Peer Review File

Article information: https://dx.doi.org/10.21037/jtd-23-346

Reviewer A

Udelsman et al. assessed the efficacy of a collaborative treatment model compared with a single institution model for trimodal management of esophageal cancer using the NCDB database. The median travel distance was two times longer in the collaborative treatment group than in the single institution model. Although the patients in the collaborative group were less likely to receive guideline-recommended multiagent chemotherapy, they had an increased rate of pCR at the time of esophagectomy Five-year and 90-day mortality rates were comparable between the groups. Based on these findings, the authors concluded that collaborative trimodal treatment is a common and reasonable practice model. This manuscript is well-written and informative and will provide useful clinical information for readers. Several changes will improve the quality of this manuscript.

Comment 1: A critical problem is that this study included only the patients who completed trimodality treatment. In other words, there may be a difference in the proportion of patients who could not undergo esophagectomy due to a failure of neoadjuvant CRT. This might be the reason for the higher pCR rate despite the lower chemotherapeutic intensity in the collaborative group.

Reply 1: We agree that this is most significant limitation to our work and deserves additional comment in our discussion section. Unfortunately, we cannot determine the reason patients may have failed neoadjuvant treatment. We also do not know how many patients may have intended to undergo neoadjuvant treatment and then defer surgical resection in favor of observation. Despite this inherent flaw, we do think that the presented data is important and meaningful. Currently, the data supports a multi-institutional model and shows that at least for individuals that reach operative resection there is no clinically meaningful differences in survival. The large database analysis shows what may be possible, but the practicalities involved, especially the identification of patients who initiate but then fail treatment will require further study. Given the nuanced nature of these decisions we believe these questions may require a more focused regional approach (as opposed to large national database) and will be the focus of future works. We have detailed these points in our revised discussion section that is excerpted below.

Changes in text 1: “Limitations of this study include those commonly associated with large retrospective observational studies. We are unable to account for differences in treatment regimens or why certain patients did or did not receive a given therapy. Most importantly, we are unable to accurately identify patients who may have started neoadjuvant treatment but then failed to proceed to surgery. We do not know if these events occurred due to treatment complications, patient preference, or a decision to defer surgery in favor of clinical monitoring. We also do not know the reasons a patient elected for treatment in a single-institution or a collaborative model, or how this may have affected their trust and satisfaction with a given healthcare network. Differences in staging modality and accuracy between patients in the collaborative and single-institution treatment groups could have further biased this study. Unfortunately, the NCDB lacks

the granularity to satisfactorily answer these questions, and this will the subject of focused qualitative and quantitative studies."

Comment 2. Isn't the R0 resection rate available in the NCDB? That is an important variable to compare both models.

Reply 2: The reviewer is correct, and we agree this is an important metric to document. The R0 resection rate in the overall cohort was 94.2%. In the single institution treatment group 93.7% had an R0 resection and in the collaborative treatment group 94.4% had an R0 resection (p-value 0.17). We have added this information to our results section (excerpted below) and to the revised figure 2 as well as supplemental table 1.

Changes in Text 2: "There was no significant difference in R0 resection between patients treated in a collaborative treatment and single-institution treatment (94.4% vs. 93.7%; p=0.17)."

Comment 3: Page 8, Line 205: cPR should be pCR.

Reply 3: Thank you for catching. We have made the indicated correction.

Change in Text 3: Multimodal therapy, R0 resection, pCR, and 90-day mortality

Reviewer B

Comment 1: The authors are to be complimented. This is an interesting and relevant analysis of the treatment of esophageal cancer patient in a network. this is an actual way of treating patients in European countries and the outcome is therefore widely applicable.

Reply 1: Thank you, we appreciate the review. We agree that compared to much of Europe the US is behind in terms of coordination of care.

Change in text 1: None indicated.

Reviewer C

Comment 1: It is unclear to me if the patients were initially involved in the decision to be treated at a single institution or at multiple institutions. This is important, because if patients had a preference to be treated at their local hospital or rather at a hospital further away (because they had more trust in that specific hospital), this also plays a role in their satisfaction.

Reply 1: This is a critical point and one that should be better emphasized in our discussion. Unfortunately, within the NCDB we do not know the specific reason why patients chose single

institutional treatment vs. multi-institutional treatment. This a limitation of this work and we have expanded upon it in our discussion/limitations section. We do think that understanding the reason behind patient decision making is critical, and based on this question we are planning additional qualitative and quantitative studies.

Change in Text 1: "We also do not know the reasons a patient elected for treatment in a single-institution or a collaborative model, or how this may have affected their trust and satisfaction with a given healthcare network. Differences in staging modality and accuracy between patients in the collaborative and single-institution treatment groups could have further biased this study. Unfortunately, the NCDB lacks the granularity to satisfactorily answer these questions, and this will the subject of focused qualitative and quantitative studies."

Comment 2: The fact that patients who received collaborative treatment were more often originated from non-academic center, does seem logical to me. Maybe those patients who origins from non-academic centers do specifically prefer an academic center (status) and therefore they preferred to be enrolled in collaborative treatment.

Reply 2: We agree this is certainly possible. We have expanded on this possibility as well as the need for further qualitative studies in our discussion.

Change in text 2: "In the treatment of pancreatic adenocarcinoma—a similar cancer requiring coordination of neoadjuvant therapy and surgical resection—sociodemographic barriers are a significant hurdle to patients receiving guideline recommended treatment, most pronounced in under-represented minorities and those of lower socioeconomic status (28,29). These factors may also contribute to distrust for specific healthcare networks and influence a patient's choice to receive single-institution or collaborative treatment. Indeed, we found that patients who received collaborative treatment were more likely to originate from non-academic centers. Further qualitative analysis will be necessary to develop an understanding behind these treatment-decision and will be the focus of ongoing studies."

Comment 3: The results show that patients who received collaborative treatment traveled almost twice as far- and therefore local treatment may prevent travel burden. However, in the conclusion the authors state that collaborative care decreases travel time for patients. This seems contractionary.

Reply 3: The patients in the collaborative treatment group traveled almost twice as far to the site of surgery. However, by receiving their CRT at local centers that collaborated with the operative center they had the opportunity to potentially reduce this travel burden. We have clarified this message in the body of the manuscript.

Change in text 3: "Collaborative care is a reasonable treatment model that can allow patients to receive CRT at local centers and alleviate barriers to accessing guideline recommended care."

Comment 4 and 5: Personally, i would appreciate the conclusion to be clearer: how can these study results be translated to clinical practice today? should we change anything? can

we give recommendations? The statement in the conclusion (last sentence): 'safe delivery .... collaborative care', is maybe a bit awkwardly formulated. Isn't it that safe delivery of treatment is critical for all patients anyway everywhere? and not especially for a subgroup of patients? What do you mean with 'disproportionally impacted' by travel distance?

Reply 3 and 4: We appreciate the reviewer's critique and agree the conclusion can be improved. We have rewritten the conclusion and have excerpted it below. Have sought to both offer more definitive guidance and to clarify the language.

Change in text 4 and 5: "Multiple institutions commonly collaborate to deliver trimodality treatment to patients with esophageal cancer without an overall reduction in R0 resection, pCR, 90-day survival, or long-term survival. Collaborative care is a reasonable treatment model that can allow patients to receive CRT at local centers and alleviate barriers to accessing guideline recommended care. Collaborative deliver of trimodality treatment should be encouraged as it may allow more patients with esophageal cancer to access and benefit from guideline recommended therapy."

Reviewer D

The authors have written an interesting manuscript on collaborative (multi-institutional) versus single-institutional treatment for patients with non-metastasized esophageal cancer who are treated with neo-adjuvant chemoradiation and resection. I do however have some critical points that should be addressed.

Comment 1: The authors state that they intend to compare compliance with guideline recommendations, but they only select patients that did receive both neo-adjuvant chemoradiation and surgery. However, by doing so they do not include patients that underwent chemoradiation with neo-adjuvant intent but that for some reason did not proceed to surgery and they do also did not include patients with a similar clinical stage but with other types of treatment (definitive chemoradiation or no treatment with curative intent). It is therefore hard to make a conclusion on whether clinical guidelines were followed in the same way in patients treated in one or multiple centers. As selection on which patients proceed to the currently included treatment types might differ between the two groups compared in the current study. The same applies to excluding incomplete clinical staging, this might differ according to where a patient is diagnosed and therefore whether the patient would have been treated in a single center or multiple centers.

Reply 1:

We appreciate the review of our work and these important points. The reviewer is correct in that there is potential bias in the selection criteria for this study. Unfortunately, the NCDB does not provide the granularity to assess why a patient received a certain treatment. We do not know how many patients intended to undergo neoadjuvant treatment and then deferred surgical resection in favor of observation.

Despite this inherent flaw, we do think that the presented data is important and meaningful. Currently, the data supports a multi-institutional model and shows that a least for individuals that reach operative resection there is no clinically meaningful difference in survival. This large database analysis shows what may be possible, but the practicalities involved, especially the identification of patients who initiate but then fail treatment will require further study. Given the nuanced nature of these decisions we believe these questions require a more focused regional approach (as opposed to large national database) and will be the subject of future works. We have detailed these points in our revised discussion section that is excerpted below.

Change in text 1: "Limitations of this study include those commonly associated with large retrospective observational studies. We are unable to account for differences in treatment regimens or why certain patients did or did not receive a given therapy. Most importantly, we are unable to accurately identify patients who may have started neoadjuvant treatment but then failed to proceed to surgery. We do not know if these events occurred due to treatment complications, patient preference, or a decision to defer surgery in favor of clinical monitoring. We also do not know the reasons a patient elected for treatment in a single-institution or a collaborative model, or how this may have affected their trust and satisfaction with a given healthcare network. Differences in staging modality and accuracy between patients in the collaborative and single-institution treatment groups could have further biased this study. Unfortunately, the NCDB lacks the granularity to satisfactorily answer these questions, and this will the subject of focused qualitative and quantitative studies."

Comment 2: There is a very large number of patients that is excluded due to incomplete clinical stage (n=3359), incomplete pathologic staging (n=3916) and treated outside NCDB reporting center (n=2494). Especially the incomplete pathologic staging is surprising, as you would expect complete pathologic staging for nearly every patient that has underwent major surgery such as an esophagectomy. They authors should discuss how the large number of patients that are excluded could have affected the outcomes of the current study.

Reply 2: We attempted to follow a rigorous selection process in order to accurately compare survival and pCR. Unfortunately, this did result in exclusion of a significant number of patients. We have expanded on this as a limitation. While we do not believe it effects the validity of our work, it is critical to note how it effects is generalizability. We appreciate the reviewer's insight and for bringing up this point.

Change in text 2: "There is also a significant amount of missing clinical and pathologic staging data. Exclusion of patients with missing data may bias these results and reduce their generalizability."

Comment 3: I am not familiar with NCDB registration procedures, but could the 2494 patients excluded due to treatment outside of NCDB reporting center be mainly patients that would have been treated in multiple centers? Because, it could have seriously impacted the outcomes of the study if exclusion is predominantly in one of the two groups of the study (single versus multi-institutional treatment).

Reply 3: This is an important point and limitation of this study. We do not know what aspect of care was delivered to patients who received care outside an NCDB center. However, we have performed a sensitivity analysis among this group of 2494 patients. We found that receiving care outside a NCDB center 1015 (41%) were designated in the single-institution group while 1,479 (59%) were designated in the collaborative treatment group. This very well approximates the breakdown among the main cohort of 39% and 61%, respectively. We have expanded on this in our limitations section.

Change in text 3: "We are unable to account for care that may have occurred outside of a NCDB reporting center and was not captured in the database. Among patients excluded due to care outside of an NCDB center we saw a similar ratio of patients receiving single-institution and collaborative treatment."

Comment 4: The authors correctly discus in both the introduction and discussion section, that there is quite some evidence that there is an association between improved surgical quality and outcomes to the resections done in higher volume centers. However, unfortunately this study does not take the annual resection volume of the hospitals into account, which is critical problem of this study. As the resection volume could have affected the primary outcomes of the study, but also differ between the two groups that are compared in this study (single versus multi-institutional treatment).

Reply 4: This is an important point and have added this analysis to our manuscript. In brief, we set a cutoff of 20 esophagectomies per year to differentiate low-volume vs. high-volume centers based on previous work by Metzger et al. Using this definition 70.8% of the patients were treated at a high-volume center. Patients treated at a single institution where slightly more likely to receive surgery at a high-volume center (72.5% vs. 69.7%; p=0.01). In multivariable logistic regression evaluating likelihood of receiving care in a collaborative model (table 2), there was a trend toward significance, but it did not cross the threshold of 0.05. We now include annual hospital volume in our Cox regressions. While the exact Hazard ratios and confidence intervals change slightly there is no significant difference in the primary trends we have identified. We suspect that treatment in collaborative model allows patients to receive CRT at local centers and then travel to high-volume centers for esophagectomy, which is the reason the groups appear so similar.

Change in Text 4:

Methods: "Based on work by Metzger et al., an annual esophagectomy volume of 20 was used to identify high-volume centers (8)."

Results (Patient Characteristics): "Slightly more patients in the single-institution group received an esophagectomy at a high-volume center (72.5% vs. 69.7%; p=0.01)."

Results (Multivariable Logistic Regression): "Esophagectomy at a high-volume center trended toward significance but did not cross the predetermined threshold for significance."

Results (5-Year Survival and Cox-Proportional Hazard Model): "Variables associated with worse survival included older age, male sex, lack of insurance, increasing Charlson-Deyo comorbidity index, esophagectomy at a low-volume center, and increasing clinical stage (**Table 3**)."

Please see revised tables for full details.

Comment 5: In addition, it would also be worthwhile to know whether the hospital in which a patient is diagnosed with esophageal cancer does perform esophagectomies and if so, what the annual resection volume is.

Reply 5: Unfortunately, we can only identify the hospital at which the patient received the esophagectomy but do not have reliable information on the hospital at which the patient was diagnosed. We have added this a limitation to our work.

Change in text 5 (Limitations): "We do not know the quality of staging, staging modality, or location at which cancer was first diagnosed. These factors may have differed between the study groups and biased the study."

Comment 6: In the tables the authors could be clearer that factors associated with 'facility' are related to the resection center and not the chemoradiotherapeutic center.

Reply 6: We apologize for the confusion and agree clarification is needed. Please see the excerpt from the revised table legend below.

Change in text 6 (excerpt from revised tables legend): Facility refers to the institution in which the esophagectomy was performed

Comment 7: The authors mention both 5-year mortality and 5-year survival in the manuscript. I would suggest to only report on 5-year survival to prevent any confusion about possible differences in the definitions between mortality and survival.

Reply 7: We agree this add unnecessary confusion. We have revised the manuscript to describe 5-year survival uniformly.

Change in text 7: We have changed 5-year mortality to 5-year survival throughout the text.