

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Gounder MM, Mahoney MR, Van Tine BA, et al. Sorafenib for advanced and refractory desmoid tumors. N Engl J Med 2018;379:2417-28. DOI: 10.1056/NEJMoa1805052

Supplementary Appendix

1. List of Investigators – Table S1: *page 2*
2. Methods:
 - a. Censoring Definitions: *page 3*
 - b. Brief Pain Inventory and PRO-CTCAE: *page(s) 4 – 11*
 - c. Evaluating MRI T2 signal and total tumor volume: *page(s) 12 - 14*
3. Figure
 - a. Figure S1: Tumor responses: *page(s) 15*
 - b. Figure S2A-C: Tumor responses: *page(s) 16 – 17*
 - c. Figure S3A-L: Graphical summary of patient-reported adverse events: *page(s) 19 – 28*
4. Figure Legend: *page 29*
5. Tables: S2 – S7 – *page(s) 30 – 41*
6. Reference: *page 42*

Table S1

Members of the Alliance (A091105) and the National Clinical Trials Network (NCTN)

Affiliations	Affiliate Centers	Study Principal Investigators	Institutional Principal Investigators	Grant Funding Numbers
ALLIANCE	Loyola University Medical Center, Maywood, IL	Meyer, Janelle	Ellen Gaynor	NA
ALLIANCE	Nevada Cancer Research Foundation NCORP, Las Vegas, NV	Grzegorz Obara	John Ellerton	UG1CA189829
ALLIANCE	Nevada Cancer Research Foundation NCORP, Las Vegas, NV	K. S. Jahangir	John Ellerton	UG1CA189829
ALLIANCE	Huntsman Cancer Institute, University of Utah, CO	Wright, Jennifer	Howard Colman	U10CA180818
ALLIANCE	University of Arkansas for Medical Sciences	Rangaswamy Govindarajan	Laura Hutchins	NA
ALLIANCE	Dayton NCI Community Oncology Research Program, Dayton, OH	Tarek Sabagh	Howard Gross	UG1CA189957
ALLIANCE	Columbia University Minority Underserved NCORP, New York, NY	Gary Schwartz	Gary Schwartz	UG1CA189960
ALLIANCE	MedStar Georgetown University Hospital, Washington, DC	Michael Pishvaian	Filipa Lynce	NA
ALLIANCE	Duke University - Duke Cancer Institute LAPS, Durham, NC	Richard Reidel	Jeffrey Crawford	U10CA180857
ALLIANCE	University of Iowa / Holden Comprehensive Cancer Center, Iowa City, IA	Mohammed Milhem	Laith Abushahin	NA
ALLIANCE	Mayo Clinic LAPS, Rochester, MN	Steven Attia	Steven Alberts	U10CA180790
ALLIANCE	University of Texas MD Anderson Cancer Center LAPS, Houston, TX	Vinod Ravi	Kelly Hunt	U10CA180858
ALLIANCE	Washington University - Siteman Cancer Center LAPS, Saint Louis, MO	Brian Van Tine	Nancy Bartlett	U10CA180833
ALLIANCE	Memorial Sloan-Kettering Cancer Center LAPS, New York, NY	Mrinal Gounder	Michael Morris	U10CA180791
ECOG-ACRIN	Froedtert and the Medical College of Wisconsin	John Charlson		NA
ECOG-ACRIN	Fox Chase Cancer Center, Philadelphia, PA	Sujana Movva		NA
ECOG-ACRIN	University of Alabama at Birmingham Cancer Center, Birmingham, AL	Robert Conry		NA
NCIC-CTG	University Health Network Princess Margaret Cancer Center, Toronto	Abha Gupta		NA
NRG	John B Amos Cancer Center, Georgia	Andrew Pippas		NA
NRG	Oklahoma Cancer Specialists & Res Institute-Tulsa	Scott Cole		NA
SWOG	Northwestern University, Chicago, IL	Mark Agulnik		NA
SWOG	Yale University, New Haven, CT	Hari Deshpande		NA
Abbreviations: ALLIANCE: Alliance Clinical Trials in Oncology Group, ECOG-ACRIN: Eastern Cooperative Oncology Group - American College of Radiology Imaging Network, NCIC-CTG: National Cancer Institute of Canada Clinical Trials Group, NRG: NSABP, RTOG, and GOG, SWOG: Southwest Oncology Group				

Statistical Methods – Censoring Definitions

1) Classifying patients relative to disease status and follow-up status:

- a. Patient disease status was evaluated using RECSIT v1.1. Patients ending treatment for symptomatic deterioration without radiographic evidence of PD, were classified as having PD (per RECIST v1.1). Otherwise, patients not yet showing disease progression were classified as having no progression at the most recent disease assessment and in the following cases: crossing over to receive sorafenib following the Alliance DSMB unblinding of the study results, date of first non-protocol directed anti-cancer therapy after ending protocol directed therapy for non-PD reasons, lost to follow-up, withdrawal of consent, and changing imaging methods from that which was used at study entry and without returning to the original imaging technique.
- b. Patient follow-up status was the most recent assessment of alive or dead at date most recently known as alive or on the date of death. Otherwise, patients considered lost to follow-up, withdrawing consent for future follow-up were classified as alive on the date of most recent follow-up.

2) Calculating time to event endpoints, based on 1 above:

- a. Duration of response for patients achieving response was calculated as the time between the date of earliest objective tumor response associated with initial treatment to date of PD during initial treatment assignment; otherwise, evaluated as non-PD on the date aforementioned.
- b. Time to response was calculated as the time between randomization and the date of first response associated with initial treatment and PD; otherwise, evaluated as continuing in response on the date aforementioned.
- c. Progression free survival (PFS) was calculated as the time between the date randomization to the date of PD to associated with initial treatment or death; otherwise, evaluated as alive on the date aforementioned.
- d. Overall Survival (OS) was calculated as the time between the date of randomization and death, including the follow-up obtained during crossover treatment; otherwise, evaluated as alive on the date aforementioned.

Brief Pain Inventory Short Form (3 questions), PRO-CTCAE (19 questions) and LASA (1 question)

Please answer the following questions about your symptoms and quality of life:

(Brief Pain Inventory Short Form questions)

1. Please rate your pain by circling the one number that best describes your pain at its WORST in the last 24 hours.

0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain As Bad As You Can Imagine

2. Circle the one number that describes how, during the past 24 hours, pain has interfered with your:

A. General Activity

0	1	2	3	4	5	6	7	8	9	10
										Completely Interferes

B. Sleep

0	1	2	3	4	5	6	7	8	9	10
										Completely interferes

PRO-CTCAE questions

In the past 7 days...

3. What was the SEVERITY of your **insomnia (including difficulty falling asleep, staying asleep, or waking up early)** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

In the past 7 days...

4. How much did **insomnia (including difficulty falling asleep, staying asleep, or waking up early)** INTERFERE with your usual or daily activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Somewhat
- ☐ Quite a bit
- ☐ Very much

5. What was the SEVERITY of your **constipation** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

6. How OFTEN did you have **pain**?

- ☐ Never
- ☐ Rarely
- ☐ Occasionally
- ☐ Frequently
- ☐ Almost Constantly

7. What was the SEVERITY of your **pain** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

8. How much did **pain** INTERFERE with your usual or daily activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Somewhat
- ☐ Quite a bit
- ☐ Very much

In the past 7 days...

9. What was the SEVERITY of your **fatigue, tiredness, or lack of energy** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

10. How much did **fatigue, tiredness, or lack of energy** INTERFERE with your usual or daily activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Somewhat
- ☐ Quite a bit
- ☐ Very much

11. How OFTEN did you have **nausea**?

- ☐ Never
- ☐ Rarely
- ☐ Occasionally
- ☐ Frequently
- ☐ Almost Constantly

12. What was the SEVERITY of your **nausea** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

13. How OFTEN did you have **vomiting**?

- ☐ Never
- ☐ Rarely
- ☐ Occasionally
- ☐ Frequently

☐ Almost Constantly

In the past 7 days...

14. What was the SEVERITY of your **vomiting** at its WORST?

☐ None

☐ Mild

☐ Moderate

☐ Severe

☐ Very Severe

15. How OFTEN did you have **loose or watery stools** (diarrhea)?

☐ Never

☐ Rarely

☐ Occasionally

☐ Frequently

☐ Almost Constantly

16. Did you have any **rash**?

☐ No

☐ Yes

17. What was the SEVERITY of your **hand-foot syndrome (a rash of the hands or feet that can cause cracking, peeling, redness, or pain)** at its WORST?

☐ None

☐ Mild

☐ Moderate

☐ Severe

☐ Very Severe

18. In the past 7 days, how much did **hand-foot syndrome (a rash of the hands or feet that can cause cracking, peeling, redness, or pain)** INTERFERE with your usual or daily activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Somewhat
- ☐ Quite a bit
- ☐ Very much

In the past 7 days...

19. What was the SEVERITY of your **decreased appetite** at its WORST?

- ☐ None
- ☐ Mild
- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

20. How much did **decreased appetite** INTERFERE with your usual or daily activities?

- ☐ Not at all
- ☐ A little bit
- ☐ Somewhat
- ☐ Quite a bit
- ☐ Very much

21. What was the SEVERITY of your **mouth or throat sores** at their WORST?

- ☐ None
- ☐ Mild

- ☐ Moderate
- ☐ Severe
- ☐ Very Severe

(LASA Overall Quality of Life Item)

Please check the number (0-10) best reflecting your response to the following that describes your feelings during the past week, including today.

How would you describe:

22. Your overall Quality of Life?

☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10

As bad
as it
can be

As good as it can
be.

Pain Medication Diary

Subject Name: _____

To be completed by study site

Subject ID: ____ Assessment Period: ____ Research Staff Signature: _____ Date: _____

Study Site Staff: Prior to each 7-day period, write in pain medication names, routes, and strengths in the columns below.

- List only one type of medication on each row. Include all medications the patient is taking for pain or that may alleviate pain symptoms
- Please write in the dates where indicated (Month/Day/Year).

Study Participant: Please record the number of units you have taken of each medication over the prior 24 hours. If you are wearing a patch, mark a “1” for every day the patch is worn.

During this 7-day period, if you start a new medication for pain or change the strength of the pills, write the new information on a new row. (Do not change the information about the medication you have already taken.)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Date:

____/____/____
(mm/dd/yy)

Medication Name
(write in name)

Route
(check one box)

Strength
(example:50 mg)

Taken

Taken

Taken

Taken

Taken

Taken

Taken

☐ Oral

☐ Patch

☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

- ☐ Oral
- ☐ Patch
- ☐ Other

Evaluating MRI T2 signal and total tumor volume

Introduction:

Although reproducible radiographic measurements by existing standards, such as RECIST 1.1, are critical to compare response rates with historical trials and other key end points, it is abundantly clear that it is challenging to reproducibly measure tumors that are asymmetric and involve multiple layers of fascia, neurovascular bundles, and complex joint spaces¹⁻³. Additionally, it is clear that the change in the maximal unidimensional measurement may not capture the actual change in tumor burden³. The optimal imaging criteria to evaluate DT/DF remain undefined and therefore it is critical to develop and validate appropriate imaging criteria that can accurately assess the clinical efficacy of drugs. Given their infiltrative nature, tumors such as desmoids do not necessarily shrink equally in three dimensions, but may shrink unevenly; a feature inadequately assessed by RECIST 1.1. Studies involving tenosynovial giant-cell tumors, also a locally infiltrative connective tissue tumor, have shown tumor volume to be a superior measure of treatment response than RECIST 1.1³. In a retrospective study of sorafenib in desmoid tumors, patients with a decrease in MRI T2 signal (30% or greater) reported symptomatic improvement without achieving a partial or complete response by RECIST 1.1.⁴ Studies have shown that areas of desmoid tumor that are hyperintense on T2-weighted images and short T1 inversion recovery (STIR) images are associated with active fibroblastic proliferation while area that are iso- or hypo-intense are collagenous⁵. In this prospective trial, we conducted an exploratory study to test whether tumor volume and MRI T2 changes are early predictors of RECIST response and whether it correlates with improvement in symptoms.

Methods: All patients undergoing MRI as part of their response evaluation were included in this exploratory study if they had at least 3 sequential images (baseline and at least 2 on-treatment scans). Protocol defined, range of MRI parameters for image acquisition included use of 1.5 T or 3T scanners with TR = 3700-6600ms, TE = 60-78 ms, echo-train length = 16-24, number of slices = 30, 4-5 mm thickness, FOV = 240-320 mm, matrix size = 320-256x192. Once the image acquisition was completed (at baseline and every 8 weeks), the scans underwent de-identification of all patient identifiers (except study ID number and protocol number) and the entire data sets in digital DICOM format, along with Alliance Adjunctive Data Form (if applicable) and Alliance Image Measurement Form (if

applicable), were electronically submitted to the ALLIANCE Imaging Core Laboratory. Following quality assurance review by Imaging Core Lab, the data was then securely sent to Memorial Sloan Kettering Cancer Center for exploratory studies. Response evaluation by RECIST was performed by local study investigators and these measurements were obtained from study database.

