

Peer Review File

Article information: <https://dx.doi.org/10.21037/jtd-22-747>

Comment 1: Why was a composite outcome used? I would recommend separating end-organ failure from death. The event rate with end organ failure is high enough that the study should maintain power to detect a difference. This would also make the analysis cleaner for readers to interpret. Are patients with higher RAPID scores more likely to die with delayed surgery?

Reply 1: Mortality and organ failure will be reported separately now. There was no significant association with mortality and surgical timing without taking RAPID scoring into context, however once mortality was evaluated with surgical timing + RAPID scores it was found that high RAPID score and early surgery was associated with mortality, whereas high RAPID score and late surgery was not.

Changes in the text: The abstract (lines 50-62), methods (lines 102-111), results (163-171, 174-181, 184- 192), discussion (203-211), conclusion (247-254), Table 2, Table 4 were changed and Table 5 was added to reflect these changes.

Comment 2: How did the authors adjust for patient acuity? Were patients in the late surgery group taken to the OR because they decompensated, thus resulting in worse outcomes? Some explanation for this adjustment or inability to account for this should be added as a limitation

Reply 2: There was no adjustment for patient acuity, we used comorbidities as a surrogate. There was not enough information in the medical records to adequately and consistently adjust for patient acuity. This was a retrospective study and the specific reasons why patients were taken to the OR at their specified time was rarely clear in the records.

Changes in the text: none

Comment 3: There are multiple analyses performed which increase the probability of obtaining a false positive result. How is multiplicity of data addressed? Additionally, I would recommend performing a 2x2 analysis High/ low score against early/late surgery. The late surgery/ high score would be the reference group to obtain your odds ratios. This would make it easier for readers to understand what the impact of each variable has on the odds of the primary outcome developing

Reply 3: To address multiplicity, we applied the Bonferroni correction to all subset analyses. The 2x2 table might help us see if there is a significant interaction in the odds of mortality between high/low and early/late groups. Using the Breslow-Day test for homogeneity of ORs (seeing if ORs significantly change across strata) and stratifying by surgical timing (early vs late), we do not find a significant difference in the ORs of mortality for high vs low RAPID scores ($p=0.1733$). Similarly, the BD test did not find a significant difference when looking at the odds of organ failure ($p=0.4833$) or 30-day readmission ($p=0.0597$).

Additionally, we looked at this using our multivariate models to see if the interaction between score category and surgical timing was significant. If it is, then it would indicate the need for separate ORs for the early and late groups. For 90-day mortality ($p=0.9524$), organ failure ($p=0.8378$), and 30-day

readmission ($p=0.1008$), we failed to detect a significant interaction between score category and surgical timing. In other words, they are dependently related.

Due to the above, we do not feel that expressing the data in a 2x2 table format would provide any further clarification on the data already presented.

Changes in the text: none

Comment 4: How many patients had stage II. vs. III. empyema (ATS classification) and how many patients had VATS vs. thoracotomy or open windows thoracostomy (OWT) How many patients died in the subgroup thoracotomy (probably Stage III), please comment.

Reply 4: 9 with stage 2, 115 with stage 3, 36 with mixed stage. 101 with VATS, 111 had thoracotomy, 52 had VATS which converted to thoracotomy and 0 had OWT. 9/111 patients who had thoracotomy died.

Changes in the text: Table 1 was updated with above data. Methods (108-109), Results (154-156, 160-161) were also updated.

Comment 5: Did you investigate the role of the art of surgery (VATS vs. thoracotomy vs. OWT). Perhaps the type of intervention influenced the results. It is known that patients in poor general condition /debilitated patients, patients with high RAPID score) do not benefit from decortication/surgery. Rather, an OWT is recommended in such cases. please comment.

Reply 5: We updated our data with the type of surgery performed as noted in reply 4. None of the patients in our cohort underwent OWT.

Changes in the text: none

Comment 6: 188: RAPID score was directly associated with mortality or organ ???(Table 4).

Reply 6: RAPID score was directly associated with organ failure.

Changes in the text: as noted in reply 1

Comment 7: 205-09: For example, the American Association for Thoracic Surgery recommends against routine use of intrapleural fibrinolytics to treat empyema. This may explain why intrapleural fibrinolytic rates in our study were relatively low (25%). The decision to use intrapleural fibrinolytics was pragmatic and didn't always involve a pulmonary consultation. So, 25 % of the patients received fibrinolytics and surgical intervention, this experience raises the question of whether the use of fibrinolytics even needed. What are the outcomes in the fibrinolytics subgroup vs. primary surgery, please comment.

Reply 7: This study was not designed to evaluate the efficacy of intrapleural fibrinolytics. This has been evaluated in many other studies as noted in the consensus statement found in *Lancet Respir Med.* 2021;9(9):1050-1064). These authors noted that intrapleural fibrinolytics can significantly reduce the need for surgical management. As our cohort was only comprised of patients who required surgical management, significant change in outcomes would not be expected. Accordingly, one of the Multivariate

Regression models assessing the composite outcome of 90-day Mortality or Organ Failure did not find a significant association with intrapleural fibrinolytics.

Changes in the text: Lines 231-235 were updated to “For example, the American Association for Thoracic Surgery recommends against routine use of intrapleural fibrinolytics to treat empyema.⁶ This may explain why intrapleural fibrinolytic rates in our study were relatively low (25%) since the decision to use intrapleural fibrinolytics was pragmatic and often at the discretion of the primary surgical team.”

Comment 8: 210-212: Our study has several limitations inherent to retrospective study design. First, data regarding whether the empyema was in the exudative, fibropurulent, or organizing stage were unavailable because radiographic details were not included in the data collection. You do not need radiographic details, just read the operation report or ask the surgeon. Please comment.

Reply 8: We appreciate the clarification and reviewed the available operation reports again to obtain this data.

Changes in the text: Table 1 was updated with above data. Methods (108-109), Results (154-156, 160-161), Discussion (237-242) were also updated.

Comment 9: 175-179: Touray et al. hypothesized that earlier intervention in patients who are good surgical candidates with a high RAPID score could decrease the risk of mortality. In the study of Sziklavari et al. 2011 for comparison, debilitated patients with delayed OWT and VAC therapy left the hospital after 31 ± 14 days. In patients with initial fenestration, however, the hospital stay was only 11.5 ± 3.5 days. This finding was consistent with Massera and colleagues 2006, who concluded that immediate creation of OWT is a significant predictor of successful thoracostomy closure. So, I think debilitated patients could benefit from early surgery, even if it is an OWT. Please comment.

Reply 9: How OWT could have affected outcomes would be interesting to evaluate, however none of our patients underwent this procedure. It is unknown if this is due to individual patient factors, individual or regional surgeon experience with this procedure, or for other reasons.

Changes in the text: none

Comment 10: The finding is interesting and important to clinicians. Please clarify how to define the timing of early and late surgery (≤ 3 days from diagnosis). [https://journal.chestnet.org/article/S0012-3692\(19\)32316-5/fulltext](https://journal.chestnet.org/article/S0012-3692(19)32316-5/fulltext).

Reply 10: The current manuscript reports: “As noted by Touray et al, there is no consensus on the definition of early versus late surgical timing with reported values between 48 hours and 2 weeks in previous literature.^{3,11-13} Also, we chose 3 days to allow for the typical dose regimen of tPa/DNase as reported in the MIST2 trial.¹⁴” Given the lack of consensus using a definition other than 3 days and the use of 3 days in other notable studies we felt it reasonable to use 3 days as our definition.

Changes in the text: none

Comment 11: In your study, there was no consistent expression in the low versus high RAPID scores in tables and supplemental tables.

Reply 11: The reported expression of low vs high RAPID scores reflected the same numbers [High (≥ 4) vs Low (< 4) is equal to Low (≤ 3) vs High (> 3)] however we agree with the above comment that this is unnecessary and confusing.

Changes in the text: All tables will now report RAPID scores as Low (≤ 3) and High (> 3)

Comment 12: A high degree of awareness is essential for perioperative management, and early surgical treatment is the benefit to patients with empyema thoracis. However, minimal invasive surgery is the trend around the world. How many patients with high RAPID scores undergoing VATS? Do you apply ERAS protocol in your patient's cohort?

Reply 12: 101/160 pts underwent minimally invasive surgery with VATS and 52/101 pts with VATS had to be converted intraoperatively to thoracotomy. ERAS protocol is followed routinely at the facilities included in the study. 27/101 pts with high RAPID scores underwent VATS.

Changes in the text: Table 1 was updated with above data. Methods (108-109), Results (154-156, 160-161) were also updated.

Comment 13: With respect to the fast recovery, there is no information about the empyema stage and severe infectious status before surgical intervention, including the levels of CRP or PCT. More data may support your explanation that the late surgery and a high RAPID score were related to a higher rate of end-organ failure or mortality.

Reply 13: Levels of CRP and PCT were not consistently obtained for the patients in our cohort. Empyema stage has now been analyzed and is now reported in the text.

Changes in the text: See comment/reply 12

Comment 14: The RAPID score in this manuscript is based on the published paper of Rahman et al (Chest 2014). The primary outcome of Rahman et al was 3 months mortality. Highrisk RAPID score in Rahman et al paper was associated with increased mortality, also in the validation cohort (MIST2). The mortality rate seems as the most significant and serious outcome for a patient with empyema. Did the authors in this manuscript check the outcome of 90-day mortality or 12-month mortality alone and not as a composite score?

Reply 14: 90-day mortality was evaluated separately and in composite with organ failure. They will now be reported separately to avoid confusion. 12-month mortality was not evaluated.

Changes in the text: See comment/reply 1

Comment 15: The composite score in the submitted manuscript includes 90-day mortality and end-organ failure. The end organ failure score is composed of 1. Need for mechanical ventilation 2. Use of

vasopressors and 3. Development of acute renal failure. What was the reason to include those parameters? Why other organs such as the liver, and cardiac were not included?

Reply 15: We included the above parameters because in our experience these measures of organ failure were more closely associated with severe disease in general and need for ICU admission. Other organ system measures were not included due to 1) lack of consistent data collection of the associated lab values and other clinical documentation to note organ failure in these systems and 2) the belief that, in general, these organ systems are not as affected in severe infections/sepsis syndromes as they are in respiratory, vascular and renal systems.

Changes in the text: none

Comment 16: *The type of surgical approach (Thoracotomy versus Thoracoscopy), (Decortication versus Washout and drainage) should be clarified. How many patients are in each group?

Reply 16: 101 with VATS, 111 had thoracotomy, 52 had VATS which converted to thoracotomy. 153/160 pts had decortication, 7/160 pts had simple washout/drainage.

Changes in the text: see comment/reply 12

Comment 17: *The authors chose 3 days to define early versus late surgery (3 days were chosen to allow for the typical dose regimen of intrapleural fibrinolysis in the MIST2 trial). The rate of intrapleural fibrinolysis in this cohort is just 25%. The MIST2 trial in 2011 demonstrated that combined use of tPA and DNase decreased surgical referral rates and length of hospital stay. A surgery-first approach will probably lead to procedures in many more patients than needed. What were the reasons for the low rate of fibrinolysis?

Reply 17: The original manuscript notes: "Chest tube drainage with intrapleural fibrinolytic therapy has been recommended to avoid surgical intervention, but there is some debate regarding its use to treat empyema.⁸ For example, the American Association for Thoracic Surgery recommends against routine use of intrapleural fibrinolytics to treat empyema.⁶ This may explain why intrapleural fibrinolytic rates in our study were relatively low (25%). The decision to use intrapleural fibrinolytics was pragmatic and didn't always involve a pulmonary consultation." One hypothesis is that the patients whose primary team was surgical (the majority of our patients) were more likely to proceed without fibrinolysis, as suggested by surgical guidelines, while other patients whose primary team was medical were more likely to pursue non-surgical options first.

Changes in the text: none

Comment 18: The authors mention that surgical practices varied between different hospitals and the timing of surgery was not decided based on any algorithm but clinical judgment. I believe that this lack of information is a significant limitation of this manuscript and its conclusions.

Reply 18: We agree as noted in our limitations. This was a pragmatic, retrospective study. The only way to avoid this would be to perform a multi-center prospective study with the creation of a standard,

consistently used algorithm for the medical/surgical treatment of empyema. To our knowledge, such an algorithm has not been reported in the literature or recommended by medical/surgical guidelines.

Changes in the text: Lines 241-242

Comment 19: *The type of surgical approach (Thoracotomy versus Thoracoscopy), (Decortication versus Washout) should be clarified. The type of surgery(Thoracotomy versus Thoracoscopy) might affect the length of stay, organ failure(respiratory failure), and readmission rate by itself.

Reply 19: The type of surgery was not significantly associated with mortality, organ failure, LOS or readmission rates

Changes in the text: lines 155-156

Comment 20: *106 patients (58%) developed new organ failure(ie. Ventilator use...). Did you include all kinds of organ failure?

Reply 20: The original manuscript notes that organ failure was defined as: “need for noninvasive or invasive mechanical ventilation, vasopressors, and/or development of acute renal failure defined per KDIGO guidelines”. Our reasons for not including other measures of organ failure are described in reply 15 above.

Changes in the text: none

Comment 21: *After accounting for demographics and comorbidities the regression analysis suggested that late surgery in combination with RAPID score>3, was associated with increased 90-day mortality or new organ failure. The authors mention that most of the association was due to the development of organ failure. Did you find a significant correlation between 90 day or 12 months mortality to time of surgery or RAPID score.? Did you check the effect of the type of surgery on new organ failure?

Reply 21: There was not a significant association between 90-day mortality and timing of surgery. There was a significant association with 90-day mortality and RAPID score. The type of surgery was not associated with organ failure.

Changes in the text: See reply 1 and 19

Comment 22: * The authors suggest that RAPID score can be useful to predict those who may benefit from early surgery in order to avoid increased 90-day mortality or new organ failure. The varied surgical practices between different hospitals in this cohort, the lack of information regarding the decision of timing of surgery, and the low use of intrapleural fibrinolytic challenge the proof of the claim.

Reply 22: We agree that the above critiques are limitations as we noted in the original manuscript. As noted in reply 18, we feel like the only way to clearly account for these limitations would be to perform a multi-center prospective study with the creation of a standard, consistently used algorithm for the medical/surgical treatment of empyema. To our knowledge, this type of study has not yet been performed.

As we suggested in the original manuscript, “Future prospective studies should examine the validity of our results to determine which patients would benefit from expedited surgery.”

Changes in the text: Lines 241-242

Comment 23: Line 188- The word failure is missing after organ.

Reply 23: This error was corrected in our revision.

Changes in the text: corrected to organ failure