Preoperative radiotherapy in patients with primary retroperitoneal sarcoma: EORTC-62092 trial (STRASS) versus off-trial (STREXIT) results

Callegaro, D.; Raut, C.P.; Ajayi, T.; Strauss, D.; Bonvalot, S.; Ng, D.; ... ; Gronchi, A.

Preoperative Radiotherapy in Patients With Primary Retroperitoneal Sarcoma

EORTC-62092 Trial (STRASS) Versus Off-trial (STREXIT) Results

Dario Callegaro, MD, Chandrajit P. Raut, MD, MSc,† Taiwo Ajayi, MSc,‡ Dirk Strauss, MD,§ Sylvie Bonvalot, MD, PhD,‖ Deanna Ng, MD,¶# Eberhard Stoeckle, MD,** Mark Fairweather, MD,† Piotr Rutkowski, MD, PhD,†† Winan J. van Houdt, MD, PhD,‡‡ Hans Gelderblom, MD, PhD,§§ Claudia Sangalli, MD,‖‖ Andrew Hayes, MD,§ Charles Honoré, MD,¶¶ Rebecca A. Gladdy, MD, PhD,¶ Magali Fau, MD,# Rick Haas, MD,##*** Dimitri Tzanis, MD,‖ Aisha B. Miah, MD,††† Peter Chung, MD,‡‡‡§§§ Elizabeth H. Baldini, MD, MPH,‖‖‖ Sandrine Marreaud, MD,¶¶¶ Saskia Litiere, MSc, PhD,‡ Carol J. Swallow, MD, PhD,¶ and Alessandro Gronchi, MD*✉*

Objective: The aim of the present study was to compare the effect of radiotherapy (RT) on abdominal recurrence-free survival (ARFS) in patients with primary retroperitoneal sarcoma treated in the EORTC-STBSG-62092 (STRASS) phase 3 randomized controlled trial (STRASS cohort) and off-trial (STREXIT cohort) and to pool STRASS and STREXIT data to test the hypothesis that RT improves ARFS in patients with liposarcoma.

Background: The STRASS trial did not show any difference in ARFS between patients treated with preoperative radiotherapy+surgery (RT+S) versus surgery alone (S).

Methods: All consecutive adult patients not enrolled in STRASS and underwent curative-intent surgery for a primary retroperitoneal sarcoma with or without preoperative RT between 2012 and 2017 (STRASS recruiting period) among ten STRASS-recruiting centres formed the STREXIT cohort. The effect of RT in STREXIT was explored with a propensity score (PS)-matching analysis. Primary endpoint was ARFS defined as macroscopically incomplete resection or abdominal recurrence or death of any cause, whichever occurred first.

Results: STRASS included 266 patients, STREXIT included 831 patients (727 after excluding patients who received preoperative chemotherapy, 202 after 1:1 PS-matching). The effect of RT on ARFS in STRASS and 1:1 PS-matched STREXIT cohorts, overall and in patients with liposarcoma, was similar. In the pooled cohort analysis, RT administration was associated with better ARFS in patients with liposarcoma [N = 321, hazard ratio (HR), 0.61; 95% confidence interval (CI), 0.42–0.89]. In particular, patients with well-differentiated liposarcoma and G1-2 dedifferentiated liposarcoma (G1-2 DDLPS, n = 266) treated with RT+S had better ARFS (HR, 0.63; 95% CI, 0.40–0.97) while patients with G3 DDLPS and leiomyosarcoma had not. At the current follow-up, there was no association between RT and overall survival or distant metastases-free survival.

Conclusions: In this study, preoperative RT was associated with better ARFS in patients with primary well-differentiated liposarcoma and G1-2 DDLPS.

Keywords: sarcoma, retroperitoneal sarcoma, radiotherapy, recurrence, STRASS, STREXIT, survival

(*Ann Surg* 2023;278:127–134)

From the *Department of Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; †Department of Surgery, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA; ‡Department of Statistics, European Organisation for Research and Treatment of Cancer, Brussels, Belgium; §Department of General Surgery, Royal Marsden NHS Foundation Trust, London, UK; ‖Department of Surgery, Institute Curie, University of Paris, Paris, France; ¶Department of Surgical Oncology, Princess Margaret Cancer Center/Mount Sinai Hospital, Toronto, ON, Canada; #Department of Surgery, University of Toronto, Toronto, ON, Canada; **Department of Surgery, Institut Bergonie, Bordeaux, France; ††Department of Soft Tissue/Bone Sarcoma and Melanoma, Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland; ‡‡Department of Surgical Oncology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; §§Department of Medical Oncology, Leiden University Medical Center, Leiden, The Netherlands; ‖‖Department of Radiation Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; ¶¶Department of Surgical Oncology, Institut Gustave Roussy, Villejuif, France; ##Department of Radiotherapy, The Netherlands Cancer Institute, Amsterdam, The Netherlands; ***Department of Radiotherapy, Leiden University Medical Center, Leiden, The Netherlands; †††Department of Radiation Oncology, Royal Marsden NHS Foundation Trust, and The Institute of Cancer Research, London, UK; ‡‡‡Department of Radiation Oncology, Princess Margaret Cancer Center/Mount Sinai Hospital, Toronto, ON, Canada; §§§Department of Surgery, University of Toronto, Toronto, ON, Canada; ‖‖‖Department of Radiation Oncology, Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School, Boston, MA; and ¶¶¶European Organisation for Research and Treatment of Cancer, Brussels, Belgium.

✉alessandro.gronchi@istitutotumori.mi.it.

This study was presented, in part, at the 2020 Annual Virtual Meeting Connective Tissue Oncology Society (CTOS); virtual; November 18–21 (abstr 17).

The authors report no conflicts of interest.

Supplemental Digital Content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, www.annalsofsurgery.com.

