigue score worsened by 4.7 (SD=9.0) between baseline and week 10, while fatigue score worsened by 9.6 (SD=11.5) during the same time period in patients receiving sunitinib. The difference in score change between the two arms were both statistically significant and above the minimal clinically important difference^{16,17}. Given the clinical status of these patients at baseline, the timing of the assessment and the inclusion of the placebo arm, the worsening of fatigue can be attributed to sunitinib. During week 10 and week 22 (cycles 3 and 4), fatigue level remained relatively stable in the two arms. These results suggest that sunitinib could cause fatigue in study patients. For patients who received sorafenib, fatigue score worsened by 5.6 and 6.4 at week 10 and week 22 assessments, respectively. It was not statistically significant compared to placebo. These results suggest that sorafenib is unlikely causing any severe fatigue in the study patients.

In the previously published report of the trial⁸, fatigue was shown to be one of the most common grade 3 or worse adverse events associated with sunitinib (18%) reported via NCI CTCAE v4.0 (7% on sorafenib, and 3% on placebo). Fatigue (all grades 36.6%) was also reported as one of the most common adverse events in patients receiving sunitinib for their locoregional RCC at high risk for tumor recurrence after nephrectomy in another trial⁹. The incidence of grade 3 or worse fatigue (4.9% on sunitinib vs. 1.3% on placebo) was much lower though⁹. The current PRO analysis provides consistent results from the patient perspective, but it also demonstrated the substantial discordance and underreporting of fatigue by CTCAE compared to PRO.

PROMIS Fatigue SF1 measure is a short form of the PROMIS Fatigue item bank; so is the FACIT Fatigue Scale. Notably, and supported by these reports, PROMIS fatigue item bank had clinical validity across diverse conditions, including cancer¹⁸. In this study, the short form PROMIS Fatigue SF1 measure was shown to have good internal consistency reliability and criterion and construct validity. PROMIS Fatigue SF1 score is highly correlated with the FACIT-Fatigue score. The distribution, overall trend over time and comparison of fatigue between treatment arms showed quite similar results using PROMIS Fatigue standardized T

score compared to that assessed via the validated FACIT Fatigue Scale. This suggests that the two measures can be used interchangeably in oncology clinical research.

The trial had excellent compliance rate for the PRO study. The most common reason for missing fatigue data was patient not given form by staff. Very few patients had missing data due to illness. Missing data did not pose a serious problem for PRO comparisons between the treatment arms in this study. The patients included in the PRO study had similar patient characteristics compared to those who did not participate in the PRO study. The inclusion of placebo arm allowed the precise evaluation of treatment-related fatigue versus disease-related fatigue. Nevertheless, the study has limitations to note. Particularly, only two follow up time points were assessed in the study, and it is not known whether fatigue would continue to deteriorate or remain relatively stable with continuous treatment or resolve after treatment discontinuation. In addition, the time point of week 10 (at the completion of 4 weeks of sunitinib therapy and during continuous sorafenib dosing in cycle 2) was expected to allow exploration of the association between dose escalation and side effects had all or most patients had their dose escalated in cycle 3 (week 10 was right prior to the dose escalation and week 22 reflected 2 cycles of treatment at the escalated full dose of both agents if done in this case). The data, however, showed that about a quarter of patients did not have dose escalation and half of patients escalated their dose starting cycle 2. Thus, the association between dose escalation and side effects was not able to be precisely evaluated in the study. The PRO study was added after the parent trial was activated for 3 years: It is possible that physician management of fatigue may have improved at that point, making fatigue toxicity worse in the early phase of the trial, than in the current estimates in the PRO study. The incidence of fatigue reported via CTCAE, however, did not differ in patients enrolled before PRO study compared to those enrolled after PRO study (data not shown). For many reasons, the results reported here are not to be taken as a definitive comparison of sorafenib and sunitinib with regard to their overall tolerability. First, the study was designed to test each drug against placebo, not as a head-to-head comparison. Second, there are other toxicities associated with this class of therapy (e.g., hand foot syndrome; diarrhea) that can have adverse effects on patients' lives and therefore influence tolerability. Finally, the starting doses tested in this study (37.5 mg sunitinib; 400 mg sorafenib), may not be the mono-therapeutic doses used in practice. Instead, the value of this report lies in the demonstration of the measurable worsening of fatigue that occurs when treating otherwise relatively asymptomatic patients with these agents, and in the demonstration of the high degree of consistency in results whether one uses FACIT Fatigue Scale or the PROMIS Fatigue SF1 measure.

In conclusion, sunitinib was associated with significant fatigue after two cycles of treatment and fatigue continues until end of cycle 4, when it was no longer statistically significant though. The increased fatigue associated with sorafenib was modest relative to placebo at both time points and did not reach statistical significance. The FACIT Fatigue Scale and the PROMIS Fatigue SF1 measure both performed well and very similarly to each other, suggesting that either one is a valid measure for future studies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding:

This study was coordinated by the ECOG-ACRIN Cancer Research Group (Robert L. Comis, MD and Mitchell D. Schnall, MD, PhD, Group Co-Chairs) and supported by the National Cancer Institute of the National Institutes of Health under the following award numbers: CA180820, CA180794, CA189828. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government.

