Reviewer A Comments:

Reply to Reviewer A:

Thank you for giving us this opportunity to revise our manuscript. We appreciate your positive comments and assistance. We have tried our best to improve the manuscript, and we hope that you will find it suitable for publication.

Comment 1: The paper is well organized and well written. However, I suggest to update References list.

Reply 1: Thanks for your kind comment and all your help. According to your suggestion, we have appropriately updated the references list.


**Reviewer B Comments:**

“This is a pertinent and relevant manuscript regarding a rare disease, which lends itself well to large registry analysis…”

Reply to Reviewer B

Thank you for your positive comments. We are grateful for the constructive comments and suggestions about our manuscript. Your comments are valuable and the suggestions have improved our paper and provided better perspectives about our study. We have studied your comments carefully and made the required corrections.

Comment 1: Large registries, including SEER have significant limitations. In particular, SEER has highly limited data regarding chemotherapy and radiation therapy, and only includes those used as part of front-line therapy. This should be expanded upon in the discussion, as this greatly curtails the conclusions that can be drawn

Reply 1:

Thank you for this suggestion. As you say, SEER, as a large cancer database, has some limitations, especially the lack of some important treatment information, such as the mode of surgery, the course of radiotherapy, chemotherapy drugs, doses, and so on. In accordance with your suggestion, we have added these limitations to the discussion sections.

(Changes in the text: Discussion section, Page 16-17, Lines 310-316.)
Comment 2: SEER does not have histologic adjudication, which somewhat questions the reliability of the data. The inclusion of a 2 year-old patient in the analysis adds to such questions. The analysis should probably be limited to adults > 18 years.

Reply 2:

Thank you very much for your valuable comments. There are some limitations in identifying the histological classification based on the SEER database. We apologize for not clarifying how the histological type was determined. SEER contains ICD-O-3 codes (International Classification of Disease for Oncology: https://apps.who.int/iris/handle/10665/96612; http://www.iacr.com.fr/index.php?option=com_content&view=category&layout=blog&id=100&Itemid=577). These ICD-O-3 codes include information on both the topography (site and subsite) and morphology (histology and behavior grade or tumor/cell type behavior differentiation) of the tumor. Moreover, morphological features include 576 types and 592 categories. Hence, we used the ICD-O-3 code 8811/3 to ensure the histological type. Other authors have previously used this method [1-3].

According to your suggestion, we checked the age distribution and found that there were 36 patients under 18 years old. Previously, we wanted to reduce censored data in our analysis of OS, so we did not exclude young patients (<18 y). However, we found that many authors agree with your conclusion [4-6]. Therefore, we have added age as an exclusion criterion (<18 y) and reanalyzed the results. The figures and tables have all been revised, and we have revised the entire manuscript to account for this updated analysis.

Changes in the text:

#1: (Figure section, Figure 1-8);
#2: (Table section, Table 1-4; Added supplementary table: S1-2);
#3: (Changes in the text: Methods section, Page 8, Lines 123-125).

References:
1. Fakhri B, Fiala MA, Tuchman SA, Wildes TM: Undertreatment of Older Patients


Comment 3: Whenever the training and validation cohorts are drawn from the same database or registry, there are concerns over the generalizability. The large registry used here helps mitigate that, as well as the non-overlapping populations, though the concerns still exist and should be mentioned in discussion.

Reply 3:

Thank you for this excellent comment. We agree. However, we believe that the large amount of data in the SEER database can dilute adverse effects of dividing data from the same database into a training cohort and a validation cohort. We specifically emphasized this issue in the discussion.
Comment 4: It is entirely unclear why the authors compared this nomogram to AJCC6 rather than AJCC8. Seeing as AJCC is predominantly a prognostic staging system, it is unclear why the authors would choose to use an older iteration rather than the most up-to-date and prognostically significant version.