T2-weighted images (with or without fat suppression) of the target lesion were manually segmented on each slice by two radiologists in consensus (R.Y and R.L) using a free open-source software package (ImageJ, version 1.48; <https://imagej.nih.gov/ij/>). Lesions were segmented by using the inner border of the lesion to minimize partial volume effects; T1-weighted images were referred to for clarification when lesion margin was unclear. When adequate fat saturation could not be achieved due to the location of the tumor, metal artifacts or patient positioning, the patient was excluded from T2-signal analysis portion of this study. Short T1 Inversion Recovery (STIR) was performed but did not replace T2 with fat saturation in the data analysis. The areas of the segmented lesion were used to calculate the volume of interest for computer-based image analysis by summing areas of the tumor on each slice and multiplying by slice thickness to obtain total tumor volume. Oval region of interests (ROIs) were also placed in the adjacent muscles to provide references. The ratio to the mean tumor signal intensity to the reference muscle signal was calculated to generate normalized T2 SI ratio. Tumor volume and T2 SI ratio relative to adjacent muscle were computed, and percent changes relative to baseline were calculated. As previously described, the ratio between the T2 signal intensity of the tumor (C1) and adjacent skeletal muscle (M1) was calculated between two consecutive scans (i.e. $[C2/M2] / [C1/M1]$). The percent change in longest dimension (RECIST 1.1), tumor volume, and T2SI ratio relative to baseline (or preceding scan) were calculated and plotted (Y-axis) with the time (cycles) on treatment (X-axis).

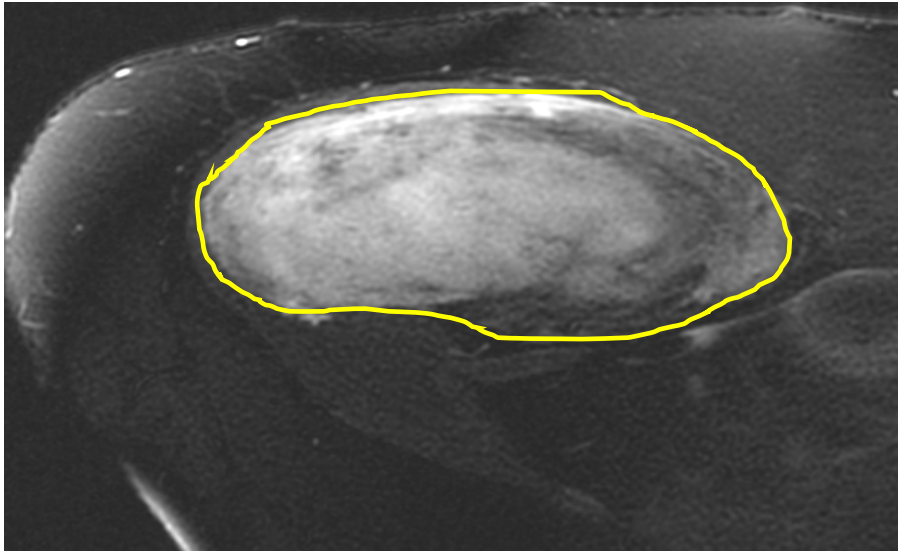
In this pilot analysis, all patients who were treated at Memorial Sloan Kettering Cancer and had at least 3 scans were included in the analysis. Patients who were initially on placebo and then crossover to open-label sorafenib had a second baseline at time of crossover. The radiologist and medical physicist were blinded to all patient demographics, treatment assignments and clinical outcomes and read each scan independent and blinded to the results of subsequent time points. The best percentage change in RECIST, tumor volume

and T2 signal were calculated. Patients initially on placebo who then crossed over to open-label placebo were included in both the placebo arm (prior to crossover) and in the sorafenib arm (after crossover). A new baseline was used at the time of crossover.

Results: At the time of writing this manuscript (May 2018), the Alliance Imaging Core Lab received a total of 498 MRI scans of 55 patients imaged exclusively by MRI scans. Patients who were imaged by CT scans were not included ($n = 22$). Images on remaining patients are pending receipt from individual sites. For the MSKCC exploratory analysis, a total of 12 patients were enrolled. All patients ($n = 12$) were evaluated by MRI but only eleven patients had a baseline scan and at least 2 follow-up scans. A total of 167 MRI scans were analyzed for MRI T2 signal and total tumor volume.

Figure S1

Baseline



Partial response

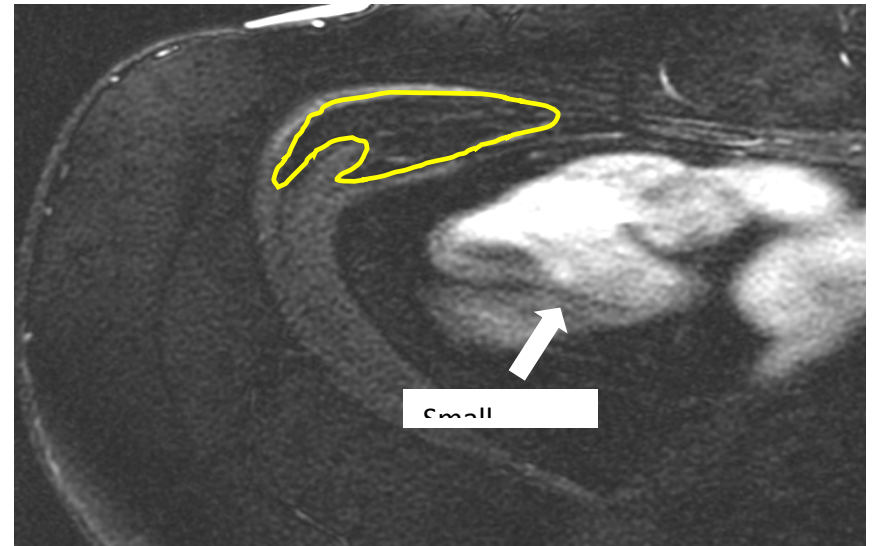
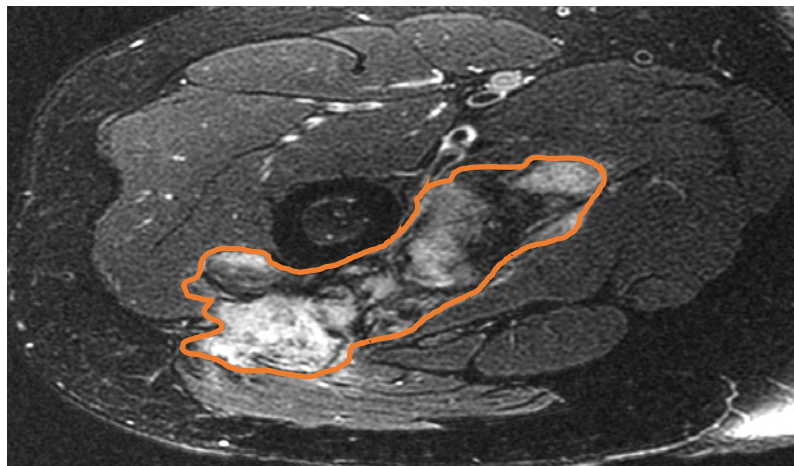


Figure S2A

Baseline (pain)



Best response (pain significantly improved)

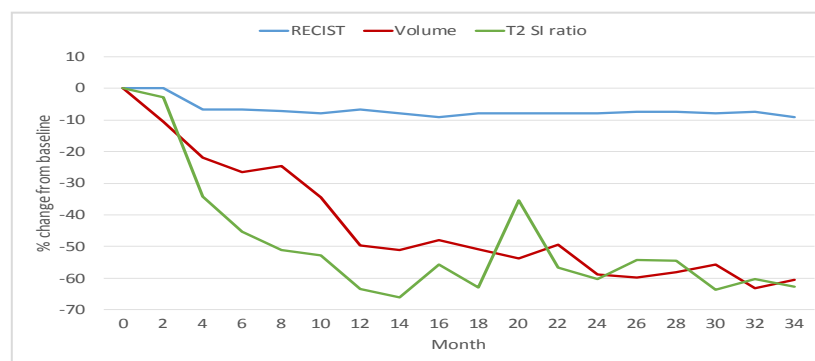
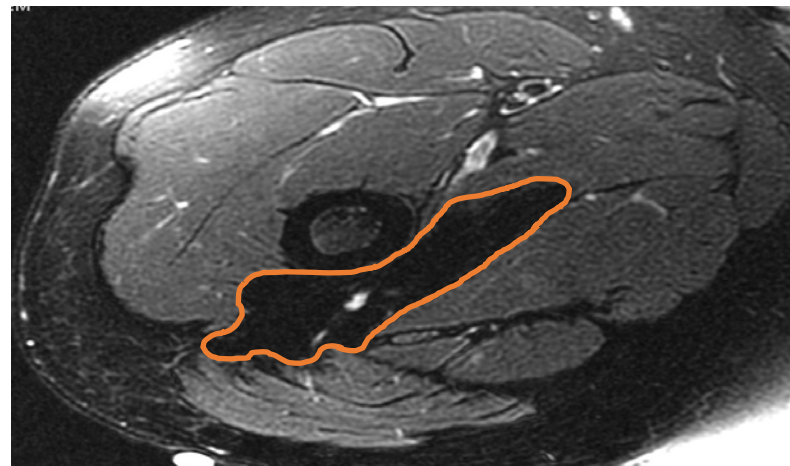
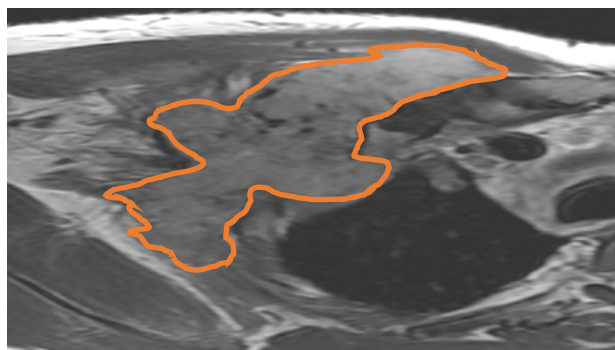
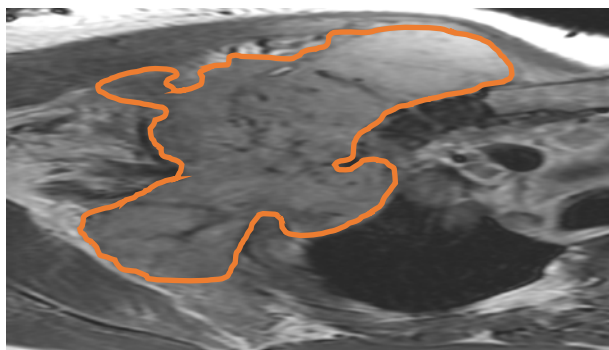


Figure S2B

Baseline
(pain, neck stiffness, decreased mobility)



Pre-crossover



Best response
(resolution of symptoms)

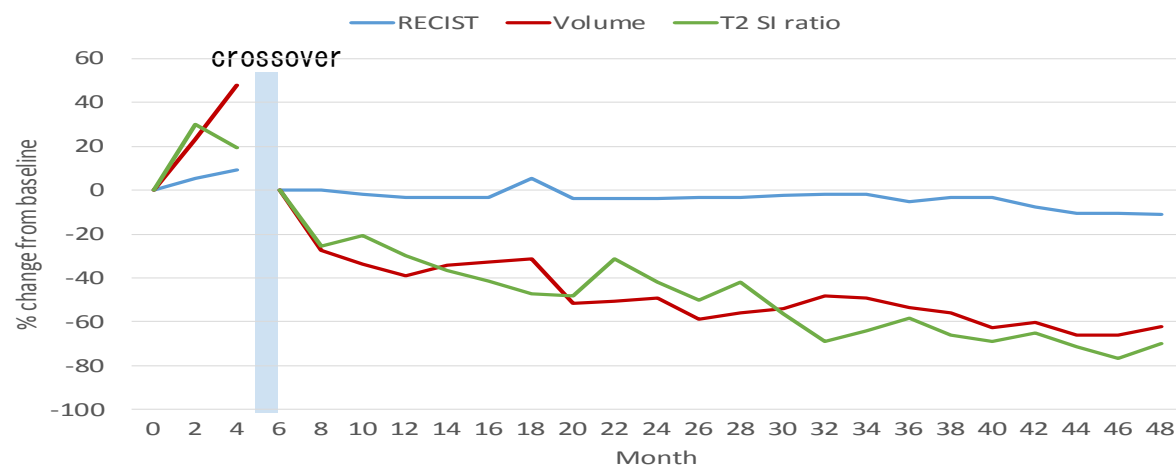
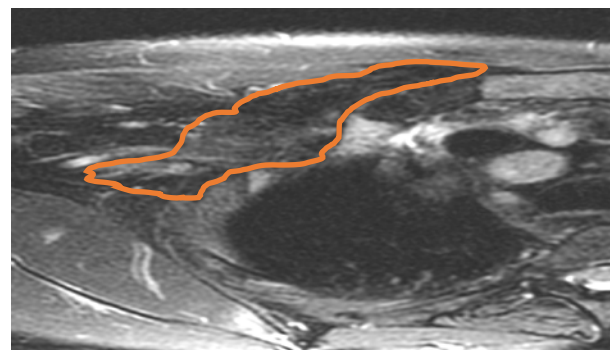


Figure S2C

Baseline (pain: difficulty with walking)

