Peer Review File

Article Information: https://dx.doi.org/10.21037/jgo-22-1005

Reviewer A

This is very interesting paper. I ask some questions to author.
Comment 1: Please explain why IC-CRT is platinum+taxane17 case (29%) 5Fu+Taxane 40 case (68%), while CRT is platinum+taxane 34 case(94%) and 5Fu+Taxane 2case (6%) in the current chemotherapy regimen.

Reply: We state in the results that the only significant difference time between the IC-CRT and CRT cohorts are treatment era (Table 1). In the discussion, we explain that before the CALGB 80803 data was presented, the primary regimen for locally advanced esophageal cancer (LA-EC) was based on the CROSS regimen which utilized carboplatin/paclitaxel-based CRT without induction. Following CALGB 80803 data, our institutional practice shifted to more routinely utilizing IC-CRT with 5FU/oxaliplatin-based induction.

Comment 2: In figure 2, progression free survival (2A) and overall survival (2B) for patients with adenocarcinoma, the IC-CRT group is superior to the CRT group, and both are good data with P = 0.16 p = 0.12. However, for SCC cases in Figure 3, the RCT group had better data than the IC-RCT group, with P = 0.28 and P = 0.01 for both. Please explain the reason for the contradictory data for adenocarcinoma and SCC.

Reply: We are happy the reviewer has noticed this important nuance in our paper. There has been work demonstrating the impact of overall treatment time (OTT) on local control and overall survival for SCC of the head and neck cancer, cervix, and anus. We hypothesize a similar interaction occurs for SCC of the esophagus and contributes to the differential impact of IC-CRT between SCC and AC-EC. We have added a sentence to the discussion to specifically articulate this.

Changes in the text:
Results: Among patients with SCC (n=18), *the PFS numerically favored CRT though did not reach statistical significance* (6.4 months IC-CRT vs. 56.6 months CRT, p=0.28), however, OS was significantly shorter for those in the IC-CRT vs. CRT cohorts (17.4 months vs. 63.0 months, p=0.01). (eFigure 3)

Discussion: "IC-CRT is not routinely utilized in patients with SCC of the esophagus, as in many other disease sites, the use induction chemotherapy and prolongation of the overall treatment time (OTT) has been detrimental to OS. 18-20 *We hypothesize the prolonged OTT (approximately 6 weeks for CRT versus 12 weeks for IC-CRT) affects SCC more significantly than AC-EC and is one of the reasons for the stark difference in outcomes between the two histologies*."

Reviewer B

Peters and colleagues from Yale University have submitted a retrospective review of 95 patients who were deemed to be eligible candidates for trimodality therapy for treatment of esophageal cancer at their institution between 2013-2019. The purpose of the study was to compare outcomes between patients who were treated according to the CROSS regimen (n=36) and patients who were treated with a protocol based on CALGB 80803 with induction Folfox chemotherapy followed by chemoradiotherapy using Folfox (n =59). The patients were not randomized. treatment decisions made based on physician preference with a shift toward an induction approach over time.

Important differences exist between this cohort and the CALGB 80803 study. All patients in the 80803 study had adenocarcinoma, whereas 18 of the 95 patients in this study had squamous cell cancer. In addition, as the authors point out, patients in this study were not evaluated by PET after completing induction therapy to see if they responded before moving on to chemoradiotherapy. Notably, Folfox was used as the induction regimen in 76% of the patients in the induction arm in this study. Furthermore, in the induction group, approximately 73% patients received 3 cycles of induction chemotherapy, whereas the median percentages of protocol induction doses were 99% in 80803. Finally, only 58% of patients went on to esophagectomy, compared to 85% in 8083.

With these important differences in mind, the authors found no significant differences in terms of overall or progression-free survival between the induction group and the CROSS group. This finding persisted even in subgroup analyses that were performed to try to closer approximate the 80803 cohort (Adenocarcinoma only, received 3 cycles of induction tehrapy). There were, however, improved rates of pCR in the induction group, despite a lower percentage of patients with squamous cell cancer in this group compared to the CROSS group. This difference was most pronounced in patients with adenocarcinoma who received 3 cycles of induction Folfox (58% v 20% p=.02). Importantly, the need for feeding tube placement was higher in the induction group, whereas overall surgical complications were not significantly different. Interestingly, patients with squamous cell cancer who underwent induction regimens seem to do worse than those treated with the CROSS regimen.

This study does provide some important real world results utilizing the Folfox induction strategy in no-selected patients. Its main weaknesses are the relatively low number of patients in each group, the heterogeneity in the induction group, and the low number of patients going on to esophagectomy. The authors attempt to deal with this heterogeneity with subgroup analyses focusing on patients with adenocarcinoma who received 3 cycles of induction Folfox and underwent esophagectomy to attempt to get closer to the 80803 population. Unfortunately, this further reduces the numbers in each group and limits the robustness of the findings.

