Peer Review File

Article information: https://dx.doi.org/10.21037/tlcr-23-266

Reviewer A

This study addresses a topic of interest, as our understanding of screening outcomes over age 75 are poorly understood. As lung cancer screening is being implemented more widely internationally, data from this population are welcome to inform new screening protocols and evolving eligibility criteria. This manuscript includes much detail on the cohort described, and the authors have gone to lengths to analyse the data available in a number of ways.

Some overarching comments:

• As acknowledged in the discussion, the sample size & event rates are limited, so interpretation of the data needs to be done cautiously.

Response:

Thank you for your valuable comments. We acknowledge that our study is limited by a small sample size, which restricts us from making definitive interpretations of the results. We agree with your perspective on this matter.(see discussion line 333)

"This led to the exclusion of a large number of participants, resulting in a small sample and statistical power."

In order to address the previously assertive tone, we have chosen to present the following description in a more subdued manner. (see conclusion line 417)

"Further research on the survival benefits of LCS with LDCT in this older population is warranted."

• The more sophisticated analyses are hampered by under powering and might benefit from simplification.

Response:

Thank you for your guidance in refining our study. We appreciate your opinion, and as a result, we have made the following changes: We have reorganized the tables, moving Table 2 and Table 3 to eTable 1 and eTable 2, respectively. We have simplified Figure 3 and made the decision to remove both Figure 3 and eTables 1, 2, and 3. Furthermore, we have decided to omit the analysis of survival between SDLC and NSDLC, and instead, we have focused on the detection of lung cancer based on age differences.

• There are grammatical errors throughout which harm readability.

Response:

Thank you for your kind comments. We would like to express our sincere apologies for the subpar quality of the initial manuscript, which contained several errors in English language usage. We have taken this matter seriously and have made significant efforts to rectify the situation. The manuscript has undergone thorough proofreading and revision, with the assistance of a native English speaker, to ensure its fluency and formal presentation. We appreciate your understanding and thank you for giving us the opportunity to improve the quality of our manuscript.

METHODS

Focusing on an age range 70-74 is reasonable, but an open-ended ≥75 age range risks being too wide to be informative, especially since some participants were aged >90. Would an upper limit be beneficial to make the groups more comparable (e.g. 75-80y), especially as screening guidelines are unlikely to dispense with age limits altogether.

Response:

Thank you for your critical comments. We appreciate your perspective, and we agree with your opinion regarding the age range for the A75 group. To address this concern, we have decided to modify our inclusion population. The A75 group will now consist of patients aged 75-80 years. By implementing this change, we aim to make the groups more comparable and improve the informativeness of our study. In line with this modification, we have conducted a re-analysis of all the comparisons associated with both the A70 and A75 groups. Additionally, we have completely rewritten the methods section, which now includes Figure 1, and restructured the results section accordingly. After adjusting these categories, we have decided to change the group names from A70 and A75 to 'early 70s' and 'late 70s,' respectively. (See methods line 117)

"Then, the patients were categorized into two age groups: the early 70s group, comprising individuals aged >70 to <75 years, and the late 70s group, comprising consisting of individuals aged >75 to <80 years."

Propensity matching was performed. Was this extra level of data manipulation necessary? Justification of this approach would be useful in the methods, including why

it was based on these 3 variables (to include BMI).

Response:

Thank you for your insightful comment. We have utilized propensity score matching (PSM) to facilitate a more robust comparison by mitigating confounding biases. PSM is a valuable technique employed in retrospective research, particularly when it is challenging to implement randomized control trials. By employing PSM, we are able to effectively adjust for confounders, minimize the impact of bias and confounding variables, and achieve greater comparability between the patient and control groups. This approach enhances the validity and reliability of our findings.

Originally, our study was designed to compare the survival outcomes between SDLC and NSDLC, with matching variables including age, sex, and BMI, which are known to be associated with survival. However, based on the insightful comments from the reviewers, it was pointed out that the comparison we initially planned may introduce lead time bias. Consequently, we have made the decision to omit the survival comparison (originally presented in Figure 4) from our analysis. Nonetheless, we have retained the analysis of lung cancer stage and characteristics comparison (as presented in the original Table 6). To ensure better comparability between the screening-detected lung cancer (SDLC) and non-screening-detected lung cancer (NSDLC) groups, we have revised our matching variables. Specifically, we now match based on age, sex, and the year of lung cancer diagnosis. This modification is significant as our study spans nearly a decade, from 2012 to 2021, during which lung cancer histology and treatment policies have undergone notable changes. Thus, we consider the year of lung cancer

diagnosis as an important factor in achieving comparability between the groups.

According to methodological considerations, a larger number of control subjects can improve study power. In our study, patients were matched using a caliper width equal to 0.01, allowing for the inclusion of five controls without replacement for each SDLC patient, taking into account the covariates of age, sex, and the year of lung cancer diagnosis. (See method line179)

"To mitigate potential imbalances in factors between patients with SDLC and NSDLC, PSM was applied to reduce selection bias between the SDLC and NSDLC groups. PSM was performed using logistic regression models, with age, sex, and year of LC diagnosis as covariates. PSM was performed at a 1:5 ratio without replacement, with a caliper value of 0.005. Following PSM, the differences between the matched groups were <0.1, indicating a successful balancing of covariates."

References.

1. Baek S, Park SH, Won E, Park YR, Kim HJ. Propensity score matching: a conceptual review for radiology researchers. Korean J Radiol. 2015 Mar-Apr;16(2):286-96. doi: 10.3348/kjr.2015.16.2.286. Epub 2015 Feb 27. PMID: 25741190; PMCID: PMC4347264.

2.Day AG. Why the Propensity for Propensity Scores? Crit Care Med. 2015 Sep;43(9):2024-6. doi: 10.1097/CCM.0000000000001175. PMID: 26274709.

3. Yu Y, Choi J, Lee MH, Kim K, Ryu HM, Han HW. Maternal disease factors associated with neonatal jaundice: a case-control study. BMC Pregnancy Childbirth.

2022 Mar 24;22(1):247. doi: 10.1186/s12884-022-04566-6. PMID: 35331174; PMCID: PMC8953140.

4. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. Pharmacoepidemiol Drug Saf. 2012 May;21 Suppl 2:69-80. doi: 10.1002/pds.3263. PMID: 22552982.

5. Rubin, 0. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41-55. 21.

A70 age range is described as 70-75 in the prose, but 70-74 elsewhere.

Response:

We deeply appreciate your valuable comments. In our LCS cohort, we categorized eligible patients into two groups based on age. The first group includes patients who are 70, 71, 72, 73, and 74 years old, while the second group comprises patients who are 75, 76, 77, 78, 79, and 80 years old. To prevent any confusion, we agree that it would be more accurate to express the age ranges as 70-74 (early 70s) and 75-80 (late 70s), respectively. We have made the necessary revisions throughout our manuscript to ensure consistency in the expression of these age ranges. (see method line 117)

"Then, the patients were categorized into two age groups: the early 70s group, comprising individuals aged >70 to <75 years, and the late 70s group, comprising consisting of individuals aged >75 to <80 years."

Line 160 – “maximum diameter ≥4mm” – does this refer to minimum diameter?

Response:

Thank you for the clarification. The revised sentence " Nodules with a diameter of ≥4 mm were defined as non-calcified lumps" effectively conveys the intended meaning and helps to avoid potential misunderstandings. (See Method line 131)

“Nodules with a diameter of ≥4 mm were defined as non-calcified lumps”

Line 242 – “significantly earlier mortality rate” – requires rephrasing

Response:

Thank you for this invaluable comment. During the revision process, we carefully deleted that part

Line 292 The curves demonstrate that patients with lung cancer had significantly earlier mortality rate than those without, across age groups, including individuals over 85 years old.

Response:

We have removed the analysis involving patients aged over 85, as you suggested in your comment.

RESULTS

Figure 1 clearly demonstrates participant flow through the study.

Response:

Thank you for your comment.

There are many Kaplan-Meier plots showing survival between different subgroups. Suggest presenting fewer, to emphasise main points, would be beneficial. Especially given the limitation of survival as an outcome in the screening setting, with lead time and length time effects limiting the validity of survival comparisons. The claim that "patients SDLC having a significantly higher survival rate than those with NSDLC (p=0.038)" is therefore misleading, especially as this is based on log-rank test rather than hazard ratio.