Best response (walking without pain)

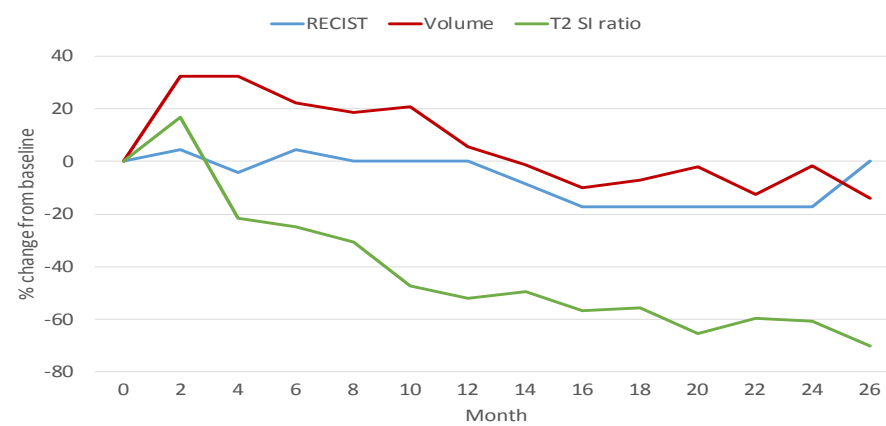
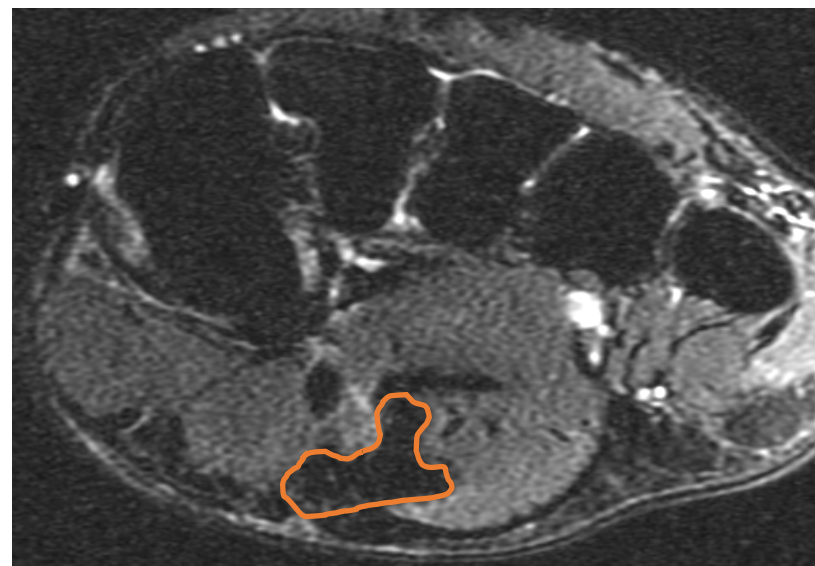
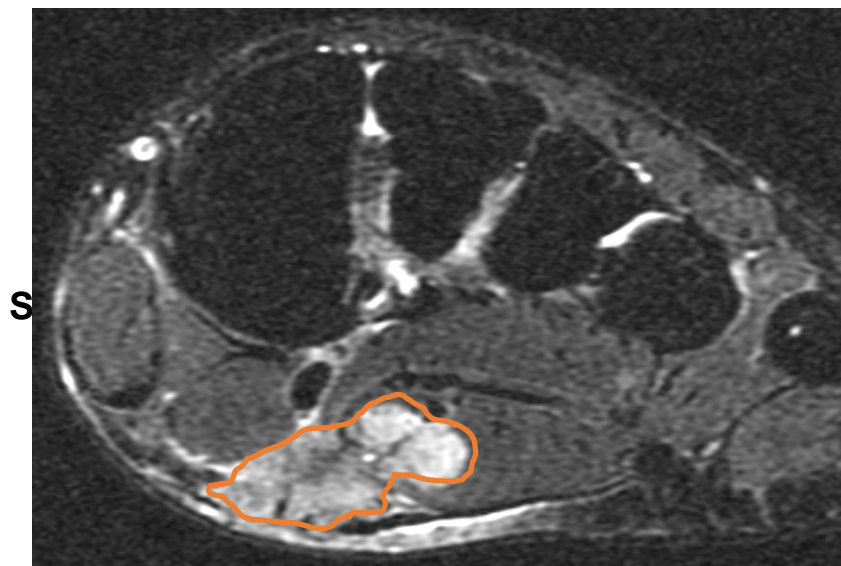
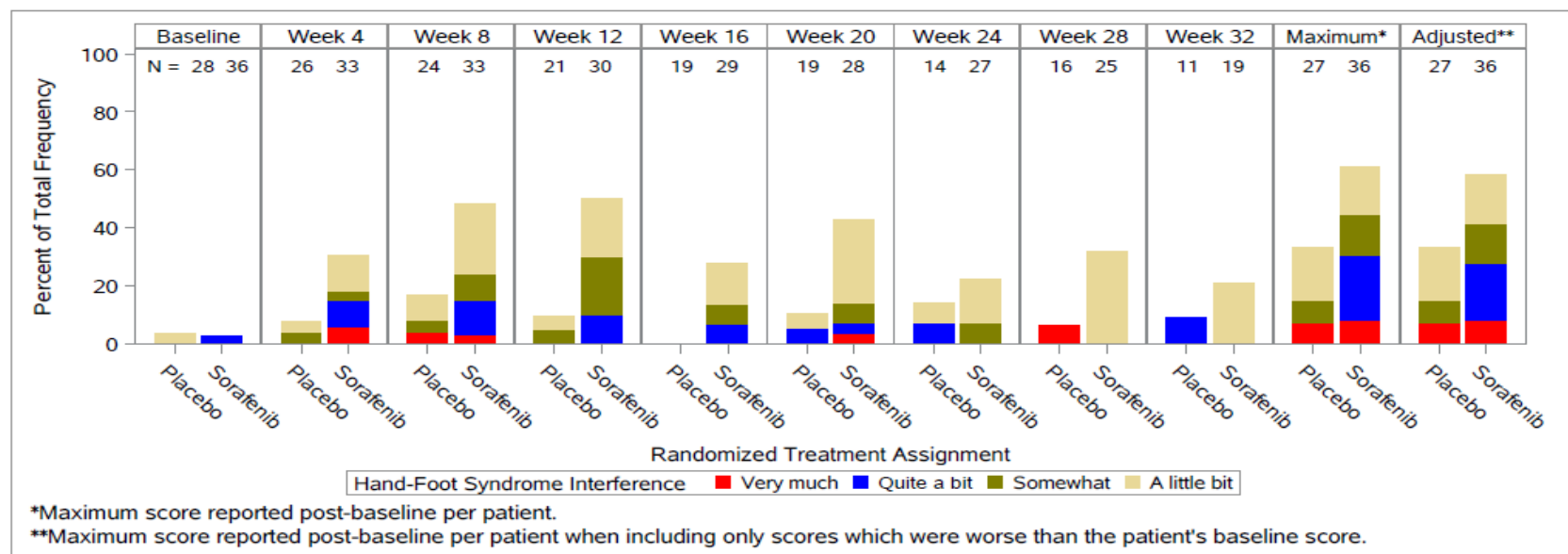
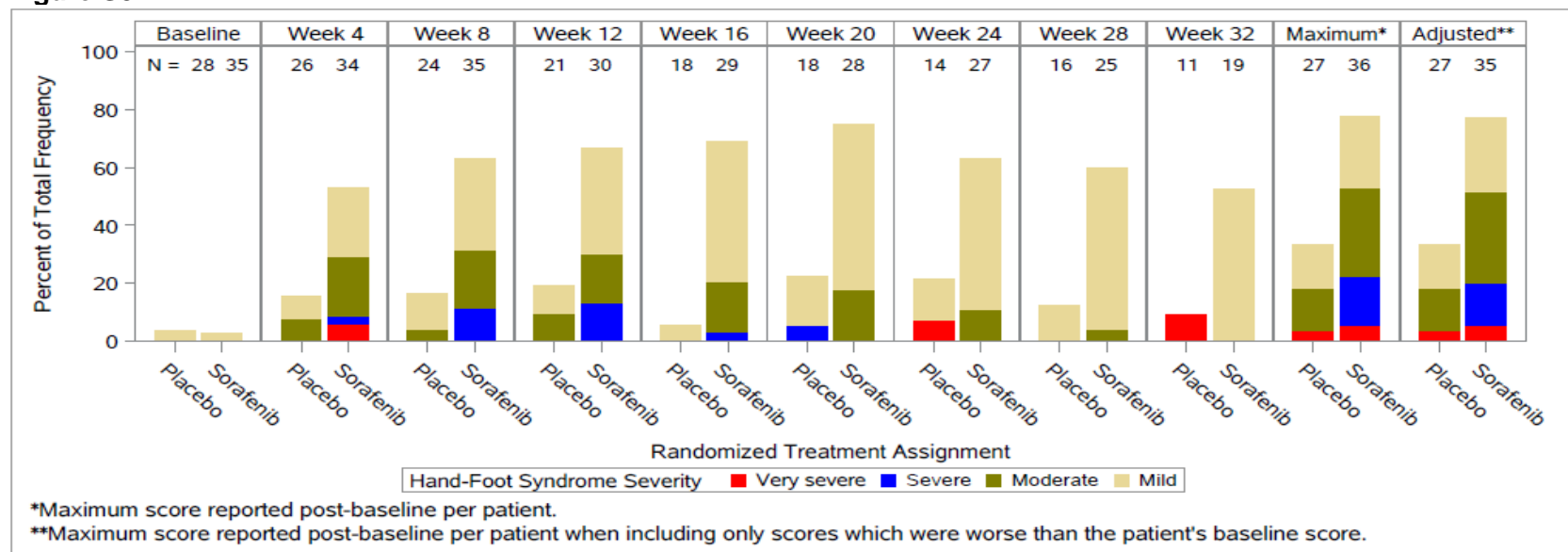