ISSN: 0003-4932/23/27801-0127
DOI: 10.1097/SLA.0000000000005492

Standard treatment of primary localized retroperitoneal sarcoma (RPS) is complete surgical excision.[1] After surgery, main prognostic factors that impact survival are age, tumor size,

tumor grade, histological type, and completeness of surgical resection.[2,3] Furthermore, histological type and subtype strongly influence pattern of tumor recurrence, with well-differentiated liposarcoma (WDLPS) and G1-2 dedifferentiated liposarcoma (DDLPS) recurring mainly locally, G3-DDLPS recurring both locally and distantly and leiomyosarcoma (LMS) recurring mainly at distant sites.[4–6] Also, the use of preoperative radiotherapy (RT) to possibly improve local control varies considerably across different sarcoma centers.[6–8]

The phase 3 EORTC-STBSG-62092 (STRASS) trial randomized patients with primary localized RPS to receive either preoperative RT followed by surgery (RT+S) or surgery alone (S). The primary endpoint was abdominal recurrence-free survival (ARFS). At median follow-up (FU) of 43 months, the trial failed to show that, overall, administration of preoperative RT is associated with better ARFS. However, the Independent Data Monitoring Committee recommended a subsequent sensitivity analysis among the subgroup of patients with liposarcoma. Specifically, for the primary endpoint where progression under RT or becoming medically unfit was not considered a failure if patients subsequently underwent a macroscopically complete resection, patients with liposarcoma receiving preoperative RT had a 10% absolute ARFS benefit at 3 years (65% vs 75%) with a hazard ratio (HR) of 0.62 [95% confidence interval (CI): 0.38–1.02]. Nevertheless, since this was a post hoc analysis and the study was not powered to detect a difference in the liposarcoma subgroup, this finding remained only hypothesis generating.[9]

The interpretation of STRASS results was heterogeneous within the sarcoma community.[10–13] In light of the results of the sensitivity analysis described above, in some sarcoma centers this trial became rationale for more frequent use of RT in patients with liposarcoma. Conversely, due to the fact that the trial did not meet its primary objective, in other sarcoma centers, this trial lead to a phasing out of RT use altogether in patients with primary RPS.

To address this clinical controversy following STRASS results, the aims of the present study were to compare outcomes of patients treated in-trial (STRASS) and off-trial (STREXIT) with respect to RT administration and to test the hypothesis that RT improves ARFS in patients with primary RPS in a pooled STRASS+STREXIT cohort, after propensity score (PS) matching of STREXIT.

METHODS

Study Design and Participants

STRASS-recruiting centers that enrolled at least 1 patient in STRASS were asked to join the current study. The following ten centres provided data:

- Brigham and Women's Hospital, Dana-Farber Cancer Institute, Harvard Medical School (Boston, MA).
- Fondazione IRCCS Istituto Nazionale dei Tumouri (Milan, Italy).
- Institut Bergonié (Bordeaux, France).
- Institut Curie (Paris, France).
- Institut Gustave Roussy (Villejuif, France).
- Leiden University Medical Centre (Leiden, the Netherlands).
- Maria Sklodowska-Curie National Research Institute of Oncology (Warsaw, Poland).
- Mount Sinai Hospital/Princess Margaret Cancer Centre, University of Toronto (Toronto, Canada).
- Royal Marsden NHS Foundation Trust (London, UK).
- The Netherlands Cancer Institute (Amsterdam, the Netherlands).

All consecutive adult (18-year-old and above) patients who were not enrolled in STRASS and underwent curative-intent surgery for a primary, localized RPS with or without preoperative RT between January 2012 and April 2017 (STRASS recruiting period) in the aforementioned 10 centers formed the STREXIT cohort. Patients receiving preoperative chemotherapy were not included in the analysis. Similar to STRASS, patients diagnosed with gastrointestinal stromal tumor, desmoid tumor, gynecological sarcoma, bone sarcoma, alveolar/embryonal rhabdomyosarcoma and Ewing family tumors were excluded. Patients enrolled in STRASS formed the STRASS cohort. The reasons why patients in the STREXIT group did not participate in the STRASS trial are the subject of a different article.

For the STREXIT cohort, data were retrieved from prospectively-maintained databases in place at each institution throughout the study period. Variable definition was concordant to what was adopted in STRASS. In particular, tumor grade was assigned according to the FNCLCC (French Federation of Centers for the Fight against Cancer) criteria (grades I, II, and III). WDLPS was graded as 1 by definition. DDLPS was predominantly graded as 2 or 3. Grade 1 DDLPS was defined as WDLPS with features of uniform fibroblastic spindle cells with mild nuclear atypia and exhibiting cellularity (previously called cellular variant of WDLPS). In STREXIT, tumor grade was defined as the grade of the surgical specimen when available. When not available (ie, after preoperative treatments), we considered the grade of the biopsy. Histological subtype was determined according to WHO criteria. Multifocality was defined as presence of discontinuous tumor foci, separated by normal tissue. Tumor rupture was defined as any discontinuation of the tumor pseudocapsule, with or without spillage of tumor, liquid or necrotic material in the operative field.

In STREXIT, decision to administer perioperative RT and/or chemotherapy was made in the context of multidisciplinary sarcoma tumor board recommendation when a high risk of relapse was foreseen. After surgery, FU scheme consisted of clinical examination and computed tomography scan of the abdomen and chest every 4 months for the first 2 years, every 6 months up to the fifth year and annually from the fifth year onwards. For the STRASS cohort, data were retrieved from the clinical database of the STRASS study.

The study was approved by the Institutional Ethics Committee at Fondazione IRCCS Istituto Nazionale Tumori, Milan, Italy (Institutional Study Protocol: INT 251/20).

Outcomes

The main study endpoints were ARFS (defined as abdominal recurrence or death of any cause, whichever occurred first), overall survival (OS; death of any cause), distant metastasis-free survival (DMFS; occurrence of distant recurrence, with or without concomitant abdominal recurrence or death of any cause, whichever occurred first) and abdominal recurrence-free interval (ARFI; abdominal recurrence or death due to abdominal recurrence; competing events included deaths not due to abdominal failure and distant metastases occurring before abdominal recurrence). For all analyses, abdominal recurrence was defined as macroscopically incomplete resection or local relapse (with or without concomitant distant relapse). Of note, this definition is different from the definition adopted for the primary endpoint in STRASS and it is in line with the definition adopted in the sensitivity analysis recommended by the IDMC

where local progression under RT not leading to inoperability was not considered as an event of interest. Reference time-point was date of randomization in STRASS cohort and date of diagnosis in STREXIT cohort. Time was censored at date of last FU for patients alive and who did not experience any events of interest or competing risks.

ARFS, OS, and DMFS curves were calculated using Kaplan-Meier method and effect of preoperative RT was estimated using the Cox model. ARFI cumulative incidence curves were calculated and effect of preoperative RT was estimated with the Fine and Gray model.

PS Matching Analysis in STREXIT

Since patients in STREXIT were not randomized, to minimize bias in the estimate of the effect of RT on study endpoints, we performed an analysis of PS-matched data.[14,15] PS was estimated using multivariate logistic regression with a binary response representing treatment, RT+S versus S, and including age, sex, tumor size, tumor grade, multifocality, and histology as covariates. Patients treated with S were matched 1:1 with patients treated with RT+S according to the PS. To remove potential for poorly matched patients, matching was based on a caliper, which defines the distant metric below which patients will be matched. In this study, a caliper of 0.25 SD of the linear PS was used, as suggested by Rosenbaum and Rubin.[14] Given the fact that the sample size of the S group was significantly larger than the sample size of RT+S group, in order to gain power and precision when estimating the treatment effect, a 2:1 matching was also performed. In this case, based on the PS and caliper of 0.25, 2 patients treated with S were matched with 1 patient treated with RT+S.

Pooled Analysis

In order to increase the power of our analysis and the precision of the treatment effect estimates, STRASS and 1:1 matched STREXIT datasets were pooled with an individual patient data approach. The Cox and Fine and Gray models were stratified by study.

Statistical analysis was performed using SAS software (SAS Institute, Cary, NC) and R software (version 4.0.2.).

RESULTS

The STRASS and STREXIT cohort consisted of 266 and 831 patients, respectively. After excluding from STREXIT patients who received preoperative chemotherapy (n = 99) or with missing data about chemotherapy administration (n = 5), STREXIT comprised 727 patients. Median FU in STREXIT was 39.4 months from diagnosis [interquartile range (IQR): 25.2–58.0] and in STRASS 43.1 months from randomization (IQR: 28.8–59.2). Clinical and pathological characteristics of the 2 cohorts are detailed in Table 1.

PS Matching

Supplementary Figure 1 (Supplemental Digital Content 1, http://links.lww.com/SLA/D988) shows the distribution of PS in STREXIT stratified by subgroup (S vs RT+S), before and after matching. After 1:1 PS matching, 202 patients were included in the 1:1 STREXIT cohort (101 in S and 101 in RT+S). Median FU from diagnosis in the 1:1 cohort was 44.1 months (IQR: 25.2.–62.0). Demographic, clinical, and pathological characteristics of the 1:1 STREXIT cohort are showed in Table 1. After 1:1 PS matching, baseline characteristics of S and RT+S subgroups were similar with regard to variables used in matching as well as performance status, tumor rupture, and tumor multifocality.

The pooled STRASS and 1:1 STREXIT cohorts included 468 patients: 234 treated with S and 234 treated with RT+S.

Outcome Analysis

Abdominal Recurrence-free Survival

ARFS events in the different cohorts are tabulated in Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/SLA/D989).

Kaplan-Meier curves for ARFS by treatment group (S vs RT+S) in STRASS, 1:1 STREXIT and pooled cohort are shown in Figure 1. Overall, including all histological types, RT administration was not associated with better ARFS in STRASS (HR: 0.83, 95% CI: 0.58–1.21; Fig. 1A) while it was associated with better ARFS in 1:1 STREXIT (HR: 0.63, 95% CI: 0.41–0.99; Fig. 1B) and in the pooled cohort (HR: 0.75, 95% CI: 0.56–1.01, Fig. 1C).

In the subgroup analysis of patients with liposarcoma, RT administration was associated with better ARFS in STRASS (Supplemental Fig. 2, Panel A, Supplemental Digital Content 3, http://links.lww.com/SLA/D990) with an HR of 0.63 (95% CI: 0.40–1.00) and in the pooled cohort (Fig. 2A) with an HR of 0.61 (95% CI: 0.42–0.89). In 1:1 STREXIT (Supplemental Fig. 2, panel B, Supplemental Digital Content 3, http://links.lww.com/SLA/D990), the association between RT and ARFS was also observed, although it was not statistically significant (HR: 0.58, 95% CI: 0.30–1.11).

To explore the effect of RT for histologic subtypes with different patterns of recurrence and sufficient sample size, further subgroup analyses were performed in the pooled cohort (STRASS+1:1 STREXIT) only. In patients with WDLPS and G1-2 DDLPS (n = 266; Fig. 2, panel B), patients treated with preoperative RT had better ARFS (HR: 0.63; 95% CI: 0.40–0.97). In particular, 5-year ARFS was 65.8% (54.7%–74.8%) in RT+S group and 56.0% (44.1–66.4) in S group. In the subgroup of patients with G3 DDLPS (n = 29; Fig. 2C) and LMS (n = 65, Fig. 2D) RT administration was not associated with ARFS, with an HR of 0.68 (95% CI: 0.22–2.16) and 0.99 (95% CI: 0.47–2.11), respectively.

Overall Survival

OS events by subgroups are detailed in Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/SLA/D989). Administration of RT was not associated with OS in STRASS, 1:1 STREXIT and pooled cohort, neither overall (all histologies) nor in liposarcoma patients (Supplemental Table 3, Supplemental Digital Content 2, http://links.lww.com/SLA/D989). Subgroup analyses of patients with WDLPS+G1-2 DDLPS, G3 liposarcoma, and LMS in the pooled cohort similarly did not show any association between RT administration and OS (Supplemental Table 3, Supplemental Digital Content 2, http://links.lww.com/SLA/D989).

Distant Metastasis-free Survival

DMFS events are shown in Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/SLA/D989). Administration of RT was not associated with DMFS in STRASS, 1:1 STREXIT and pooled cohort, overall nor in subgroup analyses (Supplemental Table 3, Supplemental Digital Content 2, http://links.lww.com/SLA/D989).