Response:

We sincerely appreciate this insightful comment. We acknowledge the presence of lead time bias when comparing SDLC and NSDLC in our study setting. Taking this into consideration, we have made the decision to remove the survival analysis between SDLC and NSDLC. Consequently, Figure 4 has been deleted, and corresponding sentences have been omitted from the manuscript.

Kaplan-Meier plots should display censoring with marks on the curves. Tables beneath the plots displaying the at-risk population in each group along the follow-up period would also aid interpretation.

Response:

We appreciate the reviewer's valuable suggestion. In response to this feedback, we have made significant improvements to the presentation of our Kaplan-Meier plots. We have now included censoring marks on the curves in Figure 4.

The age groups described in Figure 3 differ from those elsewhere without justification, with ranges ≥75, ≥80, ≥85. The information added by so many versions is limited.

Response:

Thank you for the careful review. We agree with your comments that we limited the age ranges to A70 and A75. We decided to delete the analysis with≥80 or ≥85. Consequently, we have removed the analysis involving age groups ≥80 or ≥85. (see figure 4)

Line 212 (and elsewhere) – "formal smokers" – does this refer to former smokers?

Response:

Thank you for your kind review. I have corrected the misspelled words. (see line 198)

DISCUSSION

"6-year cumulative probability rates of lung cancer detection were also similar between the two groups, with 2.4% in the A70 group and 6.1% in the A75 group"

– there is quite a notable difference of lung cancer detection rates of 2.4% and 6.1%. It is plausible that the older group would have a truly higher incidence, as this would be

consistent with the general epidemiology of lung cancer. To say these detection rates are similar, presumably because $p>0.05$, misses this point. The lack of statistical significance here more likely reflects the limited sample size & event rate.

Response:

Thank you for your feedback. I also agree with your opinion. According to the revised research results, the lung cancer detection probability in the early 70s group was 2.35%, while in the late 70s group it was 6.38%. Although there is no statistically significant difference, it is important to consider the limited sample size of the study, which may have affected the statistical power. In the revised manuscript, this part has been deleted.

"…the high rate of successful treatment in the A75 group…"

– Numerical differences in treatment proportions are shown in the table, especially for surgery/SABR, which are important to acknowledge. Limited sample size does not mean these differences are not real, as they are very plausibly real and likely to reach statistical significance in a bigger cohort.

Response:

Thank you for your comments. I appreciate your input, and I agree that the limited number (36) of patients diagnosed with lung cancer in our study may have impacted the statistical power. Therefore, we made the decision to focus on comparing the proportion of patients who underwent surgery between the early and the late 70s

groups, rather than discussing the overall differences in treatment approaches. We have incorporated this information into the discussion section to provide a more specific and clear analysis, thereby avoiding the ambiguity that may have arisen in the previous discussion. (see discussion, line 314)

"The proportion of patients who received surgical treatment and best supportive care based on stage was lower in the late 70s group than in the early 70s group, without showing statistical significance. Moreover, the SDLC group had earlier LC stages than the NSDLC group. Thus, this study indicates that LC can be diagnosed in the early stage even in individuals in their late 70s through LCS, although LC survival data were not shown appropriately. This finding highlights the need for further research regarding LCS in this age group."

– Caution must be used with the terminology "successful treatment" – specific treatment outcomes are not described so treatment success cannot be determined (which would likely be beyond the scope of this paper). If just stating that they received treatment appropriate for stage, this should be rephrased accordingly.

Response:

Thank you for noting this error. Instead of using the terminology "successful treatment," we have rephrased the sentence to reflect that the patients received treatment appropriate for their stage of cancer. (see result line 278)

"Moreover, the proportion of patients who received appropriate LC treatment based on stage was higher in the early 70s group (n = 23, 85.2%) than in the late 70s group (n =

6, 66.7%), without a statistical difference (p = 0.216) (Table 4).”

“Clinicians should carefully evaluate the individual patient's health status and weigh the risks and benefits of invasive procedures”

– This is a fundamental issue. Were there characteristics of the patients who received treatment in this cohort that were atypical and associated with being suitable for radical treatment? Might the authors suggest any approach to evaluating health status in this context? Were measures of health status/frailty captured in this cohort?

Response:

Thank you for your insightful comments. We appreciate your perspective on the issue. While we did discuss the importance of evaluating the individual patient's health status and considering the risks and benefits of invasive procedures, we acknowledge that our study did not specifically identify the characteristics of patients who received invasive work-up. In this cohort, measures of health status or frailty were not captured, limiting our ability to provide specific guidance on evaluating health status in this context. We recognize that further research is needed to develop criteria or risk factors for determining which patients would be more appropriate for invasive work-up. It is important for future studies to investigate eligible subjects for invasive work-up and explore potential risk factors associated with aggressive work-up. We have decided to remove those comments from our manuscript.

“These findings suggest that LCS can lead to the detection of lung cancer at an earlier stage in elderly patients, which may improve their survival rates.”

– Such a claim is possible, but it is a stretch from the data available. Competing causes

of mortality mount in older cohorts, and the implications for lung cancer screening remain poorly understood.

Response:

We greatly appreciate your comments. We are fully aware of the limitations of our data regarding survival analysis. Therefore, we have removed any claims suggesting that screening can improve survival rates. Instead, we acknowledge the need for further research to better understand the potential survival benefits associated with screening in elderly patients. Thank you for highlighting this important point. (see discussion line 318)

"This finding highlights the need for further research regarding LCS in this age group."

"it may not be appropriate to apply the smoking history criteria for NLST to the elderly population."

– This is perhaps at odds with the observation in the previous paragraph that describes a strong association with current smoking. P=0.069 may well just be due to sample size limitation, so should not be over-interpreted with respect to existing guidelines.

Response:

Thank you for your comments.we have omitted the discussion section and removed the statement suggesting that it may not be appropriate to apply the smoking history criteria for NLST to the elderly population.

Reviewer B

Title:

- consider removing "scan"

Response:

Thank you for your suggestion. We have removed the word "scan" from the title.

- running title implies the lung cancer is elderly- consider rephrasing.

Response:

Thank you for this important comment. We changed running title to “Lung cancer screening for the older population” (See line 23)

Abbreviation:

NELSON is Nederlands–Leuvens Longkanker Screenings Onderzoek

Response:

Thank you for the careful review.

Highlight box

- The authors' implications are too strong for the level of evidence described in the study. This was a small retrospective cohort study, the authors should not state that lung cancer screening should be considered in this demographic.

Response:

Thank you for your valuable comments. We have taken your feedback into consideration and made appropriate adjustments to the highlighted section. The revised

text to ensure that the implications are appropriately toned down in accordance with the study's level of evidence. (see highlight line 62)

Key findings

Lung cancer screening using low-dose chest computed tomography (LDCT) showed significant lung cancer detection rate in older people.

What is known and what is new?

- The effectiveness of lung cancer screening (LCS) using LDCT has been proven in heavy smokers aged 50–75 years.
- This study demonstrates the significant early-stage lung cancer detection from LCS in the older population, including individuals aged 75–80 years.

What is the implication, and what should change now?

- Given that LCS with LDCT could improve the early detection of lung cancer in the older population, more studies on lung cancer screening are needed in older smokers, considering the potential benefits of improved survival.

Abstract:

- There appears to be overlapping age ranges. Are participants aged 75 included in A70 or A75? It should be 70-74 if the other group is >/=75

Response:

Thank you for your kind comments. We have noted the classification of eligible patients in the LCS cohort into two age groups: the early 70s (70-74 years old) and the

late 70s (75-80 years old). To ensure clarity and consistency, we have made the necessary revisions in our manuscripts to reflect this classification. (see abstract 43)

"They were categorized into the early 70s (70–74 years) and late 70s (75–80 years) groups based on their age."

Introduction

- Line 110-112 on page 3 is not correct. The lung cancer-related mortality did not decrease after the NLST and NELSON, their long-term follow up data demonstrated a reduction in lung cancer-related mortality with LDCT screening compared to their control groups.