Reply 4:

Thank you for this question. The sixth edition of the AJCC Cancer Staging Manual was published in 2002 [1], the seventh edition was published in 2009 [2], and the eighth edition was published in 2017 [3]. We included patients diagnosed from 2004 to 2015. Patients diagnosed in the earlier period only had data corresponding to that of the sixth edition and patients diagnosed later had data corresponding to the eighth edition; therefore, we would have had missing or censored data if we had used the eighth edition. MFS is a rare disease and only 1270 patients in the SEER database in this time period met our criteria. Because the incidence of MFS is extremely low, more censored data would have affected the establishment of a prediction model. Although we would have liked to have manually converted the AJCC sixth edition to the eighth edition, it would have been biased. Hence, we used the sixth edition.

References:

Comment 5: It is unclear why, and how the authors utilized the decision curve analysis. There are no decisions that are indicated by the retrospective prognostic nomogram development. Similarly, the abstract and discussion suggests that the current analysis affects treatments. There is no indication from the current study that there is any treatment consideration that can be gleaned from this data. This should either be very well explained, or more likely would best be removed from the manuscript. Either way, and a mention in the abstract or discussion of its effect on treatment should be removed.

Reply 5:

Thank you for your comment. In this study, DCA was used to compare the clinical applicability of the nomogram and the AJCC staging system. We used ROC curve analysis and C-indexes to evaluate the accuracy of the nomogram, but these cannot tell us whether the model is worth using or which of two or more models is preferable. A false-negative result can be much more harmful than a false-positive result (1-2). Therefore, DCA was developed to assess the clinical usefulness of a diagnostic test, marker, or predictive model and considers a weighted sum of true and false positives (1-3). DCA can also be used to assess the value of prognostic models (1). In our specific example, we assume that patients receive an intervention when the expected 3-year survival rate is >40% and refuse the intervention when it is <40%. In the training cohort, if the nomogram is used to predict which patients will survive for 3 years, approximately 6 additional patients per 100 are identified. Therefore, these 6 patients can then receive intervention. The AJCC only identifies 2 of these patients. However, this study can only show that the nomogram is superior to the AJCC staging system in clinical applicability. As you stated, we cannot draw a conclusion from DCA that the model can guide treatment, because the specific decision is not indicated. We have deleted and revised some content to reflect this.

(Changes in the text: Abstract section, Page 4, Lines 58-59; Results section, Page 13, Lines 241-243; Discussion section, Page 14, Lines 261-262; Conclusions section, Page
References:


Comment 6: The authors correctly assert that overall survival is a significant problem. However, one of the particular nuances of managing MFS is its uncommonly high predilection for local recurrence. That is why most prior studies of this particular histologic subtype have focused on disease-free survival or local recurrence free survival. While we acknowledge that overall survival is perhaps the most important, SEER is not well designed to answer this question. Of all of the soft tissue sarcoma subtypes, this one is perhaps most influenced by local recurrence, and the role of local recurrence on overall survival is not addressed in the nomogram. This is a major limitation of the current analysis.

Reply 6:

Thank you for this comment. We agree that local recurrence is important in MFS. However, the purpose of our research is to fill the research gap or deficiencies in predicting OS in patients with MFS. Therefore, examining local recurrence-free survival or disease-specific survival as outcome variables would answer different questions. Many authors are using the SEER database to study OS in patients with different subtypes of sarcoma [1-3]. Moreover, some variables related to OS are not necessarily related to local recurrence, and the relationship between local recurrence and OS is uncertain. Regarding a nomogram to predict local recurrence, we would like to discuss the relationship between local recurrence and related independent variables in a future article. Further, because MFS is a rare disease, the data are not particularly
complete (there are some censored data). We explained this limitation in the discussion. The outcome variables of the novel nomogram were 3-year and 5-year OS. We have described it in the text, but we are very sorry that it is not clearly described in the figure. Therefore, we have re-made the nomogram (Figure 4 shows the 3-year and 5-year OS probability).

(Changes in the text: Discussion section, Page 17, Lines 316-319.)

References:


Comment 7: As with many registry studies, there is a very high number of patients that are excluded due to incomplete data. This limits the voracity of the conclusions that can be drawn.

Reply 7:

As you said, we excluded many patients because of the strict inclusion and exclusion criteria, resulting in a 43% reduction in sample size, which to some extent limits the voracity of the conclusions that can be drawn. We emphasized this issue in the limitations.