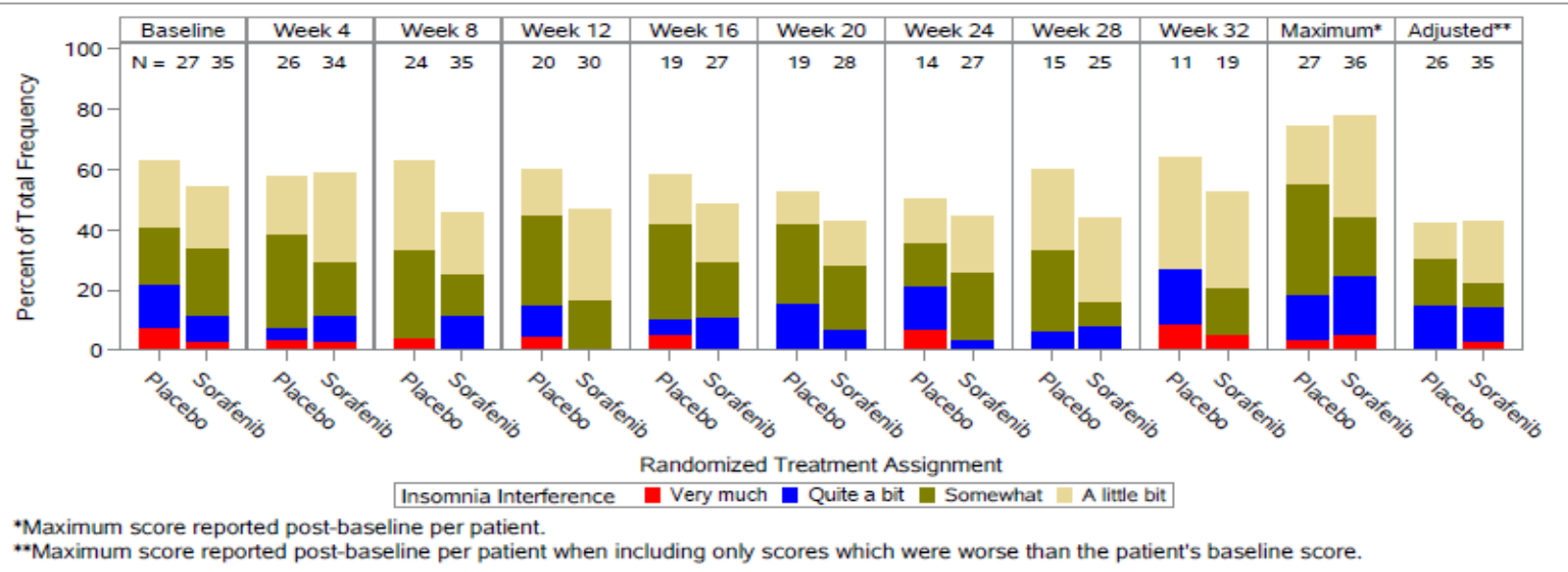
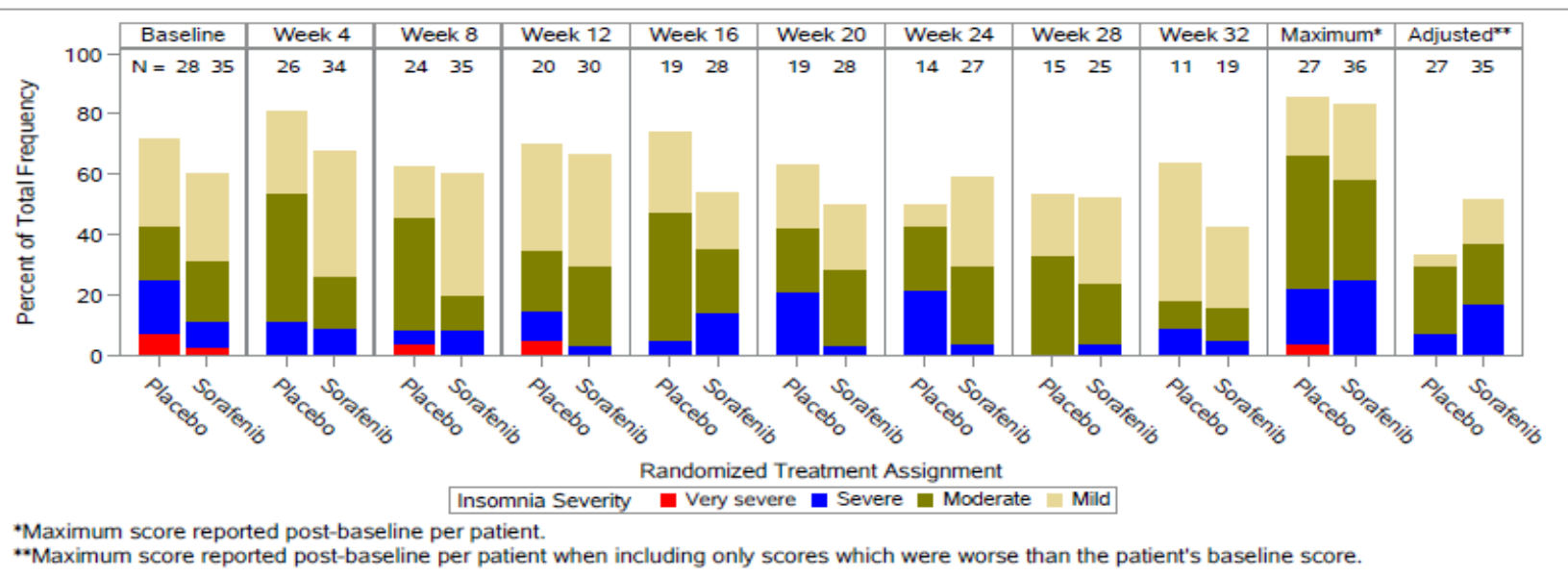
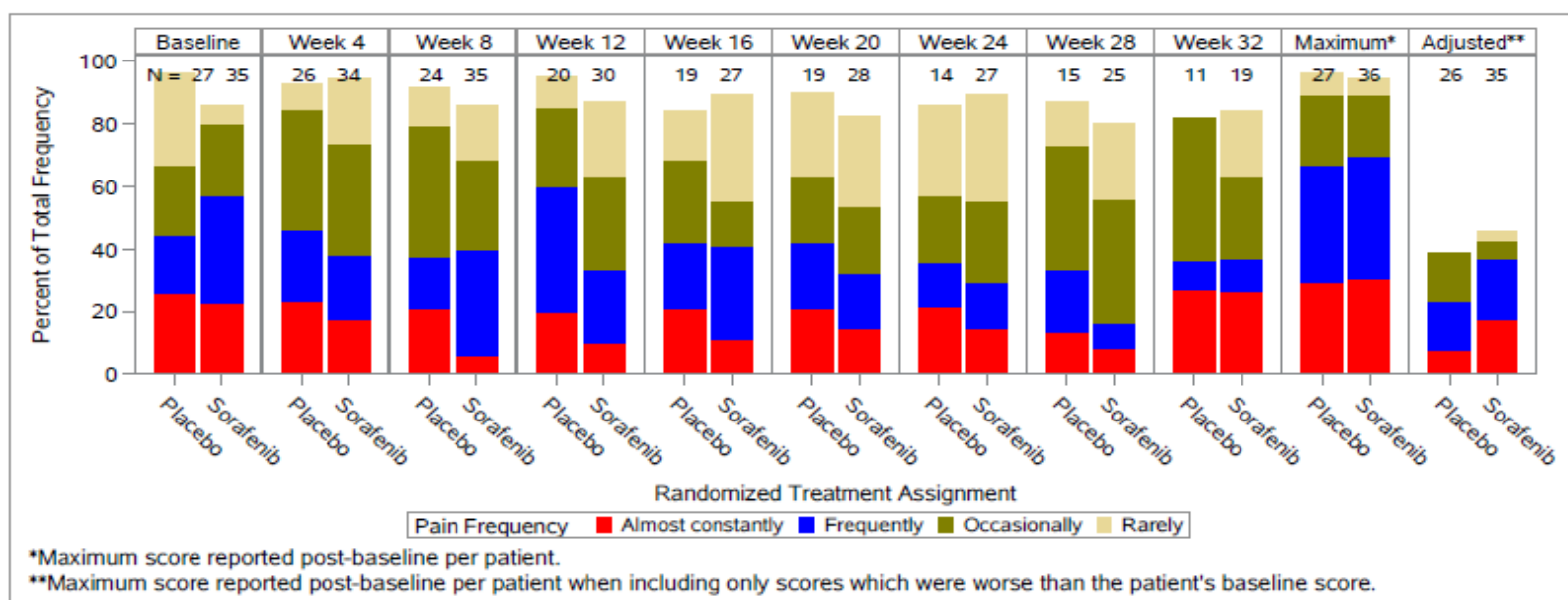
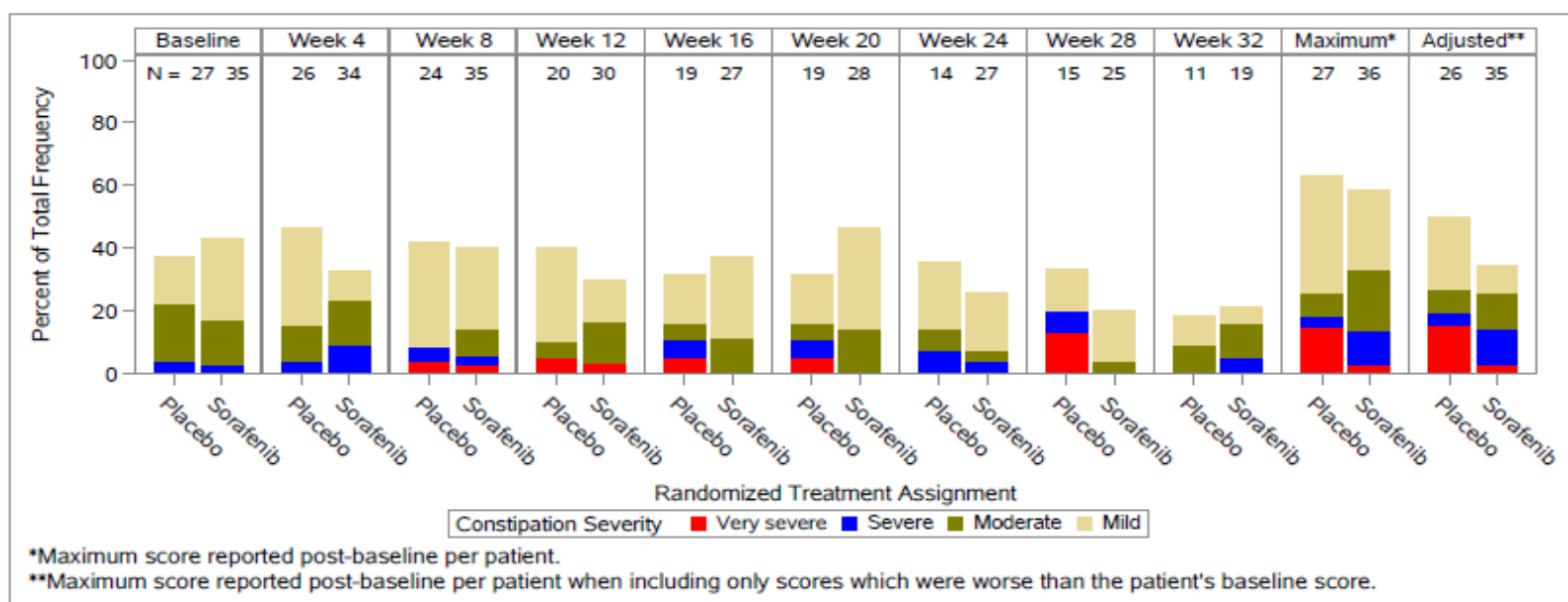
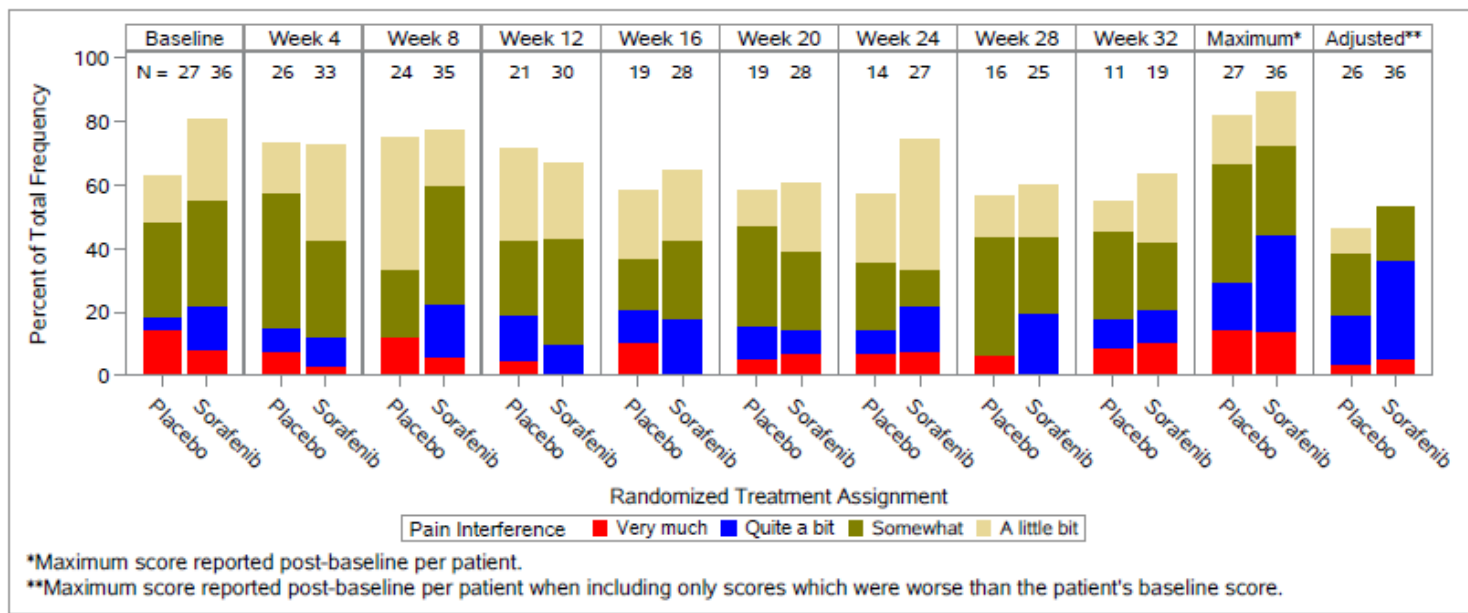
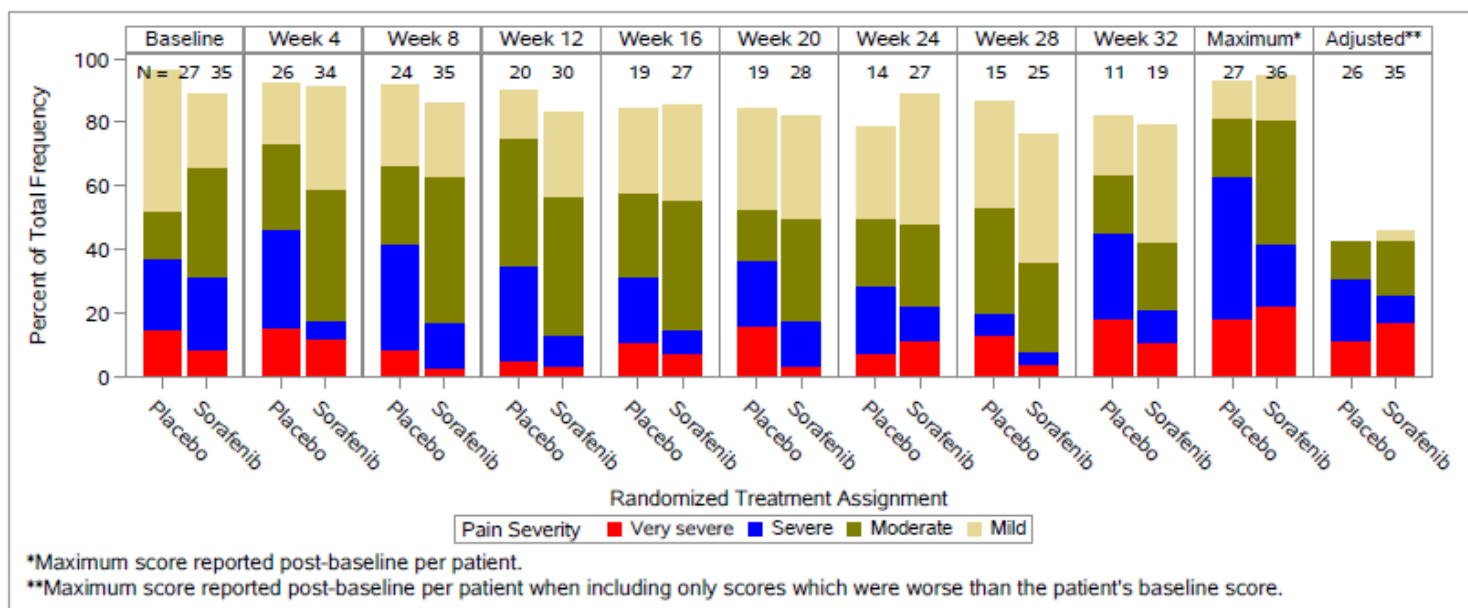