Response:

Thank you for your critical comments. We found the mistake and rewrite introduction part (see Introduction line 68)

"The National Lung Screening Trial (NLST) and Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) study provided compelling evidence on the reduction of LC-related mortality among individuals aged 50–75 years with a significant smoking history who underwent LC screening (LCS) using low-dose computed tomography (LDCT) compared to those undergoing chest radiography and those not undergoing screening,respectively."

- line 119 sentence needs correction re grammar and language

Response:

Thank you for your kind comments. We sincerely apologize for the subpar quality of

the initial manuscripts written in poor English. We have since thoroughly proofread and revised the manuscript with the assistance of a native English speaker, ensuring its fluency and formal presentation.(See introduction line 82)

"Similarly, the *CHEST* journal recommends screening individuals aged 55–77 years"

Methods:

- Study design line 142- can the authors please clarify what any abnormalities refers to

Response:

Thank you for your kind comments. We appreciate your comment regarding the clarification of "any abnormalities" mentioned in line 142 of the study design. In our study, "any abnormalities" refers to any suspicious findings of lung cancer observed in the X-rays taken prior to patients visiting the hospital. These patients were excluded from the study to focus on individuals from the general population rather than those already seeking medical attention. We have provided a more detailed explanation in the methods section of the manuscript for further clarity (see methods, line 111).

"To avoid including patients who underwent LDCT for purposes other than LCS, the following patient groups were excluded: those with a previous history of LC, those who underwent LDCT due to abnormalities found at an earlier chest radiography or other medical images (such as lung nodules, focal inflammation, haziness in the lung field, or other abnormalities suggestive of LC), and those who underwent LDCT for

symptoms such as severe cough, dyspnea, hemoptysis, and weight loss."

- Can the authors please specify the inclusion criteria for lung cancer screening for this cohort in the methods. Was it only age? Line 155 suggests it was requested by the patient or at the discretion of the clinician.

Response:

Thank you for your comments. In our study, the inclusion criteria for the lung cancer screening cohort were as follows: Patients who visited a pulmonologist's clinic specifically for lung cancer screening. Patients who underwent low-dose computed tomography (LDCT) between January 2012 and June 2020.Patients aged between 70 and 80 years. To minimize potential selection bias, we excluded patients who met any of the following exclusion criteria: Patients with a history of lung cancer. Patients who underwent LDCT after the detection of any abnormalities such as lung nodules, focal inflammation, haziness in the lung area, or other abnormalities found in previous chest radiography or medical images. Patients who underwent an LDCT scan after the onset of lung cancer symptoms such as severe cough, dyspnea, hemoptysis, or unexplained weight loss. (see methods line109- 116)

"Among participants who visited the pulmonologists' clinic, only patients who had their first LCS with LDCT between January 2012 and June 2020 and whose age at the initial screening was between 70 and 80 years were included. To avoid including patients who underwent LDCT for purposes other than LCS, the following patient groups were excluded: those with a previous history of LC, those who underwent LDCT due to

abnormalities found at an earlier chest radiography or other medical images (such as lung nodules, focal inflammation, haziness in the lung field, or other abnormalities suggestive of LC), and those who underwent LDCT for symptoms such as severe cough, dyspnea, hemoptysis, and weight loss."

- The definition of lung nodules on line 160 is not correct. Do you mean to say that a positive scan was a nodule >/= 4mm (i.e. minimum diameter) as per the NLST? Was there any other CT abnormalities classified as a positive scan?

Response:

Thank you for the careful review. We apologize for any confusion caused. We have revised the sentence to provide a clearer definition. The updated sentence now states, " Nodules with a diameter of ≥4 mm were defined as non-calcified lumps " (see methods line 131). This classification of positive scans only included masses or nodules, and other CT abnormalities were not considered.

- Was survival data only collected from medical records? Were any central databases/registries contacted. The follow up is quite short for participants towards the 2020 side of screening- was there a minimum follow-up time planned?

Response:

Thank you for your inquiry. I would like to clarify the data collection and follow-up process in our study. Survival data for the participants were collected from the medical records of the Veterans Health Service Medical Center (VHSMC). The VHSMC medical records are linked to the central national database maintained by the

Ministry of Patriots and Veterans Affairs, which provides comprehensive survival data for veterans in Korea (see method line136). Regarding the follow-up duration, we noted that the follow-up period towards the 2020 side of screening appears relatively short. In our study, the primary objective was to assess lung cancer detection rather than survival outcomes. Therefore, we designed a minimal follow-up period of 15 months to allow for adequate evaluation of cancer detection rates. The intention was to enroll patients who had undergone at least 1 year of follow-up and had a sufficient time frame for lung cancer work-up (approximately 3 months).

- The authors have included never-smokers - this is a completely different cohort to any published RCTs results for lung cancer screening with LDCT.

Response:

Thank you for your comments. You are correct that including never-smokers in our study introduces a difference compared to published randomized controlled trials (RCTs) on lung cancer screening with LDCT, as those trials typically focused on high-risk individuals with a significant smoking history. In our study, we aimed to investigate the potential benefits of lung cancer screening in the elderly population, regardless of their smoking history. To account for the potential influence of smoking and comorbidities, we conducted subgroup analyses and presented the results in tables 1 and e-table 2. This allows for a more detailed understanding of the impact of these factors on lung cancer detection rates within our study population.

Results

- The 510 patients >/=75 years old appear to have had screening outside of the National Cancer Screening program of Korea. Was there any explanation for this? It is likely going to be a source of bias as this population may not represent the general population of >/=75 years old.

Response:

Thank you for your comments.

Within the Korean medical system, individuals have the option to undergo lung cancer screening for health check-ups, regardless of the criteria set by the national cancer screening program. At VHSMC, we offered a lung cancer screening program specifically for participants who expressed their desire to undergo such screening. This information has been briefly mentioned in the method section of our manuscript. (see methods line108). The study population is not sufficient to fully represent the general population of individuals aged 75 years or older. We have addressed this limitation and its implications in the discussion section (see discussion line 326).

"As a result, the cohort primarily consisted of male participants who were smokers and had a higher prevalence of comorbidities compared with the general population, which introduces potential selection bias and limits generalizability."

- The authors should provide fitness/ECOG/performance status data at time of enrolment.

Response:

Thank you for your comment. We acknowledge that we did not collect fitness/ECOG/performance status data at the time of enrollment, and we understand that this information is important for a comprehensive analysis. Unfortunately, this is a limitation of our study as we do not have data available for fitness/ECOG/performance status in our cohort. We have discussed this limitation in the discussion section (See discussion line 334)

". In addition, no data on the fitness/Eastern Cooperative Oncology Group (ECOG) performance status of the participants at the time of screening were analyzed, which hinders the comprehensive evaluation of the older population."

Limitation

- Very high percentage of male participants (>90%)

Response:

Thank you for your comment. We acknowledge the high percentage of male participants in our study, with over 90% being male. This is primarily due to the nature of our study population, as our patients are veterans receiving healthcare services at our hospital. The veteran population is predominantly male, which contributes to the imbalance in gender representation. While we recognize this as a limitation, it is inherent to the specific setting of our study and beyond our control. We have discussed this limitation in the discussion section (See discussion line 323)

"The screening was primarily conducted at a pulmonologist's clinic, and this study was conducted at the largest veteran health service center in Korea, resulting in a

predominantly male group. As a result, the cohort primarily consisted of male participants who were smokers and had a higher prevalence of comorbidities compared with the general population, which introduces potential selection bias and limits generalizability”

- Line 212- formal should be former

Response:

Thank you for pointing out the error. I have corrected the misspelling of "formal" to "former" in the results section. (see result page 197)

- Do the authors have any information regarding decision making for treatment?

Response:

Thank you for your comments. The treatment decisions were made by the attending physicians, including pulmonologists, oncologists, and thoracic surgeons. The decisions were based on the NCCN guidelines and were often made in the context of multidisciplinary discussions. We have included this information in the methods section for clarity. (methods line141)

“Treatment was determined by each pulmonologist, oncologist, or surgeon according to National Comprehensive Cancer Network guidelines and decision from multidisciplinary discussions.”

- How many cancers were diagnosed during screening intervals (i.e. not related to screening)?