(Changes in the text: Discussion section, Page 17, Lines 319-321.)

Comment 8: If using AJCC6 as a comparison, it is unclear why the authors utilized a 100 mm cut off for size. The authors should provide some analysis as to why this
particular cutoff is utilized. In the current manuscript, this otherwise appears somewhat arbitrary as a line of distinction. AJCC6 utilizes a 5 cm cutoff, and AJCC8 utilizes a tiered system of 5, 10, and 15 cm.

Reply 8:

Thank you for your excellent comment and advice. We recoded the “Tumor Size” variable according to the cutoff value (5 cm). We also reanalyzed tumor size in the univariate analysis ($p < 0.05$) and multivariate analysis ($p = 0.02$).

*(Changes in the text: Methods section, Page 8, Lines 140-141.)*

Comment 9: The decision as to which patients are going to undergo frontline chemotherapy, and therefore included in SEER as having undergone chemotherapy is likely a reflection of high risk, and resultant poor prognosis. As a result, chemotherapy may not be a prognostic factor in and of itself, but rather a reflection of other prognostic factors. This cannot be adequately remedied by the current analysis, but must be addressed in the manuscript.

Reply 9:

Thank you for your valuable comment. We re-emphasized the limitations of our conclusions based on chemotherapy administration in accordance with your comments. However, in our study, the results of multivariate Cox regression analysis showed that chemotherapy was a significant predictor of poor outcomes ($p < 0.05$, HR > 1). Moreover, the VIF was < 5, indicating there was no obvious collinearity with other variables. Therefore, we speculated on the possible reasons for the poor prognosis associated with chemotherapy. However, as you say, we cannot examine chemotherapy alone in the current analysis. Thorough examination of the effect of chemotherapy on survival requires a randomized controlled trial or propensity score matching analysis to balance the effects of baseline clinicopathological differences [1]. We included this limitation in the discussion and envisioned future research topics.

*(Changes in the text: Discussion section, Page 16, Lines 304-310.)*

Reference:

Comment 10: It appears as though the authors utilized a Kaplan Meier analysis as a univariate screen for variables to be included in their multivariable analysis. However, it is unclear why her age was not included in the Kaplan Meier analyses. In addition, given that socioeconomic status and inclusion of radiation were not demonstrated to have a significant effect on Kaplan Meier analysis, it defies logic as to why these were included in the multivariable analysis. The multivariable analysis should probably be revised, utilizing only those variables identified by a predefined screen of either Kaplan Meier or formal univariate analysis.

Reply 10:
Thank you for your valuable advice. The Kaplan-Meier method and log-rank test were used to perform univariate prognostic analysis [1], as this could show whether the curves crossed during follow-up. Some authors choose to process continuous variables as categorical variables for univariate analysis [2]. However, we did not want to artificially convert continuous variables into categorical variables. Therefore, in accordance with your suggestions, we performed univariate Cox regression as a formal univariate analysis and analyzed age and other variables (novel univariate analysis, Table 2). After revising the exclusion criteria, we conducted a univariate analysis and found that SES was still not statistically significant. Therefore, as you state, we have no direct evidence that SES affects the prognosis of MFS, and SES was not included in the multivariate analysis.

(Changes in the text: Abstract section, Page 3, Lines 40-41; Methods section, Page 9, Lines 149-150; Results section, Page 11, Lines 199-204.)

Selecting statistically significant variables in univariate analysis and then including those variables in the multivariate analysis is a common statistical method. However, the actual or clinical significance of the variables should also be considered with regard to their inclusion in the multivariate analysis. MFS has high rates of relapse after surgery; therefore, local adjuvant treatment may be beneficial. However, the clinical
benefit of RT in MFS or soft tissue sarcomas in the extremities remains controversial [3,4]. Some authors reported that RT was associated with a better outcome [5,6]. Therefore, adjuvant RT has been suggested for such patients [7]. Based on our clinical analysis and these previous studies, we believed that RT may have had an impact on prognosis. Many authors suggest that extremity myxofibrosarcoma is not clinically inherently radioresistant (a previous study reported $p = 0.157$ for RT, and in our univariate analysis, $p = 0.16$ for RT) [8]. Therefore, we incorporated RT in the multivariate analysis, which confirmed our hypothesis that RT had an impact on OS. Here, we focus on the conclusions of the multivariate analysis rather than those of the univariate analysis.