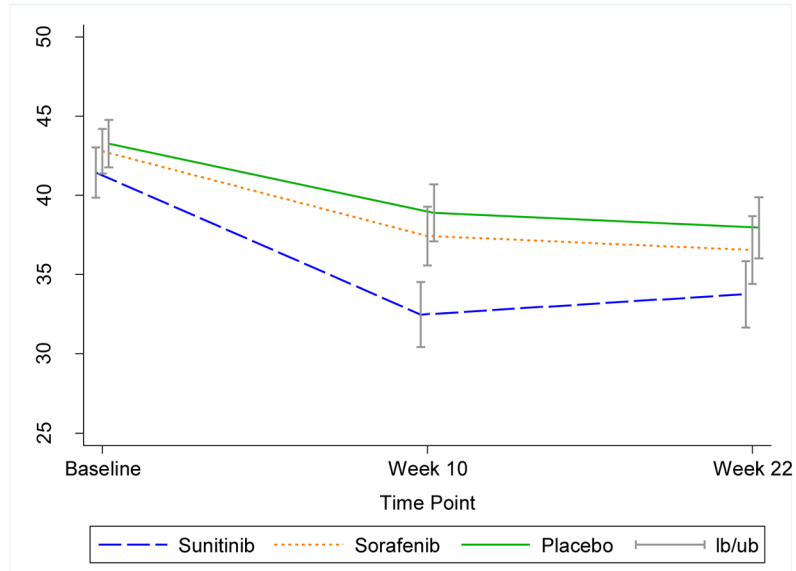
The authors thank all the patients who participated in the study.

References

- Harding G, Cella D, Robinson D Jr, Mahadevia PJ, Clark J, Revicki DA. Symptom burden among patients with renal cell carcinoma (RCC): content for a symptom index. *Health Qual Life Outcome*. 2007; 5:34.
- Hutson T, Quinn D. Cytokine therapy: a standard of care for metastatic renal cell carcinoma? *Clin Genitourin Cancer*. 2005; 4:181–186. [PubMed: 16425986]
- Malik UR, Makower DF, Wadler S. Interferon-mediated fatigue. *Cancer*. 2001; 92:1664–8. [PubMed: 11598884]
- Escudier B, Eisen T, Stadler WM, et al. Sorafenib in Advanced Clear-Cell Renal-Cell Carcinoma. *N Engl J Med*. 2007; 356:125–134. [PubMed: 17215530]
- Motzer R, Hutson T, Tomczak P, et al. Sunitinib versus interferon alfa in metastatic renal-cell carcinoma. *N Engl J Med*. 2007; 356:115–24. [PubMed: 17215529]
- Motzer RJ, Hutson TE, Cella D, et al. Pazopanib versus Sunitinib in Metastatic Renal-Cell Carcinoma. *N Engl J Med*. 2013; 369:722–731. [PubMed: 23964934]
- Escudier B, Pluzanska A, Koralewski P, et al. Bevacizumab plus interferon alfa-2a for treatment of metastatic renal cell carcinoma: a randomised, double-blind phase III trial. *The Lancet*. 2007; 370:2103–2111.
- Haas N, Manola J, Uzzo R, et al. Adjuvant sunitinib or sorafenib for high-risk, non-metastatic renal-cell carcinoma (ECOG-ACRIN E2805): a double-blind, placebo-controlled, randomised, phase 3 trial. *Lancet*. 2016; 387:2008–2016. [PubMed: 26969090]
- Ravaud A, Motzer RJ, Pandha HS, et al. Adjuvant Sunitinib in High-Risk Renal-Cell Carcinoma after Nephrectomy. *N Engl J Med*. 2016; 375:2246–2254. [PubMed: 27718781]
- Kollmannsberger C, Soulieres D, Wong R, et al. Sunitinib therapy for metastatic renal cell carcinoma: recommendations for management of side effects. *Can Urol Assoc J*. 2007; 1:S41–S54. [PubMed: 18542784]
- Chandwani KD, Zhao F, Morrow GR, et al. Lack of Patient-Clinician Concordance in Cancer Patients: Its Relation With Patient Variables. *Journal of Pain and Symptom Management*. 2017; 53:988–998. [PubMed: 28185892]
- Gravis G, Marino P, Joly F, et al. Patients' self-assessment versus investigators' evaluation in a phase III trial in non-castrate metastatic prostate cancer (GETUG-AFU 15). *Eur J Cancer*. 2014; 50:953–62. [PubMed: 24424105]
- Cella D, Lai J-S, Chang C-H, et al. Fatigue in cancer patients compared with fatigue in the general United States population. *Cancer*. 2002; 94:528–538. [PubMed: 11900238]
- Lai JS, Crane PK, Cella D. Factor analysis techniques for assessing sufficient unidimensionality of cancer related fatigue. *Qual Life Res*. 2006; 15(7):1179–90. [PubMed: 17001438]
- Lai J-S, Cella D, Yanez B, et al. Linking fatigue measures on a common reporting metric. *Journal of Pain and Symptom Management*. 2014; 48:639–648. [PubMed: 24698661]