TABLE 1. Demographic, Clinical and Pathological Characteristics of the STREXIT and STRASS Cohorts, Stratified by Preoperative RT Administration

	STREXIT			1:1 Matched STREXIT			STRASS		
	S (N = 620)	RT+S (N = 107)	Total (N = 727)	S (N = 101)	RT+S (N = 101)	Total (N = 202)	S (N = 133)	RT+S (N = 133)	Total (N = 266)
Sex, n (%)									
Female	304 (49.0)	41 (38.3)	345 (47.5)	33 (32.7)	40 (39.6)	73 (36.1)	66 (49.6)	62 (46.6)	128 (48.1)
Male	316 (51.0)	66 (61.7)	382 (52.5)	68 (67.3)	61 (60.4)	129 (63.9)	67 (50.4)	71 (53.4)	138 (51.9)
Age at diagnosis (y)									
Median (IQR)	63 (53–71)	63 (54–70)	63 (53–71)	65 (55–72)	63 (54–70)	63 (54–71)	62 (54–68)	62 (53–68)	62 (53–68)
WHO PS, n (%)									
0	182 (29.4)	22 (20.6)	204 (28.1)	33 (32.7)	20 (19.8)	53 (26.2)	100 (75.2)	110 (82.7)	210 (78.9)
1	186 (30.0)	23 (21.5)	209 (28.7)	32 (31.7)	20 (19.8)	52 (25.7)	33 (24.8)	22 (16.5)	55 (20.7)
2	19 (3.1)	5 (4.7)	24 (3.3)	6 (5.9)	5 (5.0)	11 (5.4)	0	1 (0.8)	1 (0.4)
3–4	6 (0.9)	—	6 (0.9)	—	—		—	—	—
Missing	227 (36.6)	57 (53.3)	284 (39.1.)	30 (29.7)	56 (55.4)	86 (42.6)	—	—	—
Tumor size, mm									
Median (IQR)	180 (110–250)	146 (100–190)	170 (110–240)	150 (100–210)	150 (100–200)	150 (100–200)	170 (124–210)	160 (114–210)	165 (117–210)
Grade, n (%)									
I	181 (29.2)	22 (20.6)	203 (27.9)	22 (21.8)	22 (21.8)	44 (21.8)	43 (32.3)	44 (33.1)	87 (32.7)
II	210 (33.9)	60 (56.1)	270 (37.1)	58 (57.4)	59 (58.4)	117 (57.9)	38 (28.6)	47 (35.3)	85 (31.9)
III	185 (29.8)	21 (19.6)	206 (28.3)	21 (20.8)	20 (19.8)	41 (20.3)	19 (14.3)	12 (9.0)	31 (11.7)
Missing/not evaluable	44 (7.1)	4 (3.7)	48 (6.6)	—	—	—	33 (24.8)	30 (22.6)	63 (23.7)
Histology									
WDLPS	152 (24.5)	9 (8.4)	161 (22.1)	9 (8.9)	9 (8.9)	18 (8.9)	42 (31.6)	46 (34.6)	88 (33.1)
DDLPS	284 (45.8)	56 (52.3)	340 (46.8)	51 (50.5)	54 (53.5)	105 (52.0)	54 (40.6)	51 (38.4)	105 (39.5)
LMS	102 (16.5)	15 (14.0)	117 (16.1)	14 (13.9)	13 (12.9)	27 (13.4)	22 (16.5)	16 (12.0)	38 (14.3)
Other	82 (13.2)	27 (25.2)	109 (15.0)	27 (26.8)	25 (24.7)	52 (25.7)	15 (11.3)	19 (14.3)	34 (12.8)
Missing	—	—	—	—				1 (0.8)	1 (0.4)
Multifocality									
Yes	31 (5.0)	4 (3.7)	35 (4.8)	2 (2.0)	4 (4.0)	6 (3.0)	2 (1.5)	2 (1.5)	4 (1.5)
No	564 (91.0)	100 (93.5)	664 (91.3)	99 (98.0)	97 (96.0)	196 (97.0)	87 (65.4)	80 (60.2)	167 (62.8)
Missing	25 (4.0)	3 (2.8)	28 (3.9)	—	—	—	44 (33.1)	51 (38.4)	95 (35.7)
Tumor rupture									
Yes	18 (2.9)	8 (7.5)	26 (3.6)	2 (2.0)	6 (5.9)	8 (4.0)	3 (2.3)	3 (2.3%)	6 (2.3)
No	602 (97.1)	99 (92.5)	701 (96.4)	99 (98.0)	95 (94.1)	194 (96.0)	126 (94.7)	120 (90.2)	246 (92.5)
Missing	—	—	—	—	—	—	4 (3.0)	10 (7.5)	14 (5.3)
Macroscopically complete resection									
Yes (R0/R1)	588 (94.8)	103 (96.3)	691 (95.0)	97 (96.9)	98 (97.0)	195 (96.5)	133 (100.0)	133 (100.0)	266 (100.0)
No (R2)	32 (5.2)	4 (3.7)	36 (5.0)	4 (4.0)	3 (3.0)	7 (3.5)	0 (0.0)	0 (0.0)	0 (0.0)
Preop RT dose (Gy)									
Median (IQR)	—	50.4 (50.0–50.4)	—		50.4 (50.0–50.4)		—	50.4 (50.4–50.4)	—
Postop RT									
Yes	4 (0.6)	—	4 (0.6)	1 (1.0)	0 (0.0)	201 (99.5)	—	—	—
No	616 (99.4)	107 (100)	723 (99.4)	100 (99.0)	101 (100.0)	1 (0.5)	133 (100)	133 (100)	266 (100)
Postop CT									
Yes	17 (2.7)*	1 (0.9)	18 (2.5)	1 (1.0)	1 (1.0)	2 (1.0)	—	—	—
No	603 (97.3)	106 (99.1)	709 (97.5)	100 (99.0)	100 (99.0)	200 (99.0)	133 (100)	133 (100)	266 (100)

*Four patients had both preoperative and postoperative chemotherapy.
CT indicates chemotherapy; WHO PS, World Health Organization Performance Status.

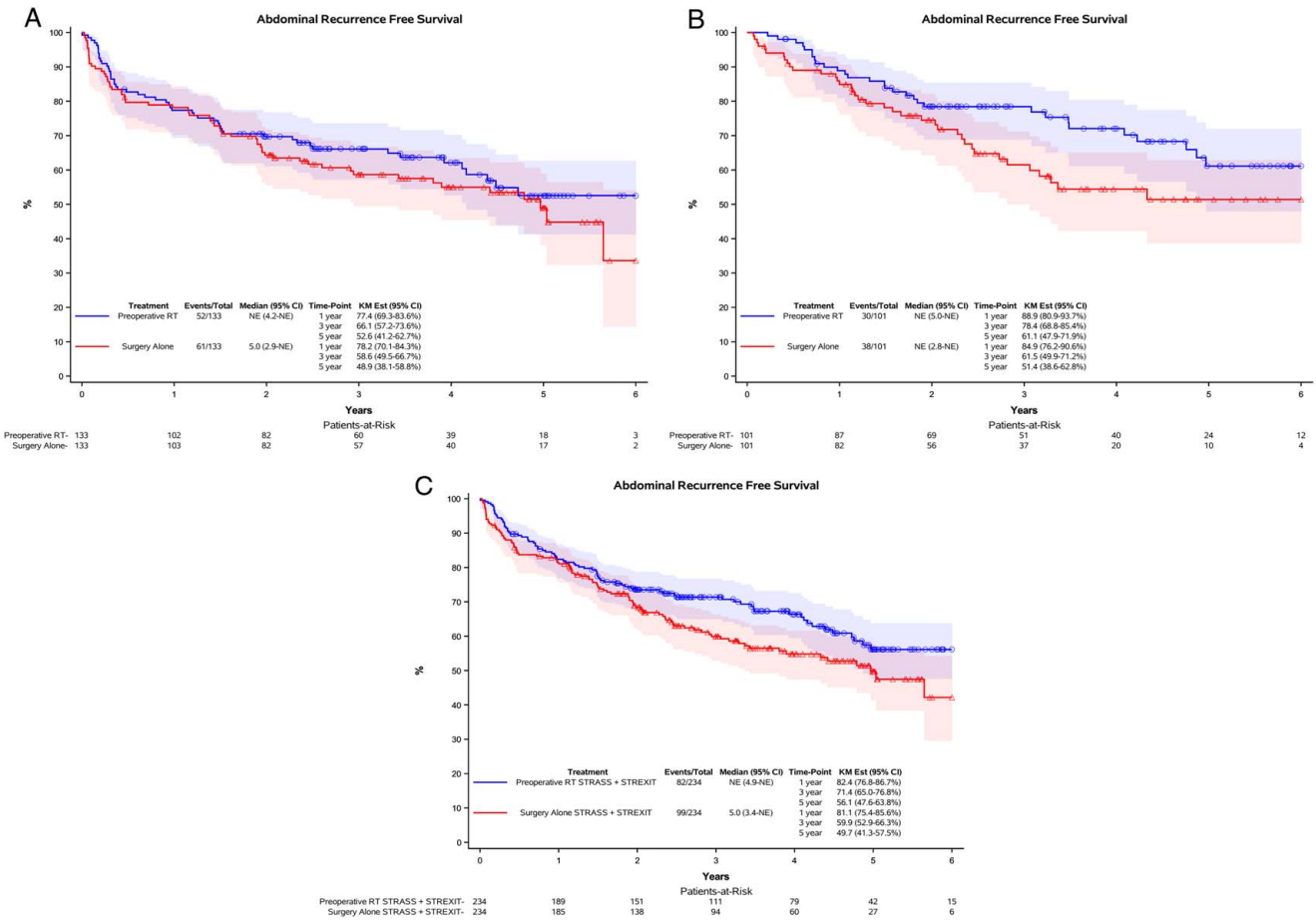

FIGURE 1. Abdominal recurrence-free survival curves according to treatment (blue: preoperative radiotherapy + surgery; red: surgery alone) in STRASS (A), STREXIT after 1:1 propensity score-matching (B), and pooled cohort (C).

Competing Risk Analysis of ARFI

Events of interest and competing risk events for ARFI are shown in Supplemental Table 2 (Supplemental Digital Content 2, http://links.lww.com/SLA/D989). In the pooled cohort, RT administration was associated with better ARFI in subgroup analyses of patients with liposarcoma (HR: 0.59, 95% CI: 0.38–0.93; Fig. 3B) and WDLPS+G1-2 DDLPS (HR: 0.59, 95% CI: 0.36–0.97; Fig. 3C) (Supplemental Table 3, Supplemental Digital Content 2, http://links.lww.com/SLA/D989). In particular, in the subgroup analysis of patients with WDLPS+G1-2 DDLPS, administration of preoperative RT was associated with an absolute 11.6% decrease from 34.6% to 23.0% in cumulative incidence of ARFI at 5 years. Cumulative incidence curves of abdominal recurrence in the pooled cohort (overall and subgroup analyses) are shown in Figure 3.

Analysis of the 2:1 Matched Cohort

Analyses of the 2:1 matching of STREXIT are presented in the Supplemental Materials (Supplementary Fig. 3, Supplemental Digital Content 4, http://links.lww.com/SLA/ D991). and Supplemental Table 1 (Supplemental Digital Content 2, http://links.lww.com/SLA/D989).

DISCUSSION

In this study of patients with primary RPS who underwent surgery within (STRASS cohort) and outside (STREXIT cohort) the EORTC-STBSG-62092 (STRASS) randomized controlled trial (RCT), we observed that RT had a similar effect on ARFS in patients treated in-trial and off-trial.[9] In addition, when we merged STRASS and PS-matched STREXIT data, in the pooled cohort we observed that RT administration was associated with better ARFS and better ARFI in patients with WDLPS and G1-2 DDLPS with a HR of about 0.6. Conversely, we did not observe any association between RT administration and ARFS or ARFI in patients with G3 DDLPS and LMS, although these latter analyses were underpowered. At the current FU, RT was not associated with OS or DMFS, overall nor in any subgroups.