Another point of view has been suggested: under the premise that the sample size is sufficient, multivariate analysis can also be performed directly. We believe that the results of large sample analysis have a certain reliability. Our statistician gave the same suggestion: the inclusion criteria should be relaxed, such that if $p$ is less than 0.1 or 0.2 in the univariate analysis, more variables can be included in multivariate analysis. The nonsignificant $p$ value for RT in the univariate analysis (0.16) may be because the interaction between independent variables is not considered. Therefore, we believe that incorporating RT into the multivariate analysis is valid. However, to thoroughly analyze the impact of RT alone, propensity score matching or a randomized controlled trial should be conducted.

(Changes in the text: Discussion section, Page 16, Lines 300-304.)

References:


3. Koshy M, Rich SE, Mohiuddin MM: Improved survival with radiation therapy in


Comment 11: Despite the authors' assertion, the demonstration of a prognostic significance of socioeconomic status on overall survival on the multivariable analysis in extremity soft tissue sarcoma has been established in the medical literature. While that is likely that the authors are correct that this may not have been proven specifically for MFS previously, the authors' suggestion of novelty of this analysis is otherwise incorrect.

Reply 11:

Thank you very much for your suggestion. We agree. In accordance with your previous comments, we revised the exclusion criteria and performed the univariate analysis again. In this re-analysis, the p-value for SES was still > 0.05 (univariate COX moderate p = 0.285, high, p = 0.683; KM, p = 0.561). Moreover, there is no previous
evidence indicating that SES influences OS in MFS patients. Therefore, we did not include in the multifactor analysis in this revision and deleted the related sections of text.

(Changes in the text: Discussion section, Page 15, Lines 280-283.)

Comment 12: From a practical standpoint, seeing as grade is an integral part of stage, the finding that grade is not independent of stage in its prognostic significance is not surprising. The finding that grade is not a prognostic factor is not entirely correct. Rather, grade has been shown repeatedly to be an important prognostic factor, which is why it is included in staging. If the authors want to determine the true factors that are associated with prognosis, in order to identify a new staging system, it would likely best be performed by either 1) eliminating stage in favor of its individual components; or 2) eliminating the components that make up stage, utilizing stage itself. Including both in the multivariable analysis defies logic and clouds the picture, presenting a confusing and inaccurate portrait of what actually is prognostically significant.

Reply 12:

Thank you for this comment. We have re-emphasized tumor grade, which is indeed an important prognostic factor in many subtypes of soft tissue sarcoma. As you state, simple multivariate analysis is not perfect. Therefore, after conducting a multivariate analysis including grade, we conducted a collinearity analysis and found that the VIF (variance inflation factor) of tumor grade was < 5. However, based on the multivariate analysis, tumor grade is not an independent variable ($p > 0.05$).

<table>
<thead>
<tr>
<th>Grade</th>
<th>Univariate analysis ($p$)</th>
<th>Multivariate analysis ($p$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td>II</td>
<td>0.314</td>
<td>0.218</td>
</tr>
<tr>
<td>III</td>
<td>&lt;0.001</td>
<td>0.211</td>
</tr>
<tr>
<td>IV</td>
<td>&lt;0.001</td>
<td>0.197</td>
</tr>
</tbody>
</table>

Moreover, in many studies, the elements in stage, such as tumor grade, have been similarly analyzed in parallel [1-4]. We thought that the AJCC stage is a more
comprehensive factor than the tumor grade, and the VIF (Grade) between them was < 5. Therefore, we believe that specific tumor subtypes should be analyzed in detail. This may be one of the reasons why the application or prognostic evaluation of AJCC stage in specific tumor subtypes is not as good as the novel established nomograms [5-6].

(Changes in the text: **Discussion** section, Page 15-16, Lines 289-294.)

Reference: