**Reviewer A**

The authors retrospectively explored the prevalence of substance use and its associations with clinical and social parameters among patients with non-small cell lung cancer. Considering the scantiness of data on this topic, the results from this study can provide some implications for clinicians. However, several revisions are warranted before considering this manuscript for publication. Please see my comments below.

Major comments:
Comment 1. Methods: The authors should provide more clear definition on "hospital admissions" in the methods section. Does it mean unintended / emergency admission (exclude scheduled ones?) etc.
**Reply:** We thank the reviewer for this comment. We included all hospital admissions regardless of type in our data analysis.
**Changes in the text:** We had edited the text on Page 6, Lines 127 - 129 to clarify that we determined the frequency of all hospital admissions for patients in the cohort.

Comment 2. Methods: I request the authors to provide the details of "substance use" (e.g., specific drug names and their number and prevalence). I think this point is important to improve the comprehensibility of this study in comparisons with other data on similar topic.
**Reply:** We thank the reviewer for this comment. We captured the route of administration and specific names of substances used by patients. Cocaine and marijuana use were the most prevalent with 45.8% (22/48) and 52% (25/48) of patients reporting cocaine and marijuana use, respectively. Intravenous heroin use was reported in 10.4% (5/48) and methamphetamine use in 0% (0/48). Polysubstance use was noted in 27.1% (13/48) of patients. Nine patients (36%) reporting marijuana use had concomitant use of other substances.
**Changes in the text:** We have edited the text on Pages 8 and 9, Lines 180 – 185 to reflect the prevalence data.

Comment 3. Substance use and guideline concordant care: I request the authors to provide the details on guideline non-concordant patients. This point can have an important clinical value to explore the impact of substance use on patients' compliance.
**Reply:** We thank the reviewer for this comment. We provided the reasons for patients not receiving guideline-concordant care on Page 10, lines 222-224.
**Changes in the text:** No additional changes were made to the text at this time.

Comment 4. Discussion: The differences between this study and preceding ones are still unclear. Line 274-277 might fail to show this point because of ambiguous descriptions.
**Reply:** We appreciate the reviewer's comments.
**Changes in the text:** We have modified the text as advised by providing clear descriptions in the

Methods section and removing ambiguous wording as described in more detail below.

Minor comments:
Comment 1. Introduction: In the introduction section, the authors introduced the preceding studies on substance use problem. There, they used the different words as followings: "substance", "non-tobacco substances", "harmful substances". What does these words exactly mean? I recommend the authors more clearly define these words to show the readers the lack of data and novelty or strength of current studies.
**Reply:** Thank you for this comment. We have used the term substance use more consistently in the text to avoid ambiguity. We define substance use in the context of our study in the Methods section (page 7, lines 140 – 143). Substance use was obtained through electronic medical record review and defined as any consumption of illicit drugs, inhalants or solvents, marijuana, and/or use of prescription drugs outside of their intended purpose. Alcohol and tobacco use were defined separately.
**Changes in the text:** We have also edited the text as advised on page 5, lines 88 to 93.

Comment 2. Methods: Please provide the ID number for IRB approval.
**Reply:** Thank you.
**Changes in the text:** We have added this information on Page 5, Line 108 of the manuscript file.

Comment 3. Patient adherence to treatment: Why the authors omitted the percentage of missed surgery appointments? Please correct, if it is a mistake.
**Reply:** We appreciate this observation.
**Changes in the text:** Information on the percentage of missed thoracic surgery appointments was added to the manuscript file on Page 10, Lines 212 - 214.

Comment 4. Substance use and overall survival: I could not find "figure 1C" in provided PDF.
**Reply:** Thank you.
**Changes in the text:** We have added Figure 1C to the figure files.

Comment 5. Discussion, line 326: I agree that interventions for substance use is highly important. Were the patients in the authors cohort provided any interventions or support for substance use? Please provide some information on this point.
**Reply:** We thank the reviewers for this comment. We collected data on whether patients with current substance use had supportive care or were undergoing treatment. However, the data was largely unreported. This could not be resolved via medication review or further detailed review of the medical record. Although we did not offer interventions as part of this retrospective cohort study, we do recommend this as a future direction for research in this area.
**Changes in the text:** We have highlighted the additional information requested in the Results section (Page 9; Lines 186 – 187) and Discussion section (Page 13, Lines 296 – 298) of the text.

Comment 6. Table 1: Pleas show data on treatment modalities, not just as stage" as patients' background in the table.
**Reply:** We thank the reviewers for this helpful comment.
**Changes in the text:** Information has been added to table 1 on planned treatment modalities in patients with and without substance use.

**Reviewer B**

Thank you for the opportunity to review the manuscript.
Although it is an interesting theme, unfortunately the patients are ambiguous and the analysis method is insufficient.

Comment 1. The "substance use" is ambiguous and difficult to understand from the title and abstract. Please correct it to a word that gives a more concrete image, such as adding "addictive".
**Reply:** Thank you for this comment. We have used the term substance use more consistently in the text and defined the terminology in the context of our study in the Methods section (page 7, lines 140 – 143) which hopefully improves ambiguity. Substance use occurs along a continuum, specifically, use, abuse, and dependence. (reference 14 in the main text) Substance use (without addiction, abuse or dependence) can cause or exacerbate physical and psychological problems: symptoms related to substance use or worsened comorbidities, and the consequences of decreased treatment adherence, being under the influence, or undermining one's social support system. (references 15-17 in the main text) Thus, our study focused on substance use, and we did not capture data separately on addiction, abuse or dependence.
**Changes in the text:** We attempt to clarify this distinction on Page 7 in Lines 136 - 143

