Peer Review File
Article information: https://dx.doi.org/10.21037/jtd-22-210

Reviewer A

Comment 1: Please check spelling and grammar again.
Reply 1: We thank the reviewer for the thorough review of our manuscript. We have combed over our manuscript for any spelling or grammar errors utilizing the "Spelling & Grammar" function of Microsoft Word. We also read through our manuscript again to detect any avoidable grammar errors.
Changes in the text: We modified line 18 of page 11 within the "DISCUSSION" section to delete a duplicate "was" in the sentence. The previous line was edited to: "Notably, 68% of patients in our cohort received GCC, which was associated with improved overall survival across all ages and stages of the disease."

Comment 2: Please include a paragraph about limitations of your study.
Reply 2: We thank the reviewer and would like to bring attention to lines 18-23 of page 12, and lines 1-2 of page 13 within the "DISCUSSION" section for our paragraph on the study's limitations.
Changes in the text: No changes were made to specifically address this question as we already included a paragraph on limitations of our study.

Reviewer B

Comment 1: The authors suggest that one of the strengths of the NCDB is that it represents a large portion of the US population. Yet with the exclusion criteria applied to the 262, 806 patients with SCLC initially identified - nearly 35% are eliminated for some of the analyses (remaining n=176,453) and even more for other analyses (remaining n=138,290). As those excluded often lacked key data (e.g. race, ethnicity) - it won't be easy to compare the original sample to those analyzed to evaluate for representativeness of the final populations reported on. Perhaps the authors could mention this (missing data in 1/3 or more of those initially identified) as a limitation in their discussion.
Reply 1: We thank the reviewer for their comprehensive analysis of the manuscript. We agree with the reviewer's input and have modified our paragraph on limitations within the "DISCUSSION" section as below.
Changes in the text: We have added the following sentences to lines 1-4 of page 13 of the "DISCUSSION" section: "While 262,806 patients who were diagnosed with SCLC between 2004-2016 were initially identified, 90,353 (34.4%) were excluded from analysis due to missing key data, which may result in unpredictable bias. An additional 38,163 excluded from survival and treatment analyses as well."

Comment 2: The authors share on the bottom of page 8 and top of page 9 that the literature only provides single institution or small sample sizes of studies of young patients with SCLC. However, they didn't cite the paper summarizing data by Lara et al (https://pubmed.ncbi.nlm.nih.gov/29191590/) from the California Cancer Registry which did involve an analysis of 22,000+ SCLC patients of whom 975 were <50. Truly not the sample size or large geography of this national study - but still a reasonable sample to compare findings in the discussion. In the Lara paper, they too found that female gender and non-White race (in their case Asian) were predictors of survival and generally those <50 had better survival than those >50. I'd suggest they at least cite and comment on the Lara paper.
Reply 2: We thank the reviewer for bringing the aforementioned publication to our attention. We agree with the reviewer's input and will comment on the paper by Lara and colleagues (Lara JD, et al. *Lung Cancer*. 2017;112: 165-168) within our "DISCUSSION" section.
Changes in the Text: We have modified line 22 of page 8 within the "DISCUSSION" section to state:
"The existing data on young individuals with SCLC are limited to older data registry or single-institution retrospective studies."

We have then added the following sentences to lines 8-12 of page 9 in the "Discussion" section: "In a 2017 analysis of 22,863 SCLC patients diagnosed between 1998-2012 in the California Cancer Registry, 975 (4.2%) were <50 years of age.(29) Age <50 years was associated with significantly better cause specific survival (CSS) than those ≥50. Among those <50 years, female sex, rural residence, and Asian/Pacific Islander race were associated with significantly improved CSS while advanced stage at diagnosis was associated with worse CSS.(29)"

We additionally cited the publication by Lara and colleagues (Lara JD, et al. *Lung Cancer*. 2017;112: 165-168) without changing the text of lines 2-3 of page 10 within the "DISCUSSION" section.

Comment 3: The authors evaluate the independent contribution of many variables using cox proportional hazards models on mortality. At least two of them may be highly correlated - having private insurance and also having higher household income. Both appear to be related to a lower hazards ratio for mortality - so perhaps they can explain how they evaluated for this potential problem and handled it in the statistical analysis section on page 5.
Reply 3: The reviewer raised an important point about collinearity between insurance status and income. Earlier studies on NCDB have excluded education from the model because of collinearity with median income given that both these variables are ordinal and were estimated based on zip code. We, therefore, did not include education in the model.

We believe that insurance status is not a linear combination of income. While income is an ordinal predictor with increasing levels, insurance status is a categorical variable with non-ordinal distinct levels. We tested the codependency between income and insurance status using the "variable inflation factor (VIF)" statistic generated from a linear regression model treating the two predictors as dummy variables. Although there are no formal criteria for deciding, a VIF of >4 is regarded as having codependency. We found no co-dependency between the two predictors (VIF=1.0). We also performed a non-parametric Spearman correlation test to look at the correlation between the two variables and found no correlation (spearman r=0.02).
Changes in the text: No specific change was made in the text.

Comment 4: What might be interesting for the discussion - would be a section devoted to the need for future research to include genomic data (beyond the phenotypic/clinical data in this paper) in these age-related comparisons. Other have noted that tumor suppressor genes are frequently inactivated in SCLC (e.g. TP53) - so a comparison of the prevalence and the understanding of what drives SCLC in young vs. old patients would be of practical use in terms of evolving treatment modalities.
Reply 4: We appreciate the reviewer's point that future research into the genomic landscape of SCLC would add valuable insight into our understanding of what drives SCLC in young versus older patients. We agree and will make a statement on the need for such research. However, we believe delving into a literature review and discussion on this topic is beyond the scope of this manuscript.
Changes in the text: We added the following sentence to lines 15-17 of page 13 within the "CONCLUSION" section: "An examination of the genomic alterations in SCLC and how they pertain to age may facilitate our understanding of disease tempo, treatment response and resistance."