16. Cella D, Eton DT, Lai JS, et al. 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. [PubMed: 12551804]
17. Yost KJ, Eton DT, Garcia SF, et al. Minimally important differences were estimated for six Patient-Reported Outcomes Measurement Information System-Cancer scales in advanced-stage cancer patients. *J Clin Epidemiol.* 2011; 64:507–516. [PubMed: 21447427]
18. Cella D, Lai J, Jensen S, et al. PROMIS Fatigue Item Bank had Clinical Validity across Diverse Chronic Conditions. *Journal of Clinical Epidemiology.* 2016; 73:128–134. [PubMed: 26939927]

1A: FACIT-Fatigue Score



1B: PROMIS Fatigue T score

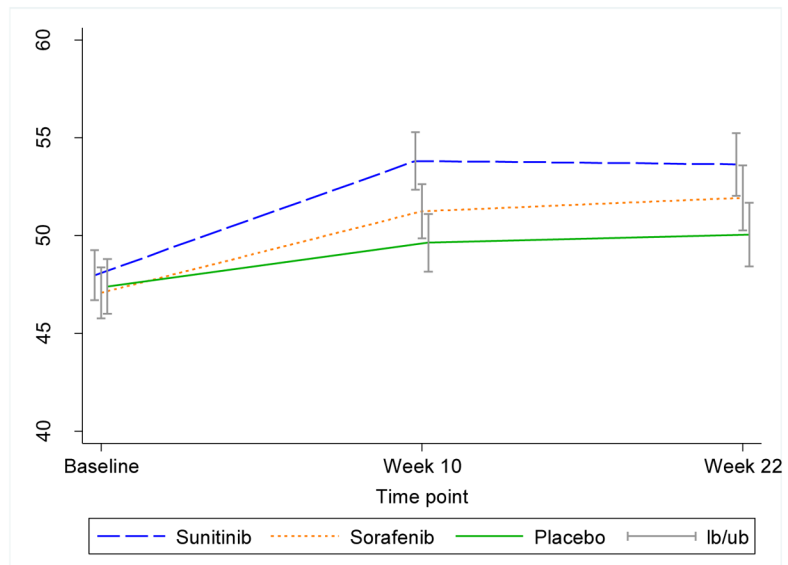


Figure 1.
Mean score and 95% CI of fatigue score by treatment arm

Table 1

Patient characteristics at study entry by treatment arm (N=321)

Variable	Sumitimb		Sorafenib		Placebo		P value
	N	%	N	%	N	%	
Age at registration (mean, SD)	53.8	10.7	54.8	10.1	56.6	9.6	0.123
Sex							0.261
Male	78	72.9	59	62.1	80	67.2	
Female	29	27.1	36	37.9	39	32.8	
Ethnicity							0.444
White	101	97.1	81	94.2	109	97.3	
Other	3	2.9	5	5.8	3	2.7	
ECOG PS							0.694
0	83	79.1	73	81.1	90	76.3	
1	22	21.0	17	18.9	28	23.7	
Fatigue present prior to surgery							0.572
No	73	79.4	67	81.7	91	85.1	
Yes	19	20.7	15	18.3	16	15.0	
Fatigue present after surgery							0.806
No	72	73.5	66	73.3	83	76.9	
Yes	26	26.5	24	26.7	25	23.2	
How was disease discovered							0.846
Incidental	37	34.6	29	30.9	40	33.6	
Symptomatic	70	65.4	65	69.2	79	66.4	
History of cardiovascular disease							0.436
No	86	81.1	73	77.7	88	74.0	
Yes	20	18.9	21	22.3	31	26.1	

Table 2

Fatigue score change between follow up and baseline

Scale	Change	Sunitinib		Sorafenib		Placebo		Sunitinib vs. Placebo			Sorafenib vs. Placebo		
		Mean	SD	Mean	SD	Mean	SD	Diff	SE	P-value	Diff	SE	P-value
FACIT	Week 10-Baseline	-9.6	11.5	-5.6	9.1	-4.7	9.0	-4.84	1.27	<0.001	-0.84	1.14	0.464
	Week 22-Baseline	-7.9	12.0	-6.4	10.3	-5.6	10.0	-2.34	1.46	0.110	-0.87	1.39	0.535
PROMIS	Week 10-Baseline	6.1	8.9	4.2	7.9	2.7	7.4	3.48	1.01	<0.001	1.49	0.98	0.130
	Week 22-Baseline	5.8	9.6	5.2	8.7	3.2	8.3	2.60	1.22	0.034	1.99	1.20	0.099