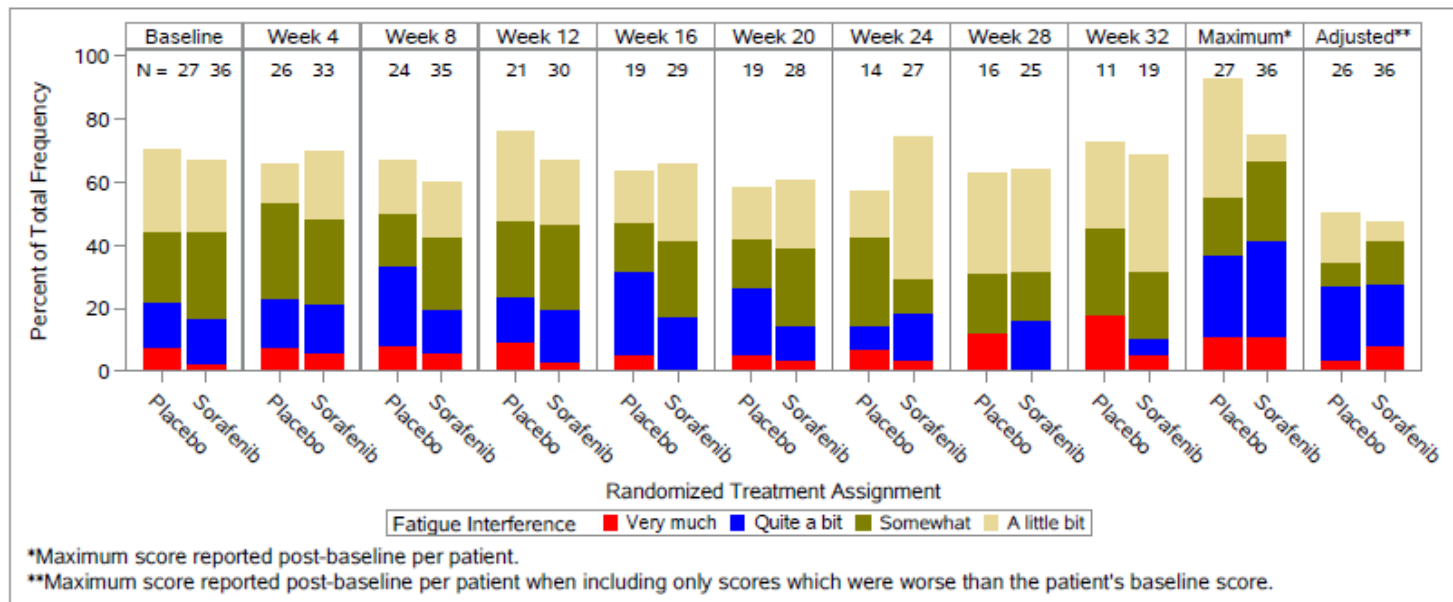
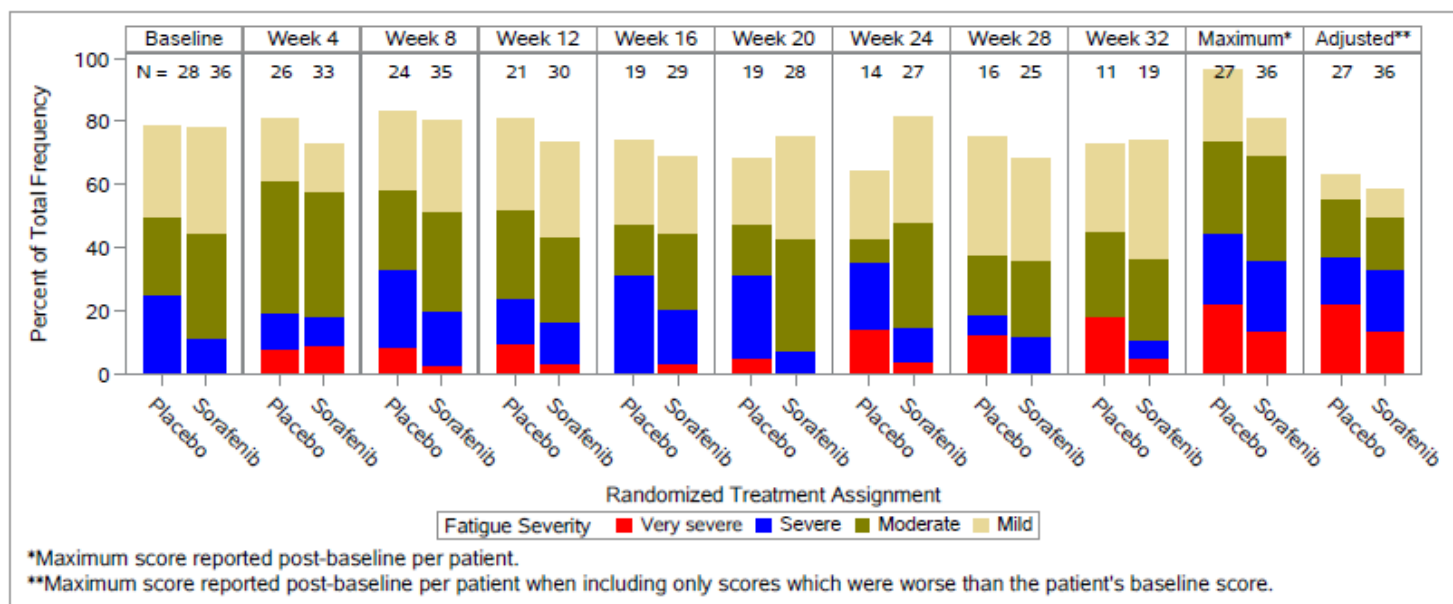
Figure S3A-L

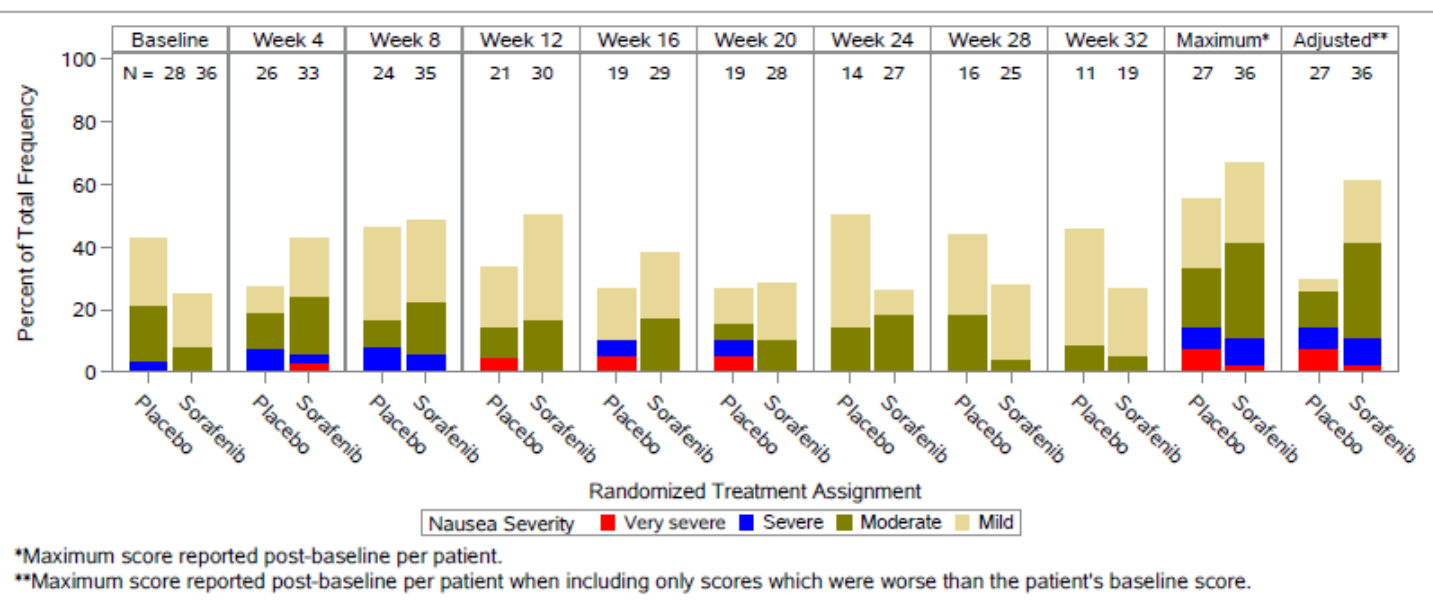
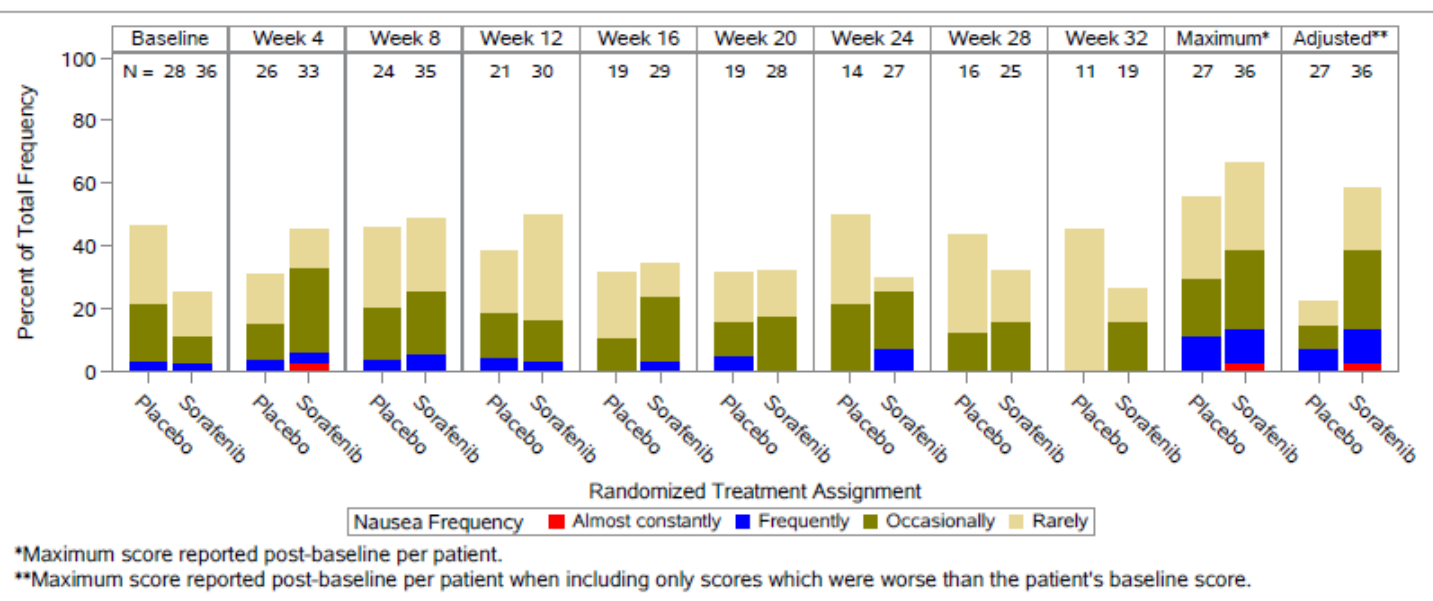