The main limitation of this study is related to the retrospective design and analysis of the observational cohort (STREXIT). To compensate for selection bias regarding treatment choice, we used PS matching. This allowed us to effectively balance baseline prognostic factors in RT+S vs S. PS-matching is indeed an efficient method to mimic randomization and in this study it accounts for the known variables that could affect both the decision to administer RT and the primary endpoint. It is inherent to this methodology that other relevant variables not captured by the study database or that do not have a defined

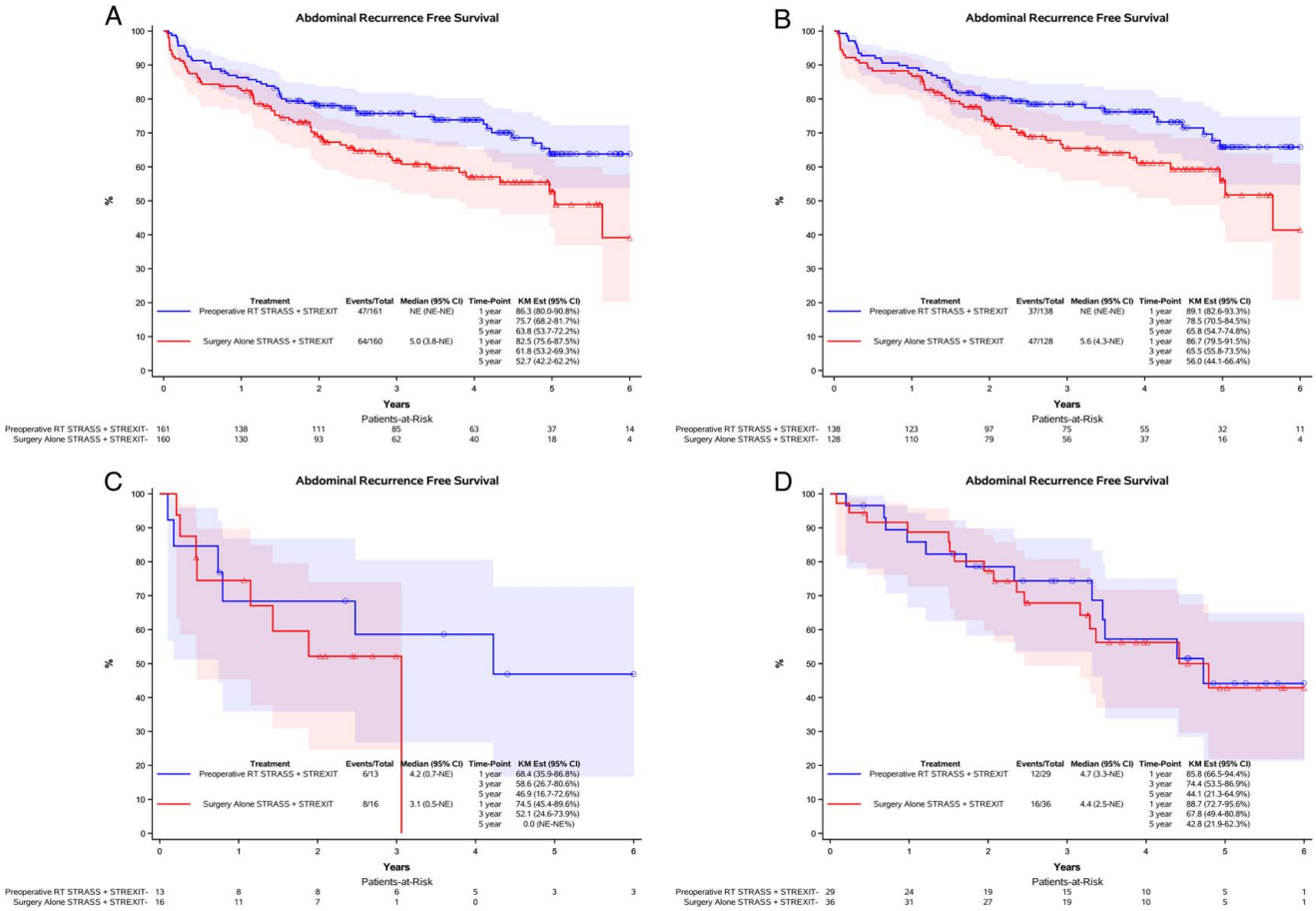

FIGURE 2. Abdominal recurrence-free survival curves in the pooled cohort subgroup analyses according to treatment (blue: preoperative radiotherapy + surgery; red: surgery alone). A, Patients with liposarcoma. B, Patients with G1-2 dedifferentiated liposarcoma and well-differentiated liposarcoma. C, Patients with G3 dedifferentiated liposarcoma. D, Patients with leiomyosarcoma.

prognostic role, such as the location of the tumor with respect to critical anatomical structures, may exist. Moreover, compared with the randomized cohort, in STREXIT we could not control for surgical and RT technique, which on the contrary were detailed in STRASS protocol (despite a high proportion of protocol deviation in terms of RT technique).[16] In other words, PS-matching is an effective way to mimic randomization but of course it is not perfect. Given these premises, pooling randomized and PS-matched observational data do not represent a standard research methodology and it is an approximation of what a larger trial could have achieved. To support this mixed randomized-observational study, in STREXIT patients were treated in the same institutions that were actively enrolling in STRASS and STRASS and PS-matched STREXIT data showed similar outcomes in terms of RT effect in different histologic type. In theory, stratification of the STRASS cohort by risk of LR could also help identifying the subgroups of patients that could benefit the most from RT but there is no validated instrument to predict the risk of LR in patients with RPS and, again, STRASS was not powered to detect differences in subgroups.

This is the first and largest combined randomized-observational study investigating the use of preoperative RT in primary RPS. By merging patients treated within and outside trial we were able to overcome the major limitation of STRASS, which is related to the low number of patients in subgroup analyses. The current study allowed us to further explore the role of RT in the treatment of patients with RPS, and to provide more data regarding the subgroups of patients that may benefit from preoperative RT. Another reason for the importance of data is that the recently launched STRASS-2 trial, a phase 3 RCT which tests the efficacy of neoadjuvant chemotherapy in G3 DDLPS and LMS, will only be analyzing the histological types that appear to derive no benefit from preoperative RT. Therefore, the subgroup of patients not included in STRASS-2 (mainly G1-2 DDLPS and WDLPS) require additional studies of adjuvant approaches that could improve local control.

The observations that PS-matched STREXIT survival curves mirrored STRASS survival curves and that patients treated outside trial outnumbered patients treated in trial are thought-provoking. Randomization provides higher quality data to test effects of a treatment, but it might also represent an obstacle to patient enrolment since patient or treating physicians might not see equipoise between treatment arms, due to the particular clinical presentation or due to preconceptions about treatment efficacy. Use of mixed observational-RCTs might help the sarcoma community enhance patient participation into trials in the future.