Suggestions:
Comment 1: Given the number of subgroups analyzed, it would be helpful to include a table

comparing median PFS and OS between induction therapy and CRT in the following 5 groups: entire cohort, adenocarcinoma patients only, adenocarcinoma patients who underwent esophagectomy, adenocarcinoma who received 3 cycles of induction therapy, and adenocarcinoma who received 3 cycles of induction therapy and underwent an esophagectomy.

Reply: Thank you for this comment. We performed multiple sensitivity analyses to help provide hypothesis-generating data related to the interplay of each of these interventions upon survival, specifically in patients with adenocarcinoma who would be hypothesized to derive greatest benefit from both induction chemotherapy and surgery. Patient numbers and duration of follow-up limit the strength of comparisons in these analyses, and for some, the median survival has not yet been reached. We have reported these data in text (below) and would favor limiting to text and eFigure2.

"When restricting the analysis to patients with adenocarcinoma histology (AC-EC, n=77), the median PFS and OS for the IC-CRT vs. CRT cohorts were 18.7 months (95% CI 9.0- not reached) vs. 22.0 months (95% CI 12.0- not reached) (p=0.93) and 38.1 months (95% CI 23.4- not reached) vs. 59.4 months (95% CI 27.9- not reached) (p=0.75), respectively (Figure 2A-B). Similarly, when comparing patients with AC-EC who received IC-CRT with ≥ 3 cycles of induction FOLFOX8 vs CRT, there was no difference in median PFS (15.1 vs 18.7 months, p= 0.93) or median OS (32.3 vs 38.1 months, p=0.86) (Figure 2C-D). In the subset of AC-EC patients who had esophagectomy (n=50), the median survival was not met in IC-CRT cohort however there was no statistical difference in outcome by log-rank in PFS or OS with 3-year PFS of 60% [95% CI 39-75%] vs 43% [95% CI 21-63%] and 3-year OS of 70% [95% CI 48-84%) vs 58% [95% CI 33-77%] (Figure 2E-F). Lastly, when restricting analysis further to AC-EC receiving at least 3 cycles of induction FOLFOX, CRT, and esophagectomy versus CRT and esophagectomy (n=39) there was still no significant difference in outcome between IC-CRT and CRT cohorts (eFigure2)."

Comment 2: In the Discussion section, please comment on why over 40% of patients did not undergo surgery. Were the patients evaluated by surgeons prior to initiation of neo-adjuvant therapy? If functional status was poor, could the patients have had surgery later, perhaps 6-12 moths after completion of therapy?

Reply: In eTable1 we specify the reasoning esophagectomy was not pursued for the 40 patients that went on to observation following the chemoradiation component of treatment, but we agree that further mention and explanation of this is warranted in the discussion. Please see below for the corresponding edits.

Changes in text- Discussion: "*While the overall cohort of patients included in this analysis had technically resectable disease at diagnosis, 42% did not ultimately proceed to esophagectomy for a variety of reasons listed in the appendix. The majority of patients either declined surgery by the time CRT +/- induction was finished, had interval progression of disease which precluded an operation, or their performance status was such that they were no*

longer eligible. This demonstrates the toll that multimodal treatment may haves on this population of patients."

Comment 3: In Table 1, please list the types of esophagectomy performed and, if feasible, the median time from completion of therapy to surgery.

Reply: We have added the median time from completion of chemoradiation to the time of esophagectomy. We did not collect surgical technique in our database, however, we have listed primary tumor location in the esophagus or gastroesophageal junction. As such, a reader may have some understanding of potential surgical techniques which may have been utilized within the cohort based upon primary location.

Changes in text- Results: "Amongst the 55 patients who underwent esophagectomy, *median interval from completion of RT to surgery was 63 days (IQR49.5-82)*."

Comment 4. Please state whether there were any post-op or peri-op deaths as this could play a role in overall survival results.

Reply: We have added an explicit statement that there was one post or peri-operative death, all complications are listed in eTable 3.

Changes in text- Results: "*There was one post-operative/peri-operative death in our cohort as a result of anastomotic leak;* surgical complications did not different between the cohorts and are detailed in eTable3."

Comment 5: Please add p values to Table 4.
Reply: Given this was an exploratory analysis without prospectively collected objective measurements of dysphagia, we opted to report the data in Table 4 descriptively without formal statistical comparison with p-values.

Comment 6: There is a discrepancy between the abstract and Tables 1 and 2 in terms of number of patients in each group.
Reply: Thank you, this has been corrected in the abstract (IC-CRT =59 and CRT=36).