Peer Review File

Article information: https://dx.doi.org/10.21037/atm-22-1667

First External Peer Review

<mark>Reviewer A</mark>

Overall well written article, which shows increased mortality in patients with COVID and active malignancy. The following suggestions would make a stronger paper regarding relationship.

1. Consider adding additional table which summarizes active chemotherapeutics used. This will demonstrate immunocompromised state, not only related to active malignancy, but to therapy as well.

   Unable to add due to data not available for patients who were on active vs. not on chemotherapy

2. Consider adding the following variables to table 1: Morbidity (survivor and non-survivors with p-values).
Data is provided in the revised manuscript Table 2.

3. Figure 1 should be a flow chart showing eligibility, excluded population and resultant two populations analyzed (i.e. cancer and noncancer.
All adult patients hospitalized with COVID-19 between March 17th, 2020 and January 18, 2021 were included.

4. Remaining figures relabeled accordingly.
   We have completed it in the manuscript
5. Table 1 requires a legend to accompany baseline characteristics.
   Legend is done in the top of Table 1

Although a significant number of studies have already been published on covid-19 and cancer, more data are needed, and the present study provides a solid ground for comparisons between cancer and non-cancer patients, something which is missing in most published studies. As covid-19 is a disease with large prognostic and clinical heterogeneity, covid-19 patients populations are also very diverse, depending on selection mechanisms. Hence, rate ratio comparisons within the patient population (presumably selected in a uniform manner) give better information than actual mortality rates or other markers of disease severity.

I have some remarks:

1. The phrasing is kind of vague as to the study design. To me, this seems to be a cohort study, at least primarily (in the comparisons between cancer and non-cancer patients). It should then be described as a cohort study throughout the manuscript, and the cohort should be called a cohort in the material and methods section.

   Thanks for the suggestions we have updated the material and methods and other sections with the language.

2. According to the methods section, there were around 15,000 admissions due to COVID-19, but only 4870 patients with COVID-19 and admitted to the health system were included in the study. Why the discrepancy? Did the study require manual chart review?

   When the pandemic started we did not imagine that we were going to treat so many patients with COVID-19 then we build a REDCAP database and manually added the data one by one, after 4870 patients we realized that was too much work and we needed the human resources to help treating COVID patients. We hope to find a way to extract all the data from our EPIC EMR for another publication.

3. The patients were collected between March 2020 and January 2021, which in calendar time is less than eleven months but in covid-19 time is half an eternity. The clinical knowledge increased dramatically during these months, and the pandemic shifted character. Was there a difference between spring 2020 and winter 2020/2021 in mortality? Were the cancer patients evenly distributed over the first waves of covid?

   The mortality rates were similar between Spring 2020 and Winter 2020/21. The cancer patients were evenly distributed. For the same reasons the reviewer is mentioning I think was good to stop the data for this comparison in January 2021 because since January 2021 we started to vaccinate patients; probably another publication can compare the

outcomes before and after vaccination of the patients.

4. In the first half of 2020, access to covid testing was also limited in many parts of the world. Was this true for southern Florida as well? Assuming that cancer patients might have been prioritized for testing in the early pandemic, would that have influenced your results?

COVID testing was not limited to in southern Florida. Our Healthcare System is governmental and we had the National Guard testing since the beginning in our hospital grounds with sometimes thousands of patients per week.

5. In the results section (lines 145-147), it is stated that "Hematologic malignancies accounted for 24.6% of COVID-19-related cancer deaths, in comparison to solid malignancies that accounted for 23.5%, but difference was found to be not statistically significant (p=0.845)." It might be my language skills that are missing, but to me it sounds like 24.6% of those with cancer and covid-19 who died had a hematologic malignancy and 23.5% had a solid tumor. I guess you mean that 24.6% died among those with hematologic malignancies and covid-19, an interpretation which would also fit the statement that there was no difference between the groups. The latter finding is, by the way, contradicting the findings of several other groups who have found patients with hematologic malignancies and lung cancer to be extra vulnerable. It might be worth a bit of discussion. Which hematologic malignancies were included?

I agree with the comments, we included all hematological malignancies meaning leukemias, lymphomas, myelomas, etc. Lung cancer also has a high mortality but it's included in the solid tumor group probably increasing the average mortality in the group. Also our sample is small we need probably a larger sample to have better statistics

6. What was the definition of "had a history of cancer"? Were those ongoing malignancies? Malignancies under treatment? Did the definition also include malignancies in a distant past?