Response:

Thank you for your valuable questions. It appears that during the screening intervals, 3 cases of interval cancer were diagnosed in the late 70s group and 1 case of interval cancer were diagnosed in the early 70s group. (See results 266)

"In the pathologic diagnoses, the four interval cancers included two small cell LC and one squamous cell carcinoma in the late 70s group and one adenocarcinoma in the early 70s group."

- How many of the total cohort were eligible for LCS based on NLST criteria?

Response:

Thank you for your comments. In total LCS population group, 586 patients were eligible for LNST criteria. It reveals in eTable1.

Discussion

-The study is not designed to answer the question of effectiveness of lung cancer screening with LDCT in the elderly.

Response:

Thank you for your comments. We acknowledge that our study design does not allow us to definitively answer the question of the effectiveness of lung cancer

screening with LDCT in the elderly population. The level of evidence provided by our study is limited and cannot establish the survival benefit of lung cancer screening in this specific demographic. We agree that further well-designed randomized trials are necessary to provide more robust evidence on the effectiveness of lung cancer screening in the elderly. However, based on our findings, which indicate a higher detection rate and a higher proportion of early-stage lung cancers in the screened population, there is a suggestion of potential benefit from lung cancer screening with LDCT in this age group.

- It is good the authors highlighted the limitation of a short follow-up period

Response:

Thank you for your feedback. We appreciate your recognition of the limitation regarding the short follow-up period in our study. We have included this limitation in the discussion section.(See discussion line337)

"Moreover, the median follow-up period of 3.7 years was relatively short, which is shorter than follow-up periods (7–11 years) in current LCS studies"

- The authors should discuss risk of overdiagnosis and lead-time bias

Response:

Thank you for your comments. we agree to existence of the lead time bias when we comparison SDLC and NSDLC in our study setting. So we decided to delete the survival analysis between SDLC and NSDLC. So Figure 4 is deleted and

Corresponding sentences have been omitted. The concerning about overdiagnosis is bestselling topic in screening fields. Our data are sufficient to assess the potential issue of overdiagnosis. (see discussion line 375)

"The issue of overdiagnosis is widely discussed in the field of screening. Previous RCTs have reported varying proportions of overdiagnosis, ranging from 0% to 67% (25). However, our data are sufficient to assess the potential issue of overdiagnosis. Unfortunately, our study did not include a control group, which makes it challenging to calculate the exact proportion of overdiagnosis."

- The authors should compare their cohorts fitness/co-morbidities to that of the general population

Response:

Thank you for your valuable comments. Unfortunately, in our study, we did not specifically investigate or compare the fitness or co-morbidities of our cohort to that of the general population, other than respiratory diseases. Therefore, it is challenging for us to provide a direct comparison as requested by the reviewer. However, based on the characteristics of our study population, we speculate that our cohort may exhibit different characteristics compared to the general population. Since our cohort consisted of screening performed patients who visited a pulmonologist's clinic, the majority of participants were smokers. Additionally, approximately 40% of the cohort had COPD. These factors suggest that our study population may have a higher prevalence of respiratory diseases and smoking-related comorbidities compared to the general

population.

- Again the authors have strong conclusions out of keeping with the level of evidence provided. for

instance line 345-"In fact, the high rate of successful treatment in the A75 group suggests that it is worthwhile to pursue a full workup and treatment plan for elderly patients with lung cancer."

Response:

Thank you for your comments. We appreciate your comments and have taken them into consideration. Upon re-evaluation, we agree that the conclusion in line 345 is not supported by the level of evidence provided in the study. Therefore, we have removed that statement from the manuscript.

- The authors should also discuss the reasonable percentage of participants who received best supportive care first line.

Response:

Thank you for your comments. After revising the data following inclusion criteria, none of the late 70s patients received best supportive care as a first-line treatment (See table 4)

Overall

- I have only pointed our a few spelling and grammatical errors, however there were many more noted throughout the manuscript.

Response:

Thank you for your kind comments. We sincerely apologize for the subpar quality of the initial manuscripts written in poor English. We have since thoroughly proofread and revised the manuscript with the assistance of a native English speaker, ensuring its fluency and formal presentation.

- The cohort included in this study needs to be better described in order to allow for the comparisons. There are concerns about participant selection in this study.

Response:

Thank you for your comments. We will make the inclusion and exclusion criteria more explicit and clear in the revised manuscript. In our study, the inclusion criteria for the lung cancer screening cohort were as follows: Patients who visited a pulmonologist's clinic specifically for lung cancer screening. Patients who underwent low-dose computed tomography (LDCT) between January 2012 and June 2020. Patients aged between 70 and 80 years. To minimize potential selection bias, we excluded patients who met any of the following exclusion criteria: Patients with a history of lung cancer. Patients who underwent LDCT after the detection of any abnormalities such as lung nodules, focal inflammation, haziness in the lung area, or other abnormalities found in previous chest radiography or medical images. Patients who underwent an LDCT scan after the onset of lung cancer symptoms such as severe cough, dyspnea, hemoptysis, or unexplained weight loss. (see methods line111) the early and late 70s had the same inclusion and exclusion criteria; however, they were categorized based on age. We acknowledge that the comparison between SDLC and

NSDLC was focused on the stage at diagnosis and initial treatment, rather than survival outcomes.

- The results of the study needs to be viewed along side its overall design

Response:

Thank you for your comments. We have reevaluated the purpose and design of this study. We have determined that comparing the survival of NSDLC patients to SDLC patients is not appropriate due to lead time bias, and therefore, survival comparison has been excluded. This study was designed to assess the detection rate of lung cancer through screening, as well as the stage at diagnosis and treatment modalities. We have made overall modifications to the objectives and methods accordingly

- There is limited discussion about many of the limitations in this study

Response:

Thank you for your valuable comments. The discussion section has been expanded to thoroughly address the limitations of our study, including the retrospective study design, the short follow-up period, the absence of a control group of non-screening populations (non-cancer), and the lack of data on ECOG/performance/fitness. These limitations are now explicitly discussed in the manuscript to provide a more comprehensive assessment of the study's findings (see discussion page line 331)

"This study has several limitations. Owing to the retrospective setting, no comparative control group such as non-screening population or chest radiography

group was employed. To minimize selection bias, patients with respiratory symptoms were excluded. This led to the exclusion of a large number of participants, resulting in a small sample and statistical power. In addition, no data on the fitness/Eastern Cooperative Oncology Group(ECOG) performance status of the participants at the time of screening were analyzed, which hinders the comprehensive evaluation of the older population. Moreover, the median follow-up period of 3.7 years was relatively short, which is shorter than follow-up periods (7–11 years) in current LCS studies"

- The authors' conclusions are too strong based on their findings.

Response:

Thank you for your comments, we agree with your opinion. Based on our findings, we recognize that the evidence is not sufficient to make definitive claims about the effectiveness of lung cancer screening in the elderly population. As a result, we have revised our conclusions to reflect this limitation and now emphasize the need for further well-designed studies to establish the efficacy of lung cancer screening in this specific population.

(see conclusion line 411)

" This study provides valuable insights into the detection of LC in the older population through LCS with LDCT. The findings highlight the comparable role of LCS with LDCT in identifying LC in patients in their early and late 70s, with LC detection rates comparable to those reported in previous studies. Moreover, a

substantial proportion of the detected LCs was at an early stage, indicating potentially favorable treatment outcomes. Importantly, the majority of older patients diagnosed with LC through screening received appropriate treatment based on their individual LC stage. Further research on the survival benefits of LCS with LDCT in this older population is warranted.“

Reviewer C

The authors of this paper report some valuable data on lung cancer screening with LDCT in an older Korean population. I believe the results are of interest to the greater scientific / medical community, but some corrections must be made prior to publication. First and foremost, the survival analysis presented in Figure 3 is flawed. They compare lung cancer survival among screen-detected patients to non-screen detected patients, ignoring the well understood phenomenon of lead time bias. Lead time bias manifests as a survival benefit observed among screened participants, even if early detection has no impact on patients' life expectancy, due simply to the fact that we are observing the patient from an earlier point in time. For this reason, RCTs measure lung cancer mortality after *randomisation* rather than diagnosis. Unfortunately, I don't believe that there is an adequate comparison group that the authors can use to measure the mortality benefits from screening. My recommendation, then, is that they remove the survival analysis between screen-detected and non-screen-detected lung cancer (in figure 4) from the paper entirely.

Response:

Thank you for your comments. We appreciate your concerns regarding the potential for lead time bias in our study when comparing SDLC and NSDLC. After careful consideration, we have decided to exclude the survival analysis between the two groups in order to avoid this bias. As a result, we have removed Figure 4 and omitted the corresponding sentences from the manuscript

The methods are somewhat vague as to the survival analysis comparing those with and without lung cancer. The authors have not indicated who the comparison group is (I presume they are screened participants), nor have they indicated what the starting point is. Is survival being counted from diagnosis, first LDCT scan, or recruitment? If they are counting from first scan / recruitment, the authors have grouped individuals based on future information which is problematic for survival analysis. If they are counting from diagnosis, then it's unclear how they can measure survival for those without lung cancer. In any case, the analysis should probably be removed.

Response:

Thank you for your comments. In our study, there are two main comparison groups. (see methods line 105)The first comparison involves the early and the late 70s subgroups within the screening cohort. The second comparison is between the SDLC cases from the screening cohort and the NSDLC cases from the lung cancer cohort. However, we have removed the survival comparison between SDLC and NSDLC during the revision process. The survival analysis between participants diagnosed with lung cancer and participants not diagnosed with lung cancer, with the survival being counted from the first LDCT scan. (see method line 177) You can find these additional details in the Methods section of the manuscript.

One noticable omission from the paper is the issue of overdiagnosis. Given the older age of this cohort, we would expect the life expectancy of participants to be shorter, reducing the potential health gains from early detection of lung cancer. For some

participants, life expectancy may be sufficiently low that screening would be futile. For others, tumour growth/progression may be sufficiently slow that it would be unlikely to cause issues for the patient within their remaining life. While I agree with the authors that clinicians "should not assume that elderly patients cannot tolerate or benefit from aggressive treatment," we should also not assume that early detection will always be of benefit.

Response:

Thank you for your comment. We agree with your comments. We should also not assume that early detection will always be of benefit. Telling about the overdiagnosis, previous randomized controlled trials (RCTs) have reported a range of overdiagnosis proportions, varying from 0% to 67%. This proportion is typically estimated by calculating the difference between the number of screen-detected cancers and the cumulative incidence of cancers in the absence of screening. Unfortunately, The concerning about overdiagnosis is bestselling topic in screening fields. Our data are sufficient to assess the potential issue of overdiagnosis. (see discussion line 375)

"The issue of overdiagnosis is widely discussed in the field of screening. Previous RCTs have reported varying proportions of overdiagnosis, ranging from 0% to 67% (25). However, our data are sufficient to assess the potential issue of overdiagnosis. Unfortunately, our study did not include a control group, which makes it challenging to calculate the exact proportion of overdiagnosis."

We acknowledge the limitation that our study cannot definitively conclude whether early detection through lung cancer screening improves patient outcomes, even in the absence of pathological overdiagnosis. Considering this, we have made the decision to remove the statement "should not assume that elderly patients cannot tolerate or benefit from aggressive treatment" from our manuscript.

Minor comments:

- Authors should specify the institution where the patients were identified.

Response:

Thank you for your comment. The patients included in this study were identified from the VHS Medical Center. We have specified this information in the Methods section of the manuscript (see methods line 104).

"This single-center retrospective cohort study was conducted at a teaching hospital, Veteran Health Service Medical Center (VHSMC), in Seoul."

- Page 3, lines 108-112:

"After the NLST and NELSON study... the lung cancer mortality rate decreased..." This phrasing makes it sound as though these studies reduced lung cancer mortality rates population-wide. Recommend saying "The NLST and NELSON study demonstrated LCS with LDCT can reduce lung cancer mortality among adults aged 50-75 with a heavy smoking history..." or something along those lines.

Response:

Thank you for your comments. We rewrite that part like this “The National Lung Screening Trial (NLST) and Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) study provided compelling evidence demonstrating a reduction in all-cause lung cancer-related mortality among individuals aged 50-75 years with a significant smoking history who underwent lung cancer screening (LCS) using low-dose computed tomography (LDCT) scans.” (see introduction line 72)

“The National Lung Screening Trial (NLST) and Nederlands-Leuvens Longkanker Screenings Onderzoek (NELSON) study provided compelling evidence on the reduction of LC-related mortality among individuals aged 50–75 years with a significant smoking history who underwent LC screening (LCS) using low-dose computed tomography (LDCT) compared to their control groups..”

- Page 6, line 249 (and elsewhere):

Relative risks should be called rate ratios (or preferably cumulative rate ratios), as the authors are comparing rates and not risks.

Response:

Thank you for your comments. We revised our manuscript as your comments. (see eTable 2)

- Page 7, line 274:

Subheading says "LC Detection Rate according to age" but text discusses more than

just age.

Response:

Thank you for your comments. We have revised the subheading as "*Characteristics and treatment approaches for older individuals with SDLC"* This change has been made in the Results section of the manuscript (see result line 259)

- Page 8, line 366:

"The lung cancer diagnosis rate of the current smoker population in the A75 group was significantly higher than that of never or former smokers in the A75 group... Therefore, it is never too late to quit smoking."

Though I don't disagree with this statement, the authors have no considered that the former smokers in the A75 group may have quit many years prior to screening. The conclusion, therefore, does not necessarily follow from the data.

Response:

Thank you for your comments. We appreciate your feedback, and we have taken it into consideration. Upon reviewing our manuscript, we acknowledge that there may have been logical leaps in our arguments. Therefore, we have made the decision to remove the aforementioned statement.

- Page 8-9, lines 372-374:

The authors are restating the results with no commentary or interpretation. Suggest removal from discussion.

Response:

Thank you for your comments. we have taken your opinion into consideration by removing the mentioned part.

- Page 9, lines 375-376:

Again, the authors are restating the results with little additional commentary. Suggest removal from discussion.

Response:

Thank you for your comments. We removed that part from discussion.

Reviewer D

• Methods: authors must specify the inclusion criteria used for screening

Response:

Thank you for your comments. The inclusion criteria for the lung cancer screening cohort were as follows: Patients who visited a pulmonologist's clinic specifically for lung cancer screening. Patients who underwent low-dose computed tomography (LDCT) between January 2012 and June 2020.Patients aged between 70 and 80 years. To minimize potential selection bias, we excluded patients who met any of the following exclusion criteria: Patients with a history of lung cancer.Patients who underwent LDCT after the detection of any abnormalities such as lung nodules, focal inflammation, haziness in the lung area, or other abnormalities found in previous chest radiography or medical images. Patients who underwent an LDCT scan after the onset of lung cancer symptoms such as severe cough, dyspnea, hemoptysis, or unexplained weight loss. (see methods line 109)

"Among participants who visited the pulmonologists' clinic, only patients who had their first LCS with LDCT between January 2012 and June 2020 and whose age at the initial screening was between 70 and 80 years were included. To avoid including patients who underwent LDCT for purposes other than LCS, the following patient groups were excluded: those with a previous history of LC, those who underwent LDCT due to abnormalities found at an earlier chest radiography or other medical images (such as lung nodules, focal inflammation, haziness in the lung field, or other abnormalities

suggestive of LC), and those who underwent LDCT for symptoms such as severe cough, dyspnea, hemoptysis, and weight loss."

• DataCollection: Authors should define interval cancers as cancers diagnosed between two rounds of screening and rather than using the term prevalent cancer use the term cancer diagnosed at baseline, in the first round, in second round….

Response:

Thank you for your valuable comments.

We have modified the definition of interval cancer, prevalence cancer in our manuscript and table4 and 5. (see methods, line 151)

"Prevalent cancer referred to LC diagnosed based on lung abnormalities such as nodules detected in round 1. Subsequent screening-detected cancer was defined as LC diagnosed based on lung abnormalities or nodules detected in subsequent screening rounds except for round 1. Conversely, an interval cancer was defined as LC that arise between scheduled LCS, with the previous screening showing a negative or "normal" result. These cancers are detected in the interval period between screenings and may be cases that were either missed or not detectable in previous screenings"

• Results:

1. Lung cancer detection rate: authors should specify that 3.1% is the cumulative

incidence. Subgroup analysis based on NLST smoking criteria is unnecessary, samples are too small

Response:

Thank you for your comments. We omitted the subgroup analysis based on NLST smoking criteria.

2. Lung cancer detection per 1000 person Year: unnecessary paragraph to be deleted

Response:

Thank you for your comments. We believe that this data provides valuable information about the effectiveness of lung cancer screening over a specific time period. We have decided to keep this part in the manuscript, as it aligns with the approach taken in other studies such as NLST and NELSON.

References>

1.National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409.

2. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. New England Journal of Medicine. 2020;382(6):503-13

. Lung nodule and lung cancer detection in each round of CT scan: the authors must

specify the percentage of participation in the different rounds of screening according to the baseline. Do not use the term prevalent but rate of cancers diagnosed in the 4th round

Response:

We have made the necessary revisions to the manuscript based on your suggestions. We have included the percentage of participants in each round relative to the number of baseline participants. Additionally, we have adopted the concepts of prevalence and interval lung cancer from the NLST team, correcting any incorrect terminology that may have been present. These changes have been incorporated into the section on lung nodule and lung cancer detection in each round of CT scan, specifically. (See Table 4)

Reference

1. Schabath MB, Massion PP, Thompson ZJ, Eschrich SA, Balagurunathan Y, Goldof D, Aberle DR, Gillies RJ. Differences in Patient Outcomes of Prevalence, Interval, and Screen-Detected Lung Cancers in the CT Arm of the National Lung Screening Trial. PLoS One. 2016 Aug 10;11(8):e0159880. doi: 10.1371/journal.pone.0159880. PMID: 27509046; PMCID: PMC4980050.

4. Cumulative lung cancer detection probability: pointless paragraph to be deleted

Response:

Thank you for your comments. However, we have decided to keep the paragraph on the cumulative lung cancer detection probability because we believe it provides valuable information about the overall impact of screening over time. In addition to capturing

the cancers detected in the initial screening, we wanted to present the cumulative occurrence of lung cancer as screening continued. We believe that this analysis contributes to a more comprehensive understanding of the screening outcomes.

5. Lung cancer detection rate according to age. The results of eTable 1 and Table 5 should be described in this paragraph.

Response:

Thank you for your comments. We agree with your assessment that eTable 1 may not be directly relevant to the study. Therefore, we have decided to remove it from the manuscript. Additionally, we have provided a more detailed description of table 5 to ensure its relevance and clarity. Also in the revised version, Table 5 has been updated and changed to Table 4. (see results line 260)

"Table 4 shows the cancer classification according to the screening round, LC diagnostic methods, LC stage, and types of treatment according to age groups. The cancer classification (prevalent, subsequent screening-detected, and interval cancer) is shown in Table 4. Four interval cancers were detected, 3 of 9 (33.3%) in the late 70s group and 1 of 27 (3.7%) in the early 70s group, with a higher proportion in the late 70s group than in the early 70s group ($p = 0.01$). In the pathologic diagnoses, the four interval cancers included two small cell LC and one squamous cell carcinoma in the late 70s group and one adenocarcinoma in the early 70s group.

Among patients diagnosed with LC in the early 70s group, only one did not undergo biopsy because of poor general medical condition. All patients with LC in the

late 70s group underwent biopsy. The proportion of patients who underwent biopsy was not statistically different between LCs detected in the early and late 70s groups (p = 0.379). Among the patients in the early 70s group, 15 (55.6%) presented with adenocarcinomas, whereas 4 (44.4%) in the late 70s group presented with squamous cell carcinomas (p = 0.442). Moreover, 15 (55.6%) and 4 (44.4%) patients in the early and late 70s groups were diagnosed with stage I LC, and 4 (14.8%) and 3 (33.3%) patients with LC in the early and late 70s groups were diagnosed with stage IV cancer, respectively. The stage distribution between the early and late 70s was not statistically different (p = 0.666). Regarding the initial treatment, 16 (59.3%) and 2(22.2%) patients in the early and late 70s groups received surgical treatment, respectively (p = 0.054). Moreover, the proportion of patients who received appropriate LC treatment based on stage was higher in the early 70s group (n = 23, 85.2%) than in the late 70s group (n = 6, 66.7%), without a statistical difference (p = 0.216) (Table 4)."

6. Overall survival of patients with lung cancer and those without lung cancer: unnecessary paragraph to be deleted

Response:

Thank you for your comments. We have decided to maintain this analysis for the following reasons. Among those who argue against the usefulness of screening in the elderly, there is a belief that elderly individuals have a limited life expectancy even without lung cancer, and that the diagnosis of lung cancer does not significantly affect their life expectancy. However, we want to emphasize that patients diagnosed with lung

cancer experience a decrease in life expectancy compared to those without a lung cancer diagnosis. It is crucial to highlight the importance of not neglecting their care. Therefore, we have chosen to keep figure 4 in its original form. However, we have made adjustments to the age groups to make the graph less cluttered. The patients have been divided into three groups: the overall patient population, the 70-75 age group, and the 75-80 age group (figure 4).

• Discussion

The discussion needs to be restructured. It is necessary to group the paragraphs concerning the incidence of cancers then to discuss the stages (the probability at 6 years of the cancer detection rate is not of interest here, specify that the 38.1% of stage I are located in the lower zone of the other RCTs) then discuss the treatments (the absence of improvement in the survival of the A75 group is probably explained by the reduction in surgery in this group) and the survival

Response:

Thank you for your response. It's great to hear that you've taken the comments into consideration and made the necessary revisions. The restructuring of the discussion section and the omission of the probability at 6 years of the cancer detection rate have been implemented.

Same remark as in the results section concerning the uselessness of the NLST subgroup and the misuse of the term prevalent

Response:

We have taken your comments into consideration. We have removed the analysis of the NLST eligible smoking group from eTable 1 and 2, as you suggested. We also acknowledge the incorrect use of the term "prevalent" and have made the necessary corrections. We have included the corrected terminology and relevant information in the manuscript under the section " Prevalent cancer referred to LC diagnosed based on lung abnormalities such as nodules detected in round 1" (See methods line 151)

Rerefence

Schabath MB, Massion PP, Thompson ZJ, Eschrich SA, Balagurunathan Y, Goldof D, Aberle DR, Gillies RJ. Differences in Patient Outcomes of Prevalence, Interval, and Screen-Detected Lung Cancers in the CT Arm of the National Lung Screening Trial. PLoS One. 2016 Aug 10;11(8):e0159880. doi: 10.1371/journal.pone.0159880. PMID: 27509046; PMCID: PMC4980050.

The paragraph on NLST subgroups is useless

Response:

Based on your comment, we have decided to remove the paragraph discussing the lung cancer detection rate according to the NLST eligible smoking criteria. (see eTable 1)

The eTable2 table is useless like the paragraph about CPCs.

Response:

We appreciate your comments. We have considered your comment and have decided to remove eTable2 from the manuscript.

In the limitations of the study, specify that it is a retrospective study.

Response:

We have taken your feedback into consideration and have revised the discussion section to explicitly mention that our study is a retrospective study. We acknowledge this limitation and emphasize it in the limitations section to provide a comprehensive understanding of the study design. (see discussion line 399)

"Because of the retrospective design, lead time bias is possible when we compare SDLC and NSDLC; thus, our data could not show the survival benefit of screening."

• The number of tables and figures should be reduced

Response:

To address this concern and improve the clarity of the manuscript, we have made the following changes:

We have removed Figure 3 from the manuscript. We have converted Table 2 and Table 3 into eTable 1 and eTable 2, respectively. We have added a renewed and simplified Table 2 to present the relevant information more effectively. These modifications aim to streamline the presentation of data, maintain readability, and avoid unnecessary distractions.

Reviewer E

The authors report on a retrospective cohort study of low-dose CT screening in an elderly population. The aim of the study was to determine the impact of screening on the detection of lung cancers in individuals aged 70 years and older. The primary endpoint was the cancer detection rate in individuals aged 70-75 and >75 years. It is well known, that that lung cancer incidence increases with age. The fact that screening is recommended up to the age of 80 by the USPSTF acknowledges that benefits are also expected for older people.