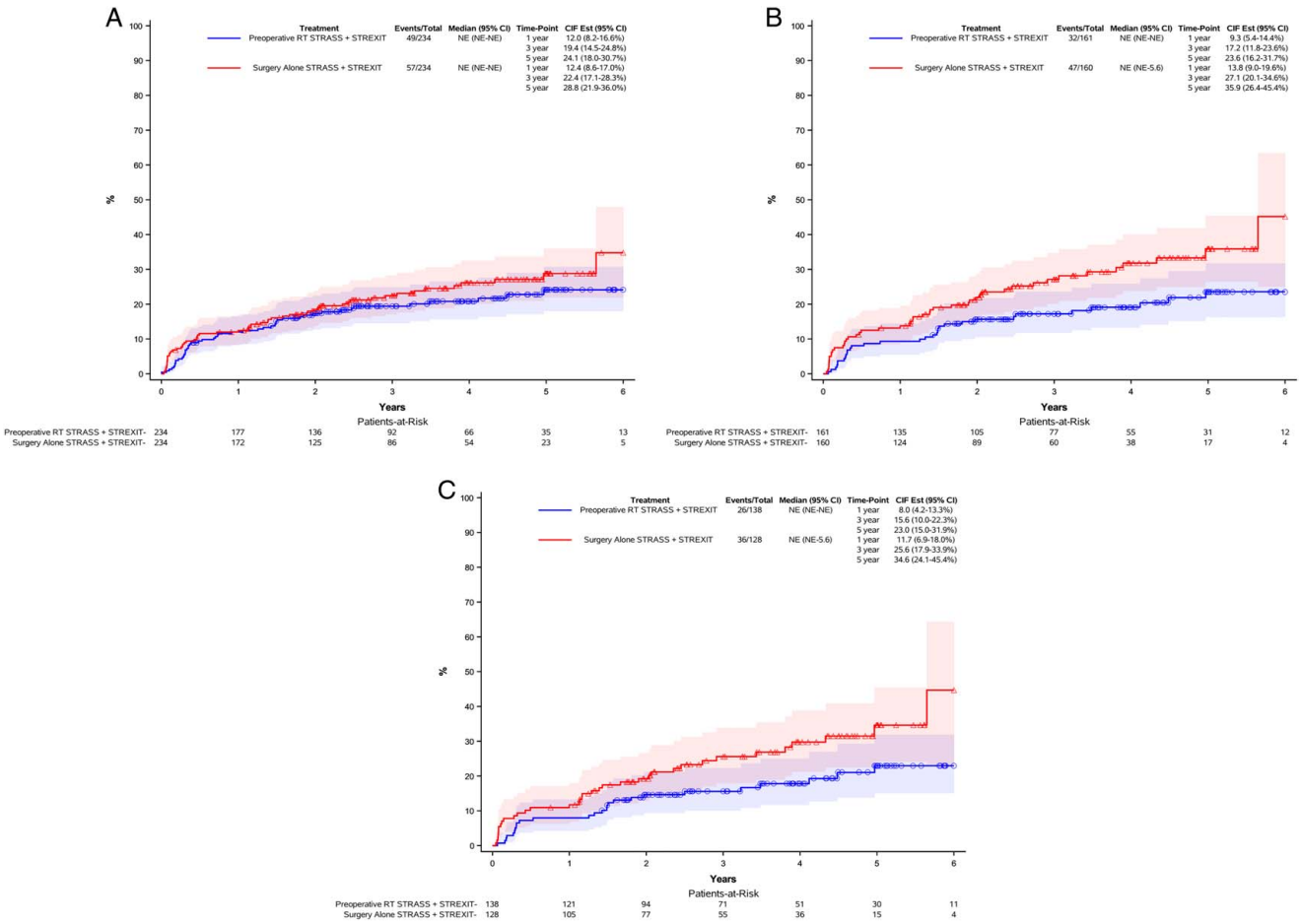

FIGURE 3. Abdominal recurrence-free interval curves according to treatment (blue: preoperative radiotherapy + surgery; red: surgery alone) in the pooled cohort overall (A), in patients with liposarcoma (B), in patients with G1-2 dedifferentiated liposarcoma and well-differentiated liposarcoma (C).

In essence, our results are consistent with STRASS, which showed a 10% absolute ARFS benefit in favor of RT in the exploratory analysis of patients with liposarcoma, and did not support the use of RT in patients with LMS. In addition, the current study helps to further dissect the effect of RT in the liposarcoma subgroup showing that RT may increase local control in WDLPS and G1-2 DDLPS, but not in G3 DDLPS.[6] Indeed the magnitude of benefit from RT on ARFS observed in the pooled cohort of the current study in patients with WDLPS and G1-2 DDLPS (HR: 0.63) is close to the HR that STRASS failed to detect as its primary endpoint (HR: 0.52). Whether the lack of an effect of preoperative RT on ARFS in patients with G3 DDLPS was related to their natural history or to the small number of patients in this series remains an open question. However the curves do overlap for the first 3 years and the separation of their tales is based on very few patients, with no significant difference.

Several older studies did not distinguish DDLPS by tumor grade, showing an overall high risk of local recurrence (LR) and moderate risk of distant metastases.[17] On the contrary, modern single and multi-institutional series have shown that G2 DDLPS maintains a largely predominant LR risk of about 40% at 5-year with very limited metastatic risk (below 10% at 5-year)[2,6,18] while G3 DDLPS has both a high risk of local and distant failure, in the 30% to 40% range. This may explain why a local treatment such as preoperative RT may benefit WDLPS and G1-2 DDLPS, which are the histological subtypes with a predominantly local pattern of failure, but much less, if at all, G3 DDLPS, characterized by a high metastatic risk.[6,19]

The STRASS trial was designed in 2010 and first results were published in 2020. The main limitations in its design are related to the lack of knowledge about RPS biology and clinical behavior when the study was conceived and the expected difficulty to enroll patients after the premature closure of the ACOSOG-Z9031 trial. In particular, the trial did not stratify for histological type or subtype. Furthermore, STRASS adopted as primary endpoint a composite metric that considered progression under RT as an event, since it was not known that most of the patients who progress during RT could still undergo complete resection. In the current study, we adopted a definition of ARFS that is in line with the one adopted in STRASS sensitivity analysis and, compared with STRASS, increased the power of the statistical analysis by doubling the number of patients in the pooled cohort. As a result, this study provides strong evidence of the association between RT administration and ARFS in patients with WDLPS and G1-2 DDLPS.