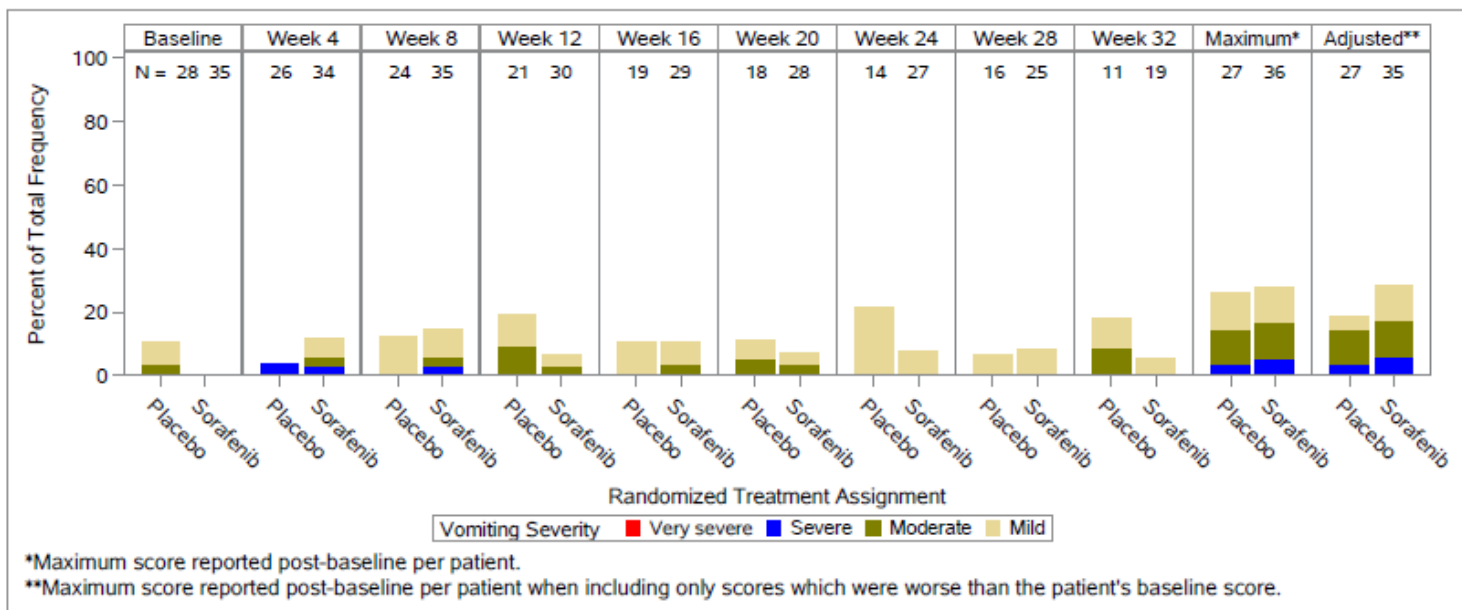
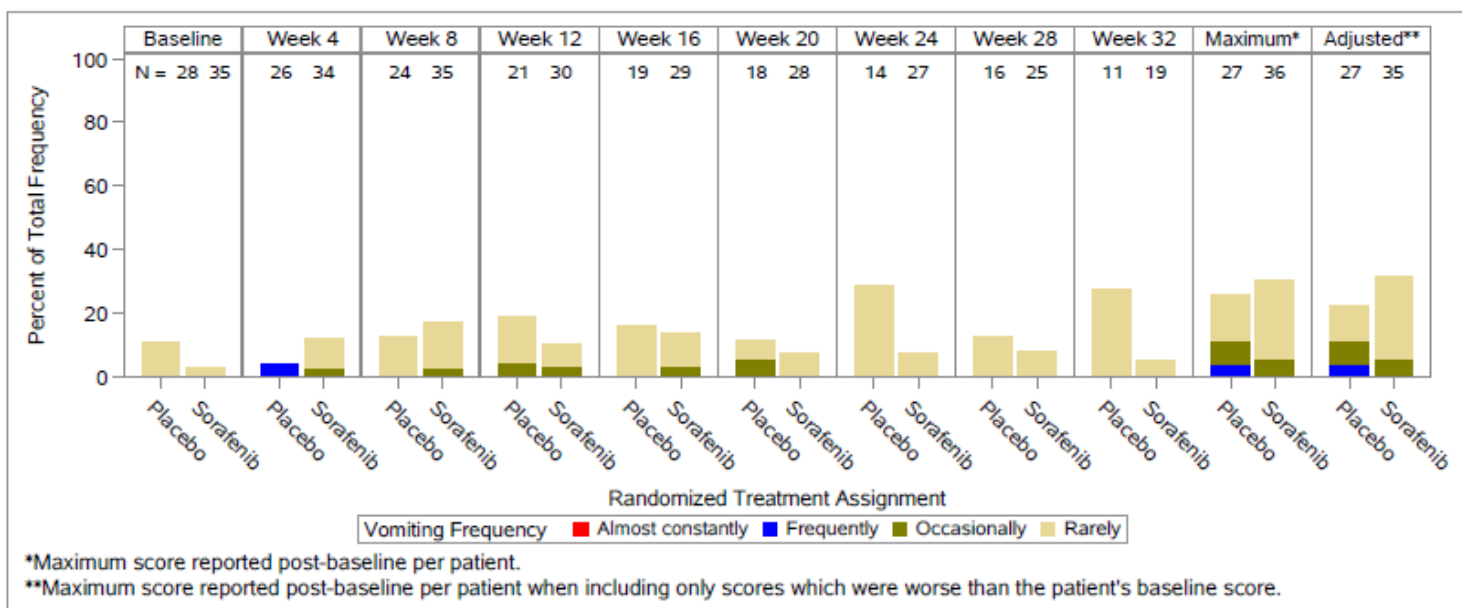


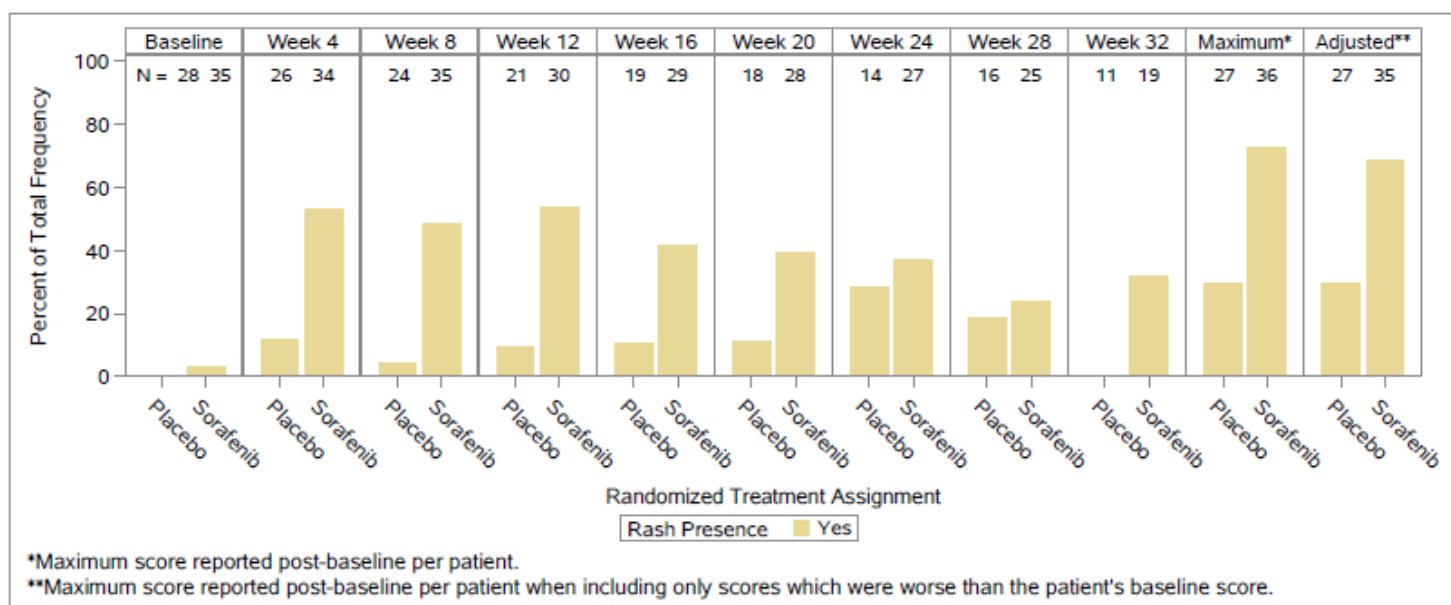
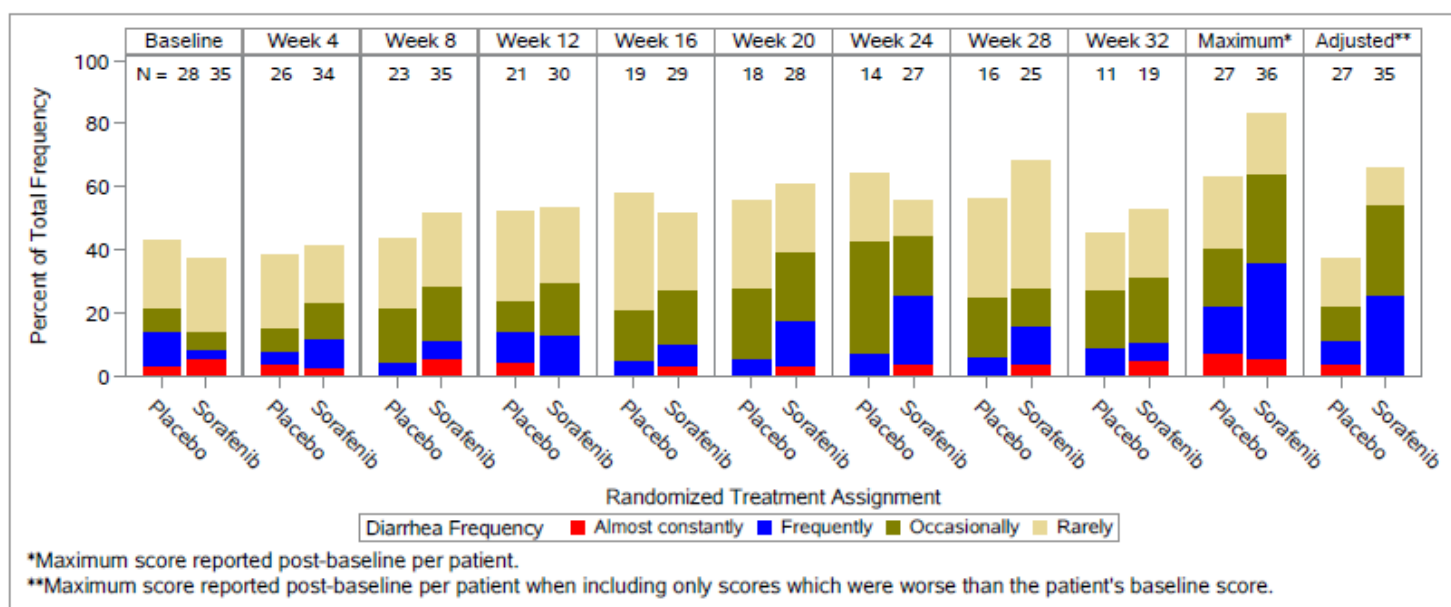


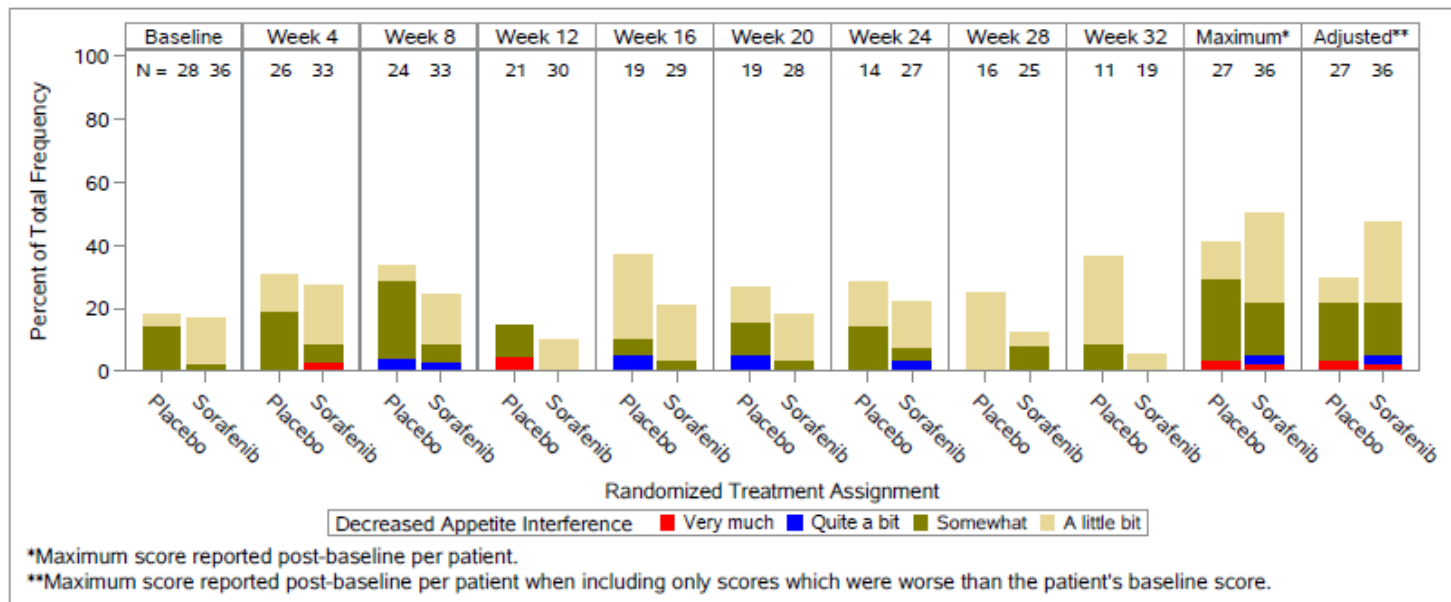
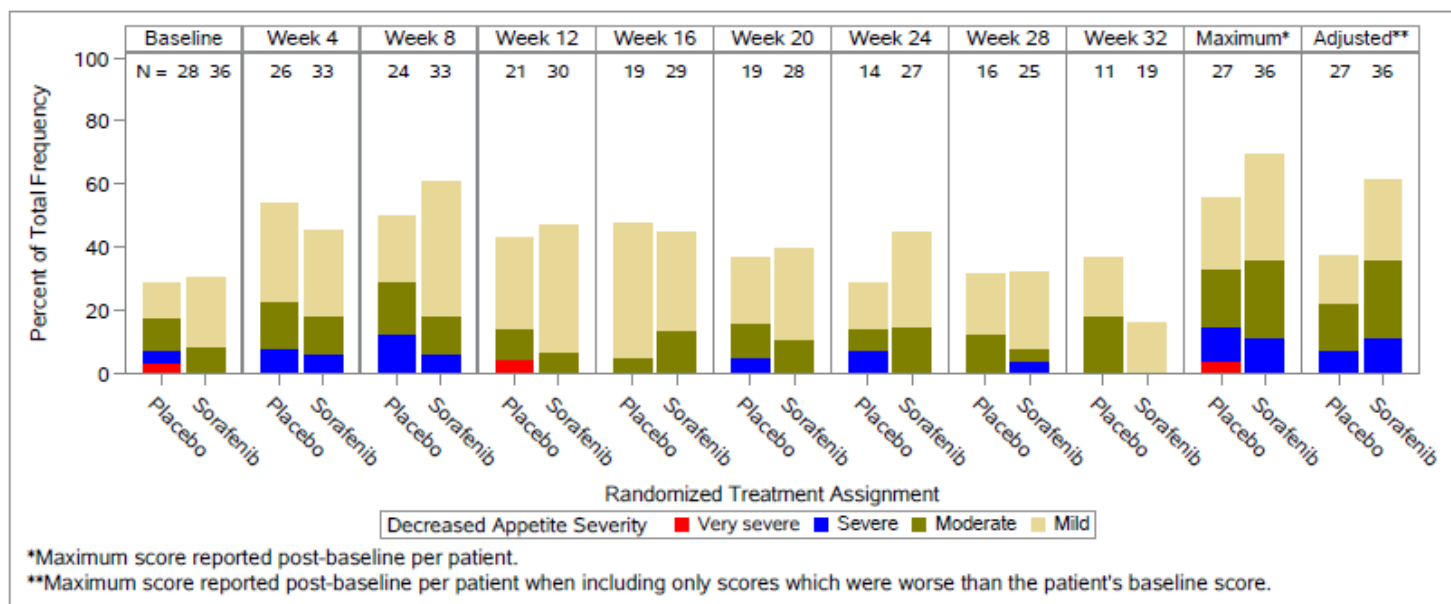












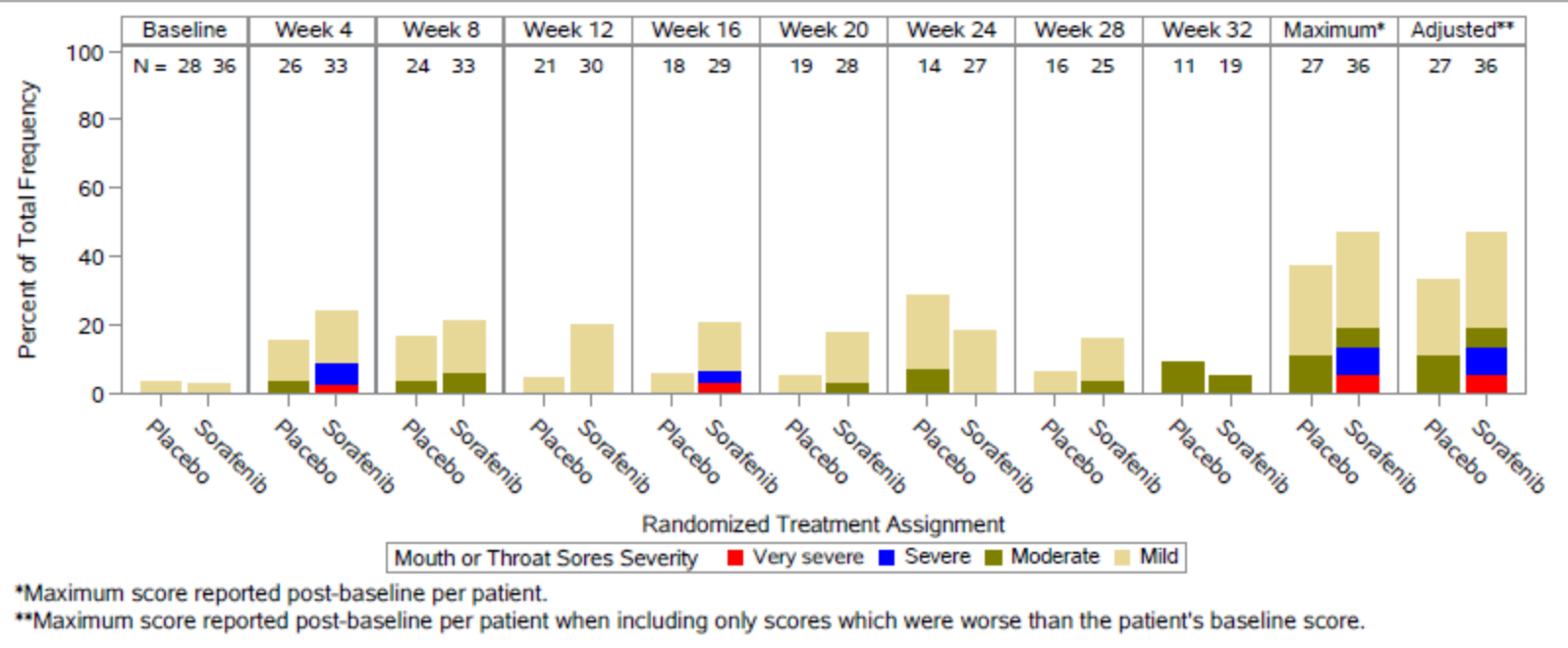


Figure S1: Radiographic images of a 30-year-old woman who presented with a symptomatic (pain and discomfort) desmoid tumor involving the right external oblique, internal oblique and transversalis muscle. The patient was randomized to sorafenib and had a RECIST partial response and complete resolution of pain. Treatment effect is also seen on MRI T2-weighted images where the tumor is hyperintense at baseline and following sorafenib there is a marked decrease in signal intensity consistent with a collagenous transformation (scar). Arrow depicts small bowel.

Figure S2A: The patient has stable disease by RECIST while on placebo. The patient experienced clinical progression without RECIST progression, but marked by an increase in tumor volume and T2W signal intensity. Following crossover to open-label sorafenib, the patient had improvement in symptoms and RECIST (overall stable disease) but a significant decrease in tumor volume. **Figure S2B:** A patient with an asymptomatic flank mass is randomized to placebo and continues to have stable disease per RECIST (+5.6%) for approximately 10 months when he notes new mild flank discomfort. The patient is unblinded for clinical deterioration and begins open-label sorafenib. Best response on treatment is stable disease per RECIST (-1.7%). The tumor undergoes a shift in T2WSI from bright (cellular) to dark (collagen) signaling treatment effect and improvement in symptoms. **Figure S2C:** A patient with an asymptomatic mass is randomized to placebo and has a spontaneous regression with RECIST partial response (-52%). Both total tumor volume and T2WSI decrease and demonstrate change from a cellular (bright) to collagenous mass.

Figure S3: Summary of patient-reported adverse events and their impact on daily living across 8 cycles. Scores are calculated for each patient per cycle. Severity of hand-foot syndrome are: 0 (none), 1 (mild), 2 (moderate), 3 (severe) or 4 (very severe) symptoms; the impact of these symptoms on daily activity are: 0 (not at all), 1 (a little bit), 2 (somewhat), 3 (quite a bit) and 4 (very much). **Figure S3A:** Compared to placebo, a significantly higher number of patients on sorafenib reported mild/moderate hand-foot syndromes (77% versus 33%, risk difference 44% CI:16.4 – 64.7%), however this did not translate to a statistically significantly increased interference in activities of daily living in the sorafenib group (58% versus 33%, CI: -0.9 – 48.3%) (**Figure S3B**). S, F and I denotes severity, frequency and interference with daily activity. To account for symptoms reported by patients at baseline, a PRO-CTCAE score during treatment was included in analysis if it was worse than the patient's baseline score or was otherwise considered a score of zero. **Figure S3C** – Insomnia, **Figure S3D** – Constipation, **Figure S3E-G:** Pain (F, S, I), **Figure S3H:** Fatigue (I), **Figure S3I:** Diarrhea (F), **Figure S3J:** Rash, **Figure S3K:** Decreased Appetite (S, I), **Figure S3L:** Mouth and Throat sore (S)

Table S2. Incidence of Treatment-Related Adverse Events (Initially Assigned Treatment) with either an incidence of $\geq 10\%$ or any grade 3 or higher. ☆

Event	Sorafenib (N=49)		Placebo (N=36)	
	Grade 1-2	Grade 3-4	Grade 1-2	Grade 3-4
	<i>number of patients (percent)</i>			
Treatment-related adverse events	33 (67)	16 (33)	26 (72)	5 (14)
Treatment-related adverse events with an incidence of $\geq 10\%$				
Palmar-plantar erythrodysesthesia syndrome	34 (69)	1 (2)	8 (22)	0
Rash - Papulopustular	22 (45)	6 (12)	3 (8)	0
Rash - maculo-papular	7 (14)	0	1 (3)	0
Rash - acneiform	6 (12)	0	0	0
Skin and subcutaneous tissue disorders - Other	6 (12)	1 (2)	2 (6)	0
Any Rash/Skin	37 (76)	7 (14)	12 (33)	0
Fatigue	30 (61)	3 (6)	17 (47)	1 (3)
Hypertension	27 (55)	4 (8)	10 (28)	0
Diarrhea	23 (47)	0	10 (28)	0
Alopecia	18 (37)	0	3 (8)	0
Nausea	19 (39)	0	13 (36)	0
Mucositis oral	11 (22)	0	6 (17)	0
Anorexia	14 (29)	0	6 (17)	0
Arthralgia	15 (31)	0	3 (8)	0
Abdominal pain	9 (18)	0	2 (6)	1 (3)
Anemia	7 (14)	1 (2)	2 (6)	1 (3)
Constipation	7 (14)	0	4 (11)	0
Vomiting	7 (14)	1 (2)	6 (17)	0
Alanine aminotransferase increased	7 (14)	0	4 (11)	0

Myalgia	7 (14)	0	4 (11)	0
Platelet count decreased	6 (12)	2 (4)	1 (3)	0
Aspartate aminotransferase increased	5 (10)	1 (2)	3 (8)	0
Blood bilirubin increased	5 (10)	0	2 (6)	1 (3)
Neutrophil count decreased	5 (10)	0	2 (6)	0
Dry skin	5 (10)	0	1 (3)	0
White blood cell decreased	3 (6)	0	6 (17)	0
Pain	2 (4)	0	1 (3)	1 (3)
Eye disorders - Other, specify	1 (2)	1 (2)	0	0
Lymphocyte count decreased	1 (2)	1 (2)	2 (6)	0
Headache	1 (2)	0	6 (17)	0
Anxiety	1 (2)	0	0	1 (3)
Blood and lymph sys disorders - Other	0	1 (2)	0	0
Hyponatremia	0	0	0	1 (3)

★ Adverse Events on initial treatment prior to crossover: grades 1 and 2 pooled, all grades 3 – 4. All 85 were assessed for safety. Included are events that were reported in at least 10% of the patients in either group, or if a grade 3 – 4 event occurred. Grade 5 adverse events of any cause occurred in 1 patient (2.0%) who received sorafenib (gastric perforation, not related) and not appearing in this table.