The authors assess the differences in survival of patients who underwent screening and those who did not. However, survival as an endpoint of screening studies must be considered with caution, as it is prone to lead time bias. When cancers are detected at an earlier time point by screening, the survival time from date of diagnosis is longer, regardless of whether the patient lives longer or not.

Response:

Thank you for your valuable comments. We appreciate your concern regarding lead time bias and the need for survival analysis. Upon careful consideration, we acknowledge the presence of lead time bias in our comparison between SDLC and NSDLC within our study setting. Consequently, we have made the decision to remove the survival analysis between SDLC and NSDLC from our study. As a result, Figure 4 has been deleted, and the corresponding sentences have been omitted.

Furthermore, the authors do not address the problem of overdiagnosis by screening,

which is an important issue, particularly in elderly patients with limited remaining life expectancy.

Response:

The concerning about overdiagnosis is bestselling topic in screening fields. Our data are sufficient to assess the potential issue of overdiagnosis. (See discussion line 375)

"The issue of overdiagnosis is widely discussed in the field of screening. Previous RCTs have reported varying proportions of overdiagnosis, ranging from 0% to 67% (25). However, our data are sufficient to assess the potential issue of overdiagnosis. Unfortunately, our study did not include a control group, which makes it challenging to calculate the exact proportion of overdiagnosis."

There is a range of research on the optimal population for screening, partly with quantitative modelling studies. This should be reflected in the paper.

The question, whether older patients receive the appropriate treatment is relevant and could probably be regarded in more depth with the available data.

Response:

Thank you for your comments. We have also investigated whether elderly patients receive appropriate treatment. In our study, we examined the proportion of patients receiving stage-appropriate treatment in both the early 70s and late 70s groups, which were 85.2% and 66.7% respectively. These results are presented in Table 4. We provided the rates of surgical resection, SABR(stereotactic ablative radiation therapy),

and best supportive care, and discussed the extent of surgical intervention in the discussion section based on these findings. And in the discussion section, we discuss the appropriate treatment in early and late 70s. (see discussion line 379)

The authors should consider the above-mentioned aspects in the interpretation of their study results.

Minor comments are made in a separate file. The manuscript should be checked for spelling (typos) and punctuation (spaces).

Response:

We deeply apologize for the initial manuscript's poor quality and the use of inadequate English language. We have since thoroughly proofread and revised the manuscript with the assistance of a native English speaker, ensuring its fluency and formal presentation.

Reviewer F

In this study, the authors investigated the effectiveness of lung cancer screening with low-dose CT scans in elderly individuals aged 70 years or older. Older individuals experience the highest risk for lung cancer incidence and mortality, and stand to be the most likely to benefit from lung cancer screening. However, the appropriate selection of older adults for screening remains a complex issue because of their limited life expectancy and multiple comorbidities and severe diseases. The present study suggests that lung cancer screening with LDCT may be effective in detecting lung cancer on elderly individuals, and the upper age limitation should be reconsidered. The conclusions seem to be appropriate and well summarized.

Response:

Thank you for your valuable comments. agree with your opinion, and as a result, we have revised the upper age limitation for the inclusion criteria in the screening group to 80 years old. Additionally, we have modified the definition of the 'late 70s' group (previously labeled as A75) to now include participants aged 75 to 80. These adjustments have been made in response to your suggestion. (see figure 1 and method line 117)

"Then, the patients were categorized into two age groups: the early 70s group, comprising individuals aged >70 to <75 years, and the late 70s group, comprising consisting of individuals aged >75 to <80 years. "

Major comments

1.study design and participants

1.1 If I understand correctly, the population the authors chose to study came from veterans. From what I understand, the majority of veterans are older men, many of whom are current or former smokers and also have multiple diseases with their own unique lung cancer characteristics (e.g. fewer patients have early-stage lung cancers). Could these results be generalizable to other older populations? I think this is a limitation that needs to be addressed. How do the authors ensure that this veteran is comparable to the matched population from the pulmonary clinic, where just age, gender and BMI do not seem to be sufficient?

Response

Thank you for your valuable comments. Unfortunately, in our study, we did not specifically investigate or compare other co-morbidities of our cohort to that of the general population. Therefore, it is challenging for us to provide a direct comparison as requested by the reviewer. The enrolled participants predominantly consisted of heavy smokers and were predominantly male, with a prevalence of approximately 40% for COPD. Therefore, it cannot be considered representative of the general elderly population. We acknowledge this as a limitation of our study and have addressed it in the discussion section. (see discussion line 325)

"As a result, the cohort primarily consisted of male participants who were smokers and had a higher prevalence of comorbidities compared with the general population, which introduces potential selection bias and limits generalizability."

1.2 In addition, the survival and mortality benefit of LC screening is difficult to estimate in the veteran population because of the high number of comorbidities. How can the authors ensure the comparability between veteran population and patients from the pulmonary clinic, where matching based on age, gender and BMI do not seem to be sufficient? How to eliminate or minimize the inherent differences in these two populations is also, I think, a limitation.

Response:

Thank you for your comments. The lung cancer patients from the matched lung cancer cohort in the screening group were also veterans, and both cohorts were derived from the same hospital, establishing their comparability.

1.3 What was difficult for me to understand. The author defined screening-detected lung cancer (SDLC) and non-screening detected lung cancer (NSDLC) cases. The NSDLC patients were diagnosed after symptomatic development. The author used a 1:10 propensity matching based on age, sex, and body mass index(BMI) to ensure comparability between the SDLC and NSDLC groups (Figure1). Is there a basis for choosing 1:10? Has a power analysis been done? I do not understand the use of the propensity score, which cases were selected? But: the conclusion is that there are 48 cases of lung cancer in de screened group and 480 in the symptomatic group.

Response:

Thank you for your valuable comments. We use propensity score matching

(PSM) for better comparison without confounding bias. PSM is used in retrospective research when it is difficult to apply randomized control. PSM can effectively adjust for confounders, reduces the effects of bias and confounding variables, make more comparability between patients and control. In our study, to make better comparability between screening detecting lung cancer and non-screening detecting lung cancer about characteristics and survivals, PSM was used.

Originally, our study was designed to compare the survival outcomes between SDLC and NSDLC, with matching variables including age, sex, and BMI, which are known to be associated with survival. However, based on the insightful comments from the reviewers, it was pointed out that the comparison we initially planned may introduce lead time bias. Consequently, we have made the decision to omit the survival comparison (originally presented in Figure 4) from our analysis. Nonetheless, we have retained the analysis of lung cancer stage and characteristics comparison (as presented in the original Table 6). To ensure better comparability between the screening-detected lung cancer (SDLC) and non-screening-detected lung cancer (NSDLC) groups, we have revised our matching variables. Specifically, we now match based on age, sex, and the year of lung cancer diagnosis. This modification is significant as our study spans nearly a decade, from 2012 to 2021, during which lung cancer histology and treatment policies have undergone notable changes. Thus, we consider the year of lung cancer diagnosis as an important factor in achieving comparability between the groups. According to methodological considerations, a larger number of control subjects can improve study power. In our study, patients were matched using a caliper width equal

to 0.01, allowing for the inclusion of five controls without replacement for each SDLC patient, taking into account the covariates of age, sex, and the year of lung cancer diagnosis. (see methods line 179)

"To mitigate potential imbalances in factors between patients with SDLC and NSDLC, PSM was applied to reduce selection bias between the SDLC and NSDLC groups. PSM was performed using logistic regression models, with age, sex, and year of LC diagnosis as covariates. PSM was performed at a 1:5 ratio without replacement, with a caliper value of 0.005. Following PSM, the differences between the matched groups were <0.1, indicating a successful balancing of covariates."

References.

1. Baek S, Park SH, Won E, Park YR, Kim HJ. Propensity score matching: a conceptual review for radiology researchers. Korean J Radiol. 2015 Mar-Apr;16(2):286-96. doi: 10.3348/kjr.2015.16.2.286. Epub 2015 Feb 27. PMID: 25741190; PMCID: PMC4347264.

2.Day AG. Why the Propensity for Propensity Scores? Crit Care Med. 2015 Sep;43(9):2024-6. doi: 10.1097/CCM.0000000000001175. PMID: 26274709.

