**Peer Review File**
**Article information: https://dx.doi.org/10.21037/jtd-24-205**

**Reviewer A**

**Comment**
The data show an increasing incidence of NSCLC in women. But this is not in keeping with literature reports which point to a decrease in both men and women howbeit a greater decline in men.

**Response**
Thank you for this comment. We have clarified the data. We did not see an increasing incidence in women. Instead, we experienced an equalization of incidence between men and women. This trend occurred because the incidence of lung cancer dropped more rapidly in men than in women. In lines 239-240, we have now said "This equalization likely occurred because the incidence of lung cancer decreased more rapidly in men than in women."

**Comment**
Arising from no 1, an explanation of the large difference in NSCLC in women is necessary to underscore the absence of selection bias of their chosen michigan state registry as authors suggest

**Response**
Thank you for this comment. We have added discussion regarding this in lines 239-241 to clarify that there was not an increased incidence in women over the years, but an equalization of incidence.

**Comment**
While travel time may explain the disparity between rural and urban center, other explanatory factors that need to be addressed include educational level, socioeconomic status and political affiliation particularly due to the fact that even longer travel time was not significant. Do these factors affect whether a patient would access screening or not?

**Response**
Thank you for this comment. We added additional discussion to these points in lines 255-260. We have now said "Additionally, we did not examine patient level socioeconomic factors that likely contribute to increased mortality in more rural areas. Patients in more rural areas may have differences in education level compared to similar patients in urban regions which can impact compliance to lung cancer screening guidelines. Political affiliations, primary payer and insurance status were not captured in this dataset, but may be important in understanding the disparity in mortality experienced for patients in more rural regions of the state compared to urban regions."

**Comment**
I have concerns about the suggestion that urban or rural areas or even time to a medical center are factors to be considered to improve target screening. Screening is ideally done on all high

risk patients. However, this study identifies hurdles to screening intervention strategies and perhaps lack of access. Secondly, an aim of screening is to reduce the incidence of advanced stage disease which did not show any significant difference between urban or rural areas.

**Response**

Thank you for this comment. We agree with you that, as a service line, we need to increase lung cancer screening broadly. We addressed the issues with screening compliance at the primary care level in lines 214-223. Also, in lines 233-234 we have added “It should be stressed, however, that lung cancer screening rates are very low across all demographics and should be increased broadly.”

**Comment**

No information is given about the state of low dose CT screening in Michigan. When was it instituted? Who has access? How and when do the authors hypothesize that it will impact outcome in Michigan? if it was stated in recent years, why is all of the data from 1985 used to assess its impact?

**Response**

This is an important point. We have now added information about lung cancer screening in Michigan. Unfortunately, rates are low and consistent with national trends. In lines 236-238 we have added “Most institutions in Michigan began their lung screening programs several years after this approval. Overall screening rates in Michigan are relatively low and consistent with national levels of screening.”

**Reviewer B**

**Comment**

This paper is well written and valid from a methodological point of view. The weak point is linked, in my opinion, to the little interest it can arouse in JTD readers if they do not belong to the state of Michigan or at most if they are not Americans. The problem is therefore represented by the doubt about the generability of these data, considering that an extremely important variable such as smoking history is missing from the study. You yourselves emphasize in the text the importance of complete data collection. Your study examines a very long period of time in which many things have changed from a diagnostic and therapeutic point of view such as the introduction into clinical practice of PET, screening programs with low-dose CT, target therapies and immunotherapy or the use of minimally invasive surgical approaches. I wonder if it wouldn't be useful to divide the study into various time periods based on the most important innovations and then compare the results.

**Response**

Thank you for this comment. We do feel that the benefit of our study is that every case of lung cancer is included. In addition, the innovations that you mention (e.g., PET-CT) are generalizable and have occurred broadly. But we addressed the differences that may have occurred over the years in lines 272-275, saying “As our study was conducted over a very long time period, it is important to note the changes in treatment patterns that has occurred. During the later part of the study more targeted treatments became available. In addition, novel techniques

such as minimally invasive surgery have continually increased in prevalence and have allowed for resection with less morbidity."

**Reviewer C**

**Comment**
This is an interesting report including relevant conclusions for lung cancer prevention and diagnosis in the authors' area. I have a few comments and suggestions and I thanks the authors for reading and considering them. Driving time is divided in five: 0-15, 15-30, 30-60, etc. Every time period includes the previous and the next ones. Thus, patients driving 15 minutes can be included in first or the second subgroups. I believe you should amend this classification.

**Response**
Thank you for pointing out our error in categorization. When we performed our analysis, we used the breakdown you mention. We have now clarified in lines 131-135 "Hospital service areas were calculated in Network Analyst ArcGIS Pro using the Michigan Street Network provided by ESRI and distance-time breakpoints of 0 to less than 15 minutes, 15 minutes to less than 30 minutes, 30 minutes to less than 60 minutes, 60 minutes to less than 90 minutes and 90 minutes to 120 minutes."

**Comment**
As in most reports including matching cases and controls by psm you are missing a high percent of patients in the analysis, more than 29,000 in your study. This limitation should be acknowledged in the text and you could include some of the references addressing this problem (I'm kindly suggesting Shiba K, et al, Using Propensity Scores for Causal Inference: Pitfalls and Tips. J Epidemiol. 2021 Aug 5;31(8):457-463. doi: 10.2188/jea.JE20210145).

**Response**
This is an excellent point and we have included this reference and discussion in the limitations. In lines 294-296 we have now added "Also, there were many patients that were excluded from analysis due to incomplete data. Some studies have suggested propensity matching for missing data, but weigh the advantages and disadvantages of this approach (33)."

**Comment**
Matching by the inverse probability of treatment weighting is an alternative to the previous comment.

**Response**
Thank you for this reference and suggestion. We added the previous reference discussing pros and cons of handling missing data. I do agree that it is important to add this to the limitations one way or another.

**Comment**

I'm wondering if an alternative to your statements in lines 200-204 could be that patients in rural areas with a driving time up to 30 minutes are misclassified and they should be considered as residing in urban zones.

**Response**
This is a good point that a 30 minute drive may not necessarily be considered "rural." We have removed that classification from those lines.

**Comment**
In Table 1, the file "Race" must be moved one line below.

**Response**
Thank you for pointing out that error. We have fixed Table 1.

**Reviewer D**

**Comment**
I consider that they have done a good job to analyze the stated objective, methodologically well planned, with good recognition of the limitations inherent to a population study of this type. However, I think the authors should make some small corrections.
• L. 162-165: Disease status at presentation: It should be clear that the percentage is based on the overall percentage of the series (141,977 p.).

**Response**
We have added the numerator and denominator to ensure that readers know that the percentages are based on the overall series.

**Comment**
L. 199-200: If advanced NSCLC were defined as regional or distant disease (l.102), then the proportion of people with advanced disease would be three-quarters.

**Response**
You are correct and we have fixed this error.

**Comment**
L. 261-267: The authors do not indicate the reference number, which is reference 32, as indicated later (l. 272).

**Response**
Thank you for pointing this out. We have added that reference number to those lines.

**Comment**
L. 286: I think the statement should be expressed in reverse: Female patients now develop the disease as frequently as male patients.

**Response**
We have changed this line per your suggestion.

**Comment**
The references do not adapt to the publication's standards, regarding page allocation.

**Response**
We have corrected all of the references to meet the publication guidelines.  Thank you.