History of cancer included all patients diagnosed with cancer regardless if on active treatment or monitoring. The definition did not include those that are in remission.

7. In the end of the results (lines 157-161), it is stated that : "Among patients with cancer, multivariable logistic regression analysis demonstrated that older age, elevated levels of creatinine and CRP were independently associated with increased risk of death in COVID-19 patients with cancer, with multivariable-adjusted ORs of 1.05 (95% CI, 1.02-1.09; p=0.006), 1.31 (95% CI, 1.04-1.66; p=0.024), and 1.05 (95% CI, 1.01- 1.10;

p=0.036) respectively (Figure 2)." Was this also true for the non-cancer patients? Did any of these measures strike you as especially relevant for the cancer patients?

This was not true for non-cancer patients. That description was only relevant for cancer patients.

For me, two methodological points need to be clarified: the concept of active malignancy being under surveillance and few missing data (metastatic status, tumor type)

Thanks for the suggestions cancer patients were classified based on their history of malignancy and we do not know their status of malignancy (active/inactive)

The manuscript from Raez and colleagues compares the survival outcomes between hospitalized COVID-19 patients with and without cancer. A retrospective, Memorial Health Care System review design was employed. In total, 4870 patients were identified who met the initial inclusion and exclusion criteria, 265 of those had a diagnosis of cancer. The main finding from this study suggests that patients diagnosed with Cancer have poorer prognosis. The study is well written, and the results obtained in the present study is consistent with previous studies and metanalyses.

Specific comments:

1. Material and Methods - the description of the study population should be further clarified, specifically a flow chart should be presented. why were so many patients excluded from the analysis? Were the data collected consecutively?

   We have done a flowchart that is included as requested.

   When the pandemic started we did not imagine that we were going to treat so many patients with COVID-19 then we build a REDCAP database and manually added the data one by one, after 4870 patients we realized that was too much work and we needed the human resources to help treating COVID patients.

Data between March 1, 2020 and January 2021 were collected consecutively and 4870 patients with COVID-19 were included.

2. Statistical analysis – Why was the selection of variables to perform the multivariate analysis so restrictive ($p<0.05$)? Why were the variables creatinine and CRP not included in the multivariate analysis? Please include a table presenting the results of the univariate logistic regression.

In the cohort of 265 COVID-19 patients with cancer, a standard significance level for hypothesis testing is used to decide which variable to include in the model based on the concept of event per variable. The purpose of such selection is to determine a set of variables that will provide the best fit for the model and to avoid noise variables so that accurate predictions can be made. Covariates for the models were chosen based on prior review of the literature and clinical judgment, focusing on variables that might be expected to confound the cancer-risk relationship. The following variables including age, sex, race/ethnicity, hypertension, diabetes, chronic obstructive pulmonary disease, chronic kidney disease, coronary artery disease, obesity, smoking, and laboratory tests (WBC, lymphocytes, D-dimer, creatinine, C-reactive protein, AST, and ALT) were included in the multivariate logistic regression model, even though those variables with a p value > 0.05 in the univariate logistic regression. Results of the univariate and multivariate logistic regression are provided in the revised manuscript (Table 3).

However, this article didn't bring new relevant information about COVID-19 patients with or without cancer.

There weren't information about stage, kind of oncological treatment or status performance from cancer patients. Furthermore, there weren't also information about admittion to the ICU, or need of invasive mechanical ventilation or other organ dysfunctions between both groups.

After adjusted analysis, just age, gender and race were considered risk factors of death. It wasn't clear to me how comorbidities were evaluated by authors during multivariable logistic analysis. Comorbidities have been often considered as an independent prognostic factors of death in COVID-19 patients.

About serum levels of creatinine, what was considered high levels of creatinine? Was it associated with more need of haemodialysis?

Thanks for the comments. Creatinine levels >1.3 mg/dL. There was no evaluation for need for HD.

About serum levels of C-reactive protein, this biomarker is inflamatory unspecific related to cases' severity or mortality.

We agree but when COVID started we did not know the disease and what biomarkers were relevant to treat these patients.

Second External Peer Review

Reviewer A

The authors adequately addressed my comments.

Reviewer B

The authors have taken the various comments into account.

1. Please add in table 1 the number of missing data (especially for biological values : IL6, CRP....). It is important to analyze results.

   A: The number of missing data were added in the revised Table 1.

2. The main limitation being that we do not know the status of the cancer (active/inactive) and the stage (localized/metastatic).

   A: Data were not available regarding the status of the cancer (active/inactive)