In the current study, with median FU of 43 months in STRASS and 39 months in STREXIT, RT did not impact OS in

any subgroup of patients, and in particular in WDLPS and G1-2 DDLPS. Whether better local control achieved with RT in these histological subtypes will translate to a subsequent survival benefit with mature FU remains to be seen. Indeed patients with WDLPS and G1-2 DDLPS are characterized by long OS (5-year OS 90% and 67%, respectively), an LR curve which does not flatten out even 10 years after surgery (5-year cumulative incidence of LR 23% in WDLPS and 43% in G1-2 DDLPS) and a potential for prolonged postrelapse survival, especially if the initial disease-free interval is long and the recurrence is resected.[6,19,20] Thus, a longer FU may inform if there is an effect of a local treatment on survival in these subgroups. For example, in the past the adoption of an extended surgical approach was initially associated with a better local control and only at a longer FU also with a survival benefit in patients with G1-2 RPS.[21,22]

In conclusion, this study provides further support for the use of preoperative RT to improve local control for patients with primary retroperitoneal WDLPS and G1-2 DDLPS while it does not support the use of RT in patients with LMS and G3 DDLPS. The potential effect of RT on survival awaits longer follow-up. Future studies of sarcoma types/subtypes should consider a mixed observational-RCT design as a valuable tool to increase sample size.

REFERENCES

1. Trans-Atlantic RPS Working Group. Management of Primary Retroperitoneal Sarcoma (RPS) in the adult: a consensus approach from the Trans-Atlantic RPS Working Group. *Ann Surg Oncol*. 2015;22:256–263.
2. Gronchi A, Miceli R, Shurell E, et al. Outcome prediction in primary resected retroperitoneal soft tissue sarcoma: histology-specific overall survival and disease-free survival nomograms built on major sarcoma center data sets. *J Clin Oncol*. 2013;31:1649–1655.
3. Tan MC, Brennan MF, Kuk D, et al. Histology-based classification predicts pattern of recurrence and improves risk stratification in primary retroperitoneal sarcoma. *Ann Surg*. 2016;263:593–600.
4. Lokuhetty D, White VA, Cree IA, eds. *WHO Classification of Tumours Editorial Board Soft Tissue and Bone Tumours*, 5th ed. Lyon (France): World Health Organization; 2020.
5. Trojani M, Contesso G, Coindre JM, et al. Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int J Cancer*. 1984;33:37–42.
6. Gronchi A, Strauss DC, Miceli R, et al. Variability in patterns of recurrence after resection of primary retroperitoneal sarcoma (RPS): a report on 1007 patients from the Multi-institutional Collaborative RPS Working Group. *Ann Surg*. 2016;263:1002–1009.
7. Nussbaum DP, Rushing CN, Lane WO, et al. Preoperative or postoperative radiotherapy versus surgery alone for retroperitoneal sarcoma: a case-control, propensity score-matched analysis of a nationwide clinical oncology database. *Lancet Oncol*. 2016;17:966–975.
8. Haas RLM, Bonvalot S, Miceli R, et al. Radiotherapy for retroperitoneal liposarcoma: a report from the Transatlantic Retroperitoneal Sarcoma Working Group. *Cancer*. 2019;125:1290–1300.
9. Bonvalot S, Gronchi A, Le Péchoux C, et al. Preoperative radiotherapy plus surgery versus surgery alone for patients with primary retroperitoneal sarcoma (EORTC-62092: STRASS): a multicentre, open-label, randomised, phase 3 trial. *Lancet Oncol*. 2020;21:1366–1377.
10. DeLaney T, Mullen JT, Wang D, et al. Preoperative radiotherapy for retroperitoneal sarcoma. *Lancet Oncol*. 2021;22:e1.
11. Chowdhary M, Spraker MB. Preoperative radiotherapy for retroperitoneal sarcoma. *Lancet Oncol*. 2021;22:e2.
12. Izzuddeen Y, Sharma DN. Preoperative radiotherapy for retroperitoneal sarcoma. *Lancet Oncol*. 2021;22:e3.
13. Bonvalot S, Gronchi A, Le Péchoux C, et al. Preoperative radiotherapy for retroperitoneal sarcoma—authors' reply. *Lancet Oncol*. 2021;22:e4.
14. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika*. 1983;70:41–55.
15. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Intern Med*. 1997;127(pt 2):757–763.
16. Haas R, Stelmes JJ, Zaffaroni F, et al. Critical impact of radiotherapy protocol compliance and quality in the treatment of retroperitoneal sarcomas: results from the EORTC 62092-22092 STRASS trial. *Cancer*. 2022;128:2796–2805.
17. Singer S, Antonescu CR, Riedel E, et al. Histologic subtype and margin of resection predict pattern of recurrence and survival for retroperitoneal liposarcoma. *Ann Surg*. 2003;238:358–370. discussion 370–371.
18. Guillou L, Coindre JM, Bonichon F, et al. Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. *J Clin Oncol*. 1997;15:350–362.
19. Gronchi A, Miceli R, Allard MA, et al. Personalizing the approach to retroperitoneal soft tissue sarcoma: histology-specific patterns of failure and postrelapse outcome after primary extended resection. *Ann Surg Oncol*. 2015;22:1447–1454.
20. MacNeill AJ, Miceli R, Strauss DC, et al. Post-relapse outcomes after primary extended resection of retroperitoneal sarcoma: a report from the Trans-Atlantic RPS Working Group. *Cancer*. 2017;123:1971–1978.
21. Gronchi A, Lo Vullo S, Fiore M, et al. Aggressive surgical policies in a retrospectively reviewed single-institution case series of retroperitoneal soft tissue sarcoma patients. *J Clin Oncol*. 2009;27:24–30.
22. Gronchi A, Miceli R, Colombo C, et al. Frontline extended surgery is associated with improved survival in retroperitoneal low- to intermediate-grade soft tissue sarcomas. *Ann Oncol*. 2012;23:1067–1073.