Comment 2. Please describe what kind of substance, how much and how long they used
**Reply:** Thank you for this comment. We were able to capture the route of administration and specific names of substances used by patients. Cocaine and marijuana use were the most prevalent with 45.8% (22/48) and 52% (25/48) of patients reporting cocaine and marijuana use, respectively. Intravenous heroin use was reported in 10.4% (5/48) and methamphetamine use in 0% (0/48). Polysubstance use was noted in 27.1% (13/48) of patients. Nine patients (36%) reporting marijuana use had concomitant use of other substances. Unfortunately, we were not able to capture how much and how long each kind of substance was used due to large amounts of missing data on this information in the electronic medical record. It is possible that providers under investigated or patients underreported such details.
**Changes in the text:** We have edited the text on Pages 8 and 9, Lines 180 – 185 to reflect the prevalence data.

Comment 3. The analysis should be considered separately whether it is current or former.
**Reply:** Thank you for this comment. We considered this in our study design and analysis but the numbers of patients in each group were too small to make reasonable descriptions, comparisons, and conclusions.
**Changes in the text:** No changes were made to the text.

Comment 4. It is natural that smoking is a significant factor in this study. It is necessary to devise an analysis such as equalizing the number of cigarettes smoked.
**Reply:** We appreciate the reviewer's comment. We accounted for the potential confounding effect of smoking via multivariable analysis as depicted in Table 3. In our descriptive analysis we found that those with substance use were more likely to be current smokers. Because of this effect and our small sample size, we deemed further analysis of this particular dataset unlikely to yield additional clues in that regard.

**Changes in the text:** No changes were made to the text at this time.

Comment 5. Comorbidity and nutritional status are not described.
**Reply:** We thank the reviewers for this important point. Understanding that individual co-morbidities could affect the outcomes being measured (health care utilization), we attempt to describe the impact of co-morbidities through the Charlson Co-morbidity index, a weighted composite score of common co-morbidities. We did not collect data on the nutritional status of the patients.
**Changes in the text:** No changes were made to the text.

Comment 6. Because there is a social background such as unemployment, they use substance. Please add consideration with the literature such as nutrition and other diseases related to substance use.
**Reply:** We thank the reviewers for this important point. Understanding that individual co-morbidities could affect the outcomes being measured (health care utilization), we attempt to describe the impact of co-morbidities through the Charlson Co-morbidity index, a weighted composite score of common co-morbidities. We did not collect data on the nutritional status of the patients.
**Changes in the text:** No changes were made to the text.

Comment 7. In the Abstract, GCC abbreviation does not be required.
**Reply:** Thank you. GCC abbreviation has been removed from the abstract.
**Changes in the text:** We have modified the text as advised on Page 4, Line 91.


**Reviewer C**

I agree that the research on the influence of substance use for patients with non-small cell lung cancer could be valuable in the U.S. where substance use has become endemic.
However, the sample size of the present study is too small as a retrospective epidemiological study. This is evident when compared to the study population in similar retrospective studies in other malignancies listed in the references.

Therefore, the paper is unfortunately not in the scope of the journal of thoracic disease.
From my point of view, further analyses with higher numbers of patients must be necessary.

**Reply:** The authors thank you for your comments. We agree that researching the influence of substance use in patients with non-small cell lung cancer is important and acknowledge the small sample size as a limitation of our work. We do suggest that this is still important to report this data from our unique population at a large safety net hospital providing care for patients with otherwise limited access to care. We also suggest that it is important to evaluate the influence of substance use on healthcare utilization, treatment, and outcomes in the context where there is substantial logistical and social support for otherwise disadvantaged populations.
**Changes in the text:** No changes were made in the text.

**Reviewer D**

I sincerely appreciate your excellent paper.
However, I found figure 1C being omitted from this paper.

**Reply:** Thank you for your comments.
**Changes in the text:** Figure 1c has been included in the paper.


**Reviewer E**

This paper investigates the impact of substance use on several factors regarding non-small cell lung cancer. The authors found the significant relationship between substance use and some patient characteristics, social determinants of health and healthy care utilization. For overall analysis, the author concluded the significance of substance use in univariate analysis, but no significance in multivariate analysis. Just as authors wrote, substance use is significantly among male patients with current smoking and alcohol. Thus, Sex, Substance use and Smoking status are highly correlated, which is multicollinearity problem in statistics. I suggest authors to further solve this problem, which may help to have a better result, such as using principal components analysis or LASSO / Ridge regression. I searched related papers and it seems this study is newly to evaluate the effect of substance on NSCLC outcomes. But I still found one, which can be cited, reviewed and compared, which is Concannon, K., Thayer, J. H., Hicks, R., Wu, V., Jenkins, I., Baik, C. S., & Linden, H. M. (2019). Outcomes among patients with a history of substance abuse in non-small cell lung cancer: A county hospital experience.

In summary, this paper indeed provides useful investigation for substance use on NSCLC, but for some statistical analysis. Some further work can be done help investigate thoroughly.

**Reply:** Thank you for your comments and the example reference. We have reviewed this reference and searched for other updated references. The literature on this topic is scarce but the most updated, related studies have been cited. The study by Concannon et al focuses on homelessness but has useful areas of comparison.
**Changes in the text:** We have edited the text as advised to include the reference on page 20, lines 489 - 491: Concannon KF, Thayer JH, Wu QV, et al. Outcomes Among Homeless Patients With Non-Small Cell Lung Cancer: A County Hospital Experience. *JCO Oncology Practice.* 2020;16(9):e1004-e1014. doi:10.1200/JOP.19.00694