Table S3. Incidence of Adverse Events According to Initially Assigned Trial Regimen Adverse Events of Any Cause, regardless of attribution with either an incidence of $\geq 10\%$ or any grade 3 or higher - Max Grade per Patient per Event ^a				
	Sorafenib (N=49)		Placebo (N=36)	
Adverse Event – no. (%)	Grade 1-2	Grade 3-5	Grade 1-2	Grade 3-4
Any adverse event	26 (53)	23 (47)	25 (69)	9 (25)
Events during trial regimen with incidence $\geq 10\%$ in any grade category or any event of grade 3-5				
Palmar-plantar erythrodysesthesia syndrome	34 (69)	1 (2)	8 (22)	0
Rash				
Papulopustular	24 (49)	6 (12)	6 (17)	0
Maculo-papular	7 (14)	0	1 (3)	0
Acneiform	6 (12)	0	0	0
Skin and subcutaneous tissue disorders - other	7 (14)	1 (2)	5 (14)	0
Any rash or skin disorder	36 (73)	7 (14)	15 (42)	0
Pruritus	7 (14)	0	0	0
Fatigue	33 (67)	3 (6)	22 (61)	1 (3)
Hypertension	27 (55)	4 (8)	14 (39)	0
Diarrhea	25 (51)	0	12 (33)	0
Nausea	24 (49)	0	14 (39)	1 (3)
Myalgia	18 (37)	1 (2)	12 (33)	0
Alopecia	18 (37)	0	3 (8)	0
Arthralgia	17 (35)	1 (2)	9 (25)	0
Abdominal pain	15 (31)	1 (2)	9 (25)	4 (11)
Anorexia	15 (31)	0	9 (25)	0

Table S3. Incidence of Adverse Events According to Initially Assigned Trial Regimen Adverse Events of Any Cause, regardless of attribution with either an incidence of $\geq 10\%$ or any grade 3 or higher - Max Grade per Patient per Event ^a				
	Sorafenib (N=49)		Placebo (N=36)	
Adverse Event – no. (%)	Grade 1-2	Grade 3-5	Grade 1-2	Grade 3-4
Mucositis oral	11 (22)	0	6 (17)	0
Constipation	11 (22)	0	4 (11)	0
Vomiting	10 (20)	1 (2)	6 (17)	2 (6)
Anemia	8 (16)	1 (2)	2 (6)	1 (3)
Alanine aminotransferase increased	7 (14)	0	4 (11)	0
Hyperglycemia	6 (12)	1 (2)	3 (8)	0
Platelet count decreased	6 (12)	2 (4)	1 (3)	0
Peripheral sensory neuropathy	6 (12)	0	1 (3)	0
Aspartate aminotransferase increased	5 (10)	1 (2)	3 (8)	0
Blood bilirubin increased	5 (10)	0	3 (8)	1 (3)
Neutrophil count decreased	5 (10)	0	2 (6)	0
Dry skin	5 (10)	0	1 (3)	0
Headache	4 (8)	0	6 (17)	0
White blood cell decreased	3 (6)	0	6 (17)	0
Musculoskeletal, connective tissue - other	3 (6)	0	4 (11)	0
Pain in extremity	3 (6)	1 (2)	2 (6)	1 (3)
Pain	3 (6)	0	2 (6)	1 (3)
General disorders and admin site conditions - other	3 (6)	0	0	1 (3)
Respiratory, thoracic, mediastinal - other	2 (4)	1 (2)	1 (3)	0
Hypernatremia	2 (4)	0	1 (3)	1 (3)

Table S3. Incidence of Adverse Events According to Initially Assigned Trial Regimen Adverse Events of Any Cause, regardless of attribution with either an incidence of $\geq 10\%$ or any grade 3 or higher - Max Grade per Patient per Event ^a				
	Sorafenib (N=49)		Placebo (N=36)	
Adverse Event – no. (%)	Grade 1-2	Grade 3-5	Grade 1-2	Grade 3-4
Eye disorders - other	2 (4)	1 (2)	0	0
Skin infection	2 (4)	1 (2)	0	0
Lymphocyte count decreased	1 (2)	1 (2)	2 (6)	1 (3)
Anxiety	1 (2)	0	0	1 (3)
Hypertriglyceridemia	0	1 (2)	0	1 (3)
Blood and lymph sys disorders - other	0	1 (2)	0	0
Leukocytosis	0	1 (2)	0	0
Gastric perforation	0	1 (2) ^b	0	0
Unintended pregnancy	0	1 (2)	0	0
Sleep apnea	0	1 (2)	0	0
Small intestinal obstruction	0	0	0	2 (6)
Flank pain	0	0	0	2 (6)
Anal fistula	0	0	0	1 (3)
Hyponatremia	0	0	0	1 (3)
Surgical and medical procedures - other	0	0	0	1 (3)

- a) Included are events that were reported in at least 10% of the patients in either group, or if a grade 3-5 event occurred.
- b) A grade 5 adverse event of any cause occurred in 1 patient (2.0%) who received sorafenib (gastric perforation, not related).

Table S4: Summary of toxicities reported by patients using PRO-CTCAE								
PRO-CTCAE Item	With adjustment for baseline							
	No. evaluable patients		Score ≥1			Score ≥3		
	Sorafenib N	Placebo N	Sorafenib N (%)	Placebo N (%)	Risk Difference (95% CI)	Sorafenib N (%)	Placebo N (%)	Risk Difference (95% CI)
Insomnia Severity	35	27	18 (51%)	9 (33%)	18.1% (-7.4 - 41.4%)	6 (17%)	2 (7%)	9.7% (-10.2 - 27.5%)
Insomnia Interference	35	26	15 (43%)	11 (42%)	0.5% (-25.4 - 25.4%)	5 (14%)	4 (15%)	-1.1% (-22.2 - 17.8%)
Constipation Severity	35	26	12 (34%)	13 (50%)	-15.7% (-40.6 - 9.6%)	5 (14%)	5 (19%)	-4.9% (-26.7 - 14.6%)
Pain Frequency	35	26	16 (46%)	10 (38%)	7.3% (-19.0 - 31.7%)	13 (37%)	6 (23%)	14.1% (-10.4 - 36.5%)
Pain Severity	35	26	16 (46%)	11 (42%)	3.4% (-22.2 - 28.5%)	9 (26%)	8 (31%)	-5.1% (-28.9 - 18.2%)
Pain Interference	36	26	19 (53%)	12 (46%)	6.6% (-18.9 - 31.6%)	13 (36%)	5 (19%)	16.9% (-7.0 - 38.4%)
Fatigue Severity	36	27	21 (58%)	17 (63%)	-4.6% (-28.7 - 20.2%)	12 (33%)	10 (37%)	-3.7% (-28.1 - 20.2%)
Fatigue Interference	36	26	17 (47%)	13 (50%)	-2.8% (-27.9 - 22.6%)	10 (28%)	7 (27%)	0.9% (-23.0 - 23.4%)
Nausea Frequency	36	27	21 (58%)	6 (22%)	36.1% (8.1 - 57.6%)	5 (14%)	2 (7%)	6.5% (-12.6 - 23.4%)
Nausea Severity	36	27	22 (61%)	8 (30%)	31.5% (4.8 - 54.0%)	4 (11%)	4 (15%)	-3.7% (-23.9 - 13.9%)
Vomiting Frequency	35	27	11 (31%)	6 (22%)	9.2% (-14.1 - 31.5%)	0 (0%)	1 (4%)	-3.7% (-19.7 - 7.0%)
Vomiting Severity	35	27	10 (29%)	5 (19%)	10.1% (-13.4 - 32.0%)	2 (6%)	1 (4%)	2.0% (-13.5 - 16.6%)
Diarrhea Frequency	35	27	23 (66%)	10 (37%)	28.7% (2.1 - 51.6%)	9 (26%)	3 (11%)	14.6% (-7.0 - 34.4%)
Rash Presence	35	27	24 (69%)	8 (30%)	38.9% (10.6 - 60.7%)	--	--	--
Hand/Foot Syndrome Severity	35	27	27 (77%)	9 (33%)	43.8% (16.4 - 64.7%)	7 (20%)	1 (4%)	16.3% (-1.2 - 33.7%)
Hand/Foot Syndrome Interference	36	27	21 (58%)	9 (33%)	25.0% (-0.9 - 48.3%)	10 (28%)	2 (7%)	20.4% (-0.5 - 39.7%)
Decreased Appetite Severity	36	27	22 (61%)	10 (37%)	24.1% (-1.4 - 47.2%)	4 (11%)	2 (7%)	3.7% (-14.5 - 19.8%)

Decreased Appetite Interference	36	27	17 (47%)	8 (30%)	17.6% (-7.5 - 40.4%)	2 (6%)	1 (4%)	1.9% (-14.0 - 15.7%)
Mouth or Throat Sores Severity	36	27	17 (47%)	9 (33%)	13.9% (-12.6 - 37.3%)	5 (14%)	0 (0%)	13.9% (-0.5 - 29.5%)
CI=confidence interval								

Table S5. Patient Demographics of Crossover Cohort at Randomization ☆

	Crossover (N=28)
Median age (range) – yr	38 (21-60)
Sex – no. (%)	
Female	19 (68)
Male	9 (32)
ECOG performance-status score – no. (%) †	
0	16 (57)
1	12 (43)
Race – no. (%) ‡	
White	22 (79)
Black or African American	4 (14)
Not reported: patient refused or not available	2 (7)
Median sum of target lesions at randomization (range) – cm	8.7 (3.4-26.5)
Brief Pain Inventory (BPI) pain score at randomization – no. (%) §§ ¶	
Score 0-2	10 (36)
Score 3-6	12 (43)

Score 7-10	6 (21)
Intra-abdominal disease – no. (%) ff	
Yes	11 (39)
No	17 (61)
Primary tumor site – no. (%)	
Chest wall	2 (7)
Abdominal wall	2 (7)
Head/Neck NOS (excluding skin and spine)	2 (7)
Upper extremities	3 (11)
Lower extremities	3 (11)
Mesentery	3 (11)
Multiple	7 (25)
Other	6 (21)
Prior radiation therapy – no. (%)	
Yes	3 (11)
No	25 (89)
Prior systemic therapy – no. (%)	
Yes	14 (50)
No	14 (50)
Prior surgical resection – no. (%)	
Yes	13 (46)
No	15 (54)
Disease status – no. (%)	
Newly diagnosed	14 (50)

Recurrent	14 (50)
-----------	---------

- ☆ The intention-to-treat population included all patients who underwent randomization with the exception of those having improper histology. Randomization was based on dynamic allocation algorithms. Randomization was stratified according to anatomical location (intra-abdominal versus extra-abdominal) and level of pain (0-2, 3-6, and 7-10) at the time of randomization; assessed using Item #3 from the Brief Pain Inventory completed by the patient within 28 days prior to randomization. There were no significant differences ($P < 0.05$) between groups at the time of randomization with the exception of criteria #2. Percentages may not total 100 because of rounding.
 - † Eastern Cooperative Oncology Group (ECOG) performance-status scores are assessed on a 5 point scale, with higher scores indicating greater disability; a score of 5 indicates death.
 - ‡ Race was self-reported by patients.
 - ¶ Brief Pain Inventory (BPI) – Short Form questions. #1: Please rate your pain by circling the one number that best describes your pain at its WORST in the last 24 hours: 0 (no pain) – 10 (Pain as bad as you can imagine), #2: Circle the one number that describes how, during the past 24 hours, pain has interfered with your general activity: 0 (does not interfere) – 10 (completely interferes), #3: Circle the one number that describes how, during the past 24 hours, pain has interfered with your sleep: 0 (does not interfere) – 10 (completely interferes). NSAIDs denotes non-steroidal anti inflammatory drugs.
 - ¶¶ Stratification factor at randomization.
-

Table S6. Clinician reported reasons for discontinuing open-label sorafenib for patients initially assigned to placebo

Reason – no. (%)	Total (N=28)
On Treatment	12 (43)
Patient Withdrawal/Refusal After Beginning Protocol Therapy (Intervention)	4 (14)
Adverse Events/Side Effects/Complications	4 (14)
Disease Progression, Relapse During Active Treatment (Intervention)	5 (18)
Alternative Therapy	1 (4)
Other: Non-Compliance (1) and New dysplastic nevi (1)	2 (7)

Table S7. Clinician reported, treatment-related adverse events – Crossover Treatment

Adverse Event – no. (%)	Crossover (N=28)	
	Grade 1-2	Grade 3-4
Palmar-plantar erythrodysesthesia syndrome	12 (52)	1 (4)
Rash - maculo-papular	5 (21)	0
Pruritus	3 (13)	0
Any Rash/Skin	15 (65)	1 (4)
Fatigue	13 (57)	0
Hypertension	9 (39)	1 (4)
Nausea	8 (35)	1 (4)
Arthralgia	8 (35)	0
Abdominal pain	7 (30)	0
Papulopustular rash	7 (30)	0
Diarrhea	6 (26)	3 (13)
Mucositis oral	6 (26)	1 (4)
Myalgia	5 (22)	0
Alopecia	5 (22)	0
Anorexia	4 (17)	1 (4)
Headache	4 (17)	0
Anemia	3 (13)	0
Alanine aminotransferase increased	3 (13)	0
Aspartate aminotransferase increased	3 (13)	0
Neutrophil count decreased	3 (13)	0
Platelet count decreased	3 (13)	0
Weight loss	3 (13)	0
White blood cell decreased	3 (13)	0

Adverse Event – no. (%)	Crossover (N=28)	
	Grade 1-2	Grade 3-4
Esophagitis	0	1 (4)
Gastroesophageal reflux disease	0	1 (4)
Appendicitis	0	1 (4)

REFERENCE

1. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228-47.
2. Gounder MM, Thomas DM, Tap WD. Locally Aggressive Connective Tissue Tumors. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2018;36:202-9.
3. Tap WD, Wainberg ZA, Anthony SP, et al. Structure-Guided Blockade of CSF1R Kinase in Tenosynovial Giant-Cell Tumor. *The New England journal of medicine* 2015;373:428-37.
4. Gounder MM, Lefkowitz RA, Keohan ML, et al. Activity of Sorafenib against desmoid tumor/deep fibromatosis. *Clin Cancer Res* 2011;17:4082-90.
5. Sundaram M, McGuire MH, Schajowicz F. Soft-tissue masses: histologic basis for decreased signal (short T2) on T2-weighted MR images. *AJR Am J Roentgenol* 1987;148:1247-50.