3. Yu Y, Choi J, Lee MH, Kim K, Ryu HM, Han HW. Maternal disease factors associated with neonatal jaundice: a case-control study. BMC Pregnancy Childbirth. 2022 Mar 24;22(1):247. doi: 10.1186/s12884-022-04566-6. PMID: 35331174; PMCID: PMC8953140.

4. Rassen JA, Shelat AA, Myers J, Glynn RJ, Rothman KJ, Schneeweiss S. One-to-many propensity score matching in cohort studies. Pharmacoepidemiol Drug Saf. 2012

May;21 Suppl 2:69-80. doi: 10.1002/pds.3263. PMID: 22552982.

5. Rubin, 0. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41-55. 21.

2.data collection

2.1 Are data obtained on the presence of a history of other diseases (e.g. cardiovascular disease)? Because older patients are a heterogeneous group with varying burden of comorbidities and functional status, it is not just lung cancer that can affect survival and mortality rates. That information should be added to table 2.

Response:

Thank you for your comments We agree with your comments that the current dataset does not allow us to determine whether the patients are representative of the general population. However, the fact that these patients visited a respiratory department in a teaching hospital suggests that they are likely to have a higher prevalence of comorbidities compared to the general population. Additionally, considering that 90% of the study participants were current or former smokers, it is reasonable to assume a higher smoking rate than the general population, indicating the inclusion of a certain proportion of high-risk elderly patients. Unfortunately, the analysis of other co-morbidities is beyond the scope of our study, as we did not collect additional data on these conditions. In that reason, we were unable to adjust or present data regarding other co-morbidities, which is a limitation of our study. As a result, we made the decision to remove the survival analysis between SDLC and NSDLC from

our analysis.

2.2 The authors mention that for subjects in poor health condition, the diagnosis of lung cancer is based on the radiological imaging findings. How accurate is this? Is there sufficient confidence that it can be distinguished from benign diseases that present similarly to lung cancer (e.g., tuberculosis)?

Response:

Thank you for your comments. In one case, the diagnosis was based on the continued growth of the lesion on follow-up CT scans and the determination by a multidisciplinary team that included a radiologist, who diagnosed it as lung cancer. Tuberculosis was ruled out through sputum tests and other diagnostic evaluations.

3. results

3.1 Overall, the results section should be framed a little more logically and organized. The part of "lung cancer detection rate"," Lung Nodule and Lung Cancer Detection in Each Round of CT Scan"," Lung Cancer Detection Rate According to Age" can be restructured in a more rational way.

Response:

Thank you for your valuable comments. Based on your guidance, we have revised the subtitles of the results section as follows:

3.1 Characteristics of the screened populations (See line 190)

3.2 LC detection during screening (see line 217)

3.3 Characteristics and treatment approaches for older individuals with SDLC (see line 259)

3.4 Survival comparison between patients with LC and individuals without LC (see line 281)

3.5 Differences in LC characteristics and treatment between the screened and non-screened groups (see line 290)

3.2 I do not understand table 2: Lung-RADS score

1 9 of 680 (1.3)

2 11 of 659 (1.7)

3 3 of 102 (2.9)

4A 10 of 59 (17.0)

4B 8 of 23 (34.8)

The 9 cases have lung cancer stage 1, but the 680? Do they also have lung cancer stage 1? Which group these are from? Then remarkeble that 20 out of 48 have an early stage lung cancer in screening where the others (28 out of 48) have a late stage in screening.

Response:

Thank you for your comments. In our study, the Lung-RADS classification was assessed based on the initial LDCT scans. We acknowledge that the ability to determine the lung cancer staging based solely on the Lung-RADS scores presented in this table is limited. Upon further consideration, we recognized that this information may not be crucial, and as a result, we have decided to delete it. To avoid any potential

misunderstandings, we only evaluated the Lung-RADS scores for patients with detected nodules, and we excluded Lung-RADS 1 (indicating negative nodules) from the table. The percentages presented in the table represent the proportion of corresponding Lung-RADS scores among the identified lung nodules. (see table 1)

3.3 Table 3 can be deleted, is not of interest

Response:

We appreciate your comments. Although this study is retrospective in nature, we aim to demonstrate the lung cancer detection through screening based on person-years, accounting for time and population. This approach has been employed in previous studies such as NLST and NELSON, and we consider it an important aspect of our study's key findings

Reference>

1. de Koning HJ, van der Aalst CM, de Jong PA, Scholten ET, Nackaerts K, Heuvelmans MA, et al. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial. New England Journal of Medicine. 2020;382(6):503-13

2. National Lung Screening Trial Research T, Aberle DR, Adams AM, Berg CD, Black WC, Clapp JD, et al. Reduced lung-cancer mortality with low-dose computed tomographic screening. N Engl J Med. 2011;365(5):395-409.

3.4 Table 4: definition of prevalent cancer is not correct, should be adapted. In addition, it should be mentioned that there are missed cancers, 8 out of 12 in the second round and 3 out 5 in the third round.

Response:

Thank you for your important comments. We adapted the concept of prevalent and interval lung cancer from the NLST team, and we corrected the incorrect terminology. We add it in the manuscripts of Lung nodule and lung cancer detection in each round of CT scan (see line 149) and table 4. As an alternative to evaluating "missed cancer," we conducted an analysis of interval cancer between screenings. In the screening population, we identified 4 cases of interval cancer. (see line 241 and figure 2)

"Regarding interval cancer, one interval cancer was found between rounds 1 and 2 in the early 70s group, and three interval cancers (1 between rounds 1 and 2, 1 between rounds 3 and 4, and 1 between rounds 4 and 5) were diagnosed in the late 70s group"

Reference>

Schabath MB, Massion PP, Thompson ZJ, Eschrich SA, Balagurunathan Y, Goldof D, Aberle DR, Gillies RJ. Differences in Patient Outcomes of Prevalence, Interval, and Screen-Detected Lung Cancers in the CT Arm of the National Lung Screening Trial. PLoS One. 2016 Aug 10;11(8):e0159880. doi: 10.1371/journal.pone.0159880. PMID: 27509046; PMCID: PMC4980050.

3.5 Tables 2-5 are about the screened cases. Overall, I do not understand why these data are stratified by age.

Response:

Thank you for your comments. Our study aimed to compare the outcomes and

characteristics of two specific groups: the screened patients (the early 70s), whose ages ranged from 70 to 74, and the other patients (the late 70s), whose ages ranged from 75 to 80. By stratifying the data by age, we were able to examine any potential differences or similarities between these two groups in terms of lung cancer detection, nodule characteristics, treatment approaches, and other relevant factors.

3.6 eTable 1 lists the differences in the initial characteristics of lung nodules in patients with and without lung cancer. However, this part of the data on lung nodules does not appear to be very directly relevant to the study.

Response:

Thank you for your consideration. We have decided to remove eTable 1 from the manuscript.

4. discussion

4.1 In the third paragraph, the author mentions the lung cancer survival rate of patients with SDLC did not significantly differ from that of patients with NSDLC (Figure 4). I think the explanation given by the authors is appropriate, but other factors that may affect life expectancy and survival duration (such as history of disease) should also be taken into account.

Response:

Thank you for your valuable comment. Upon careful revision, we have come to the conclusion that a direct comparison between SDLC and NSDLC in our study setting is not feasible. Therefore, we have made the decision to delete the survival analysis between SDLC and NSDLC, including Figure 4, from our discussion.

Additionally, we have omitted the corresponding sentences to ensure the accuracy and clarity of our revised manuscript.

4.2 In the fourth paragraph, the lung cancer detection rates at initial screening from the NELSON and NLST study which mentioned by the author are inclusive of the full age range. In contrast, the data from this study used by the authors for comparison were for those aged 70 years or older. I believe that this should not be used for direct comparison and that lung cancer detection rates for the same age group should be recalculated. Please consider refreshing these data as much as possible.

Response:

Thank you for your comments. Unfortunately, we were unable to find specific data on lung cancer detection rates for individuals aged 70 years or older. While the NELSON study and NLST study include patients in the age range of 70-74 years, they do not provide separate data specifically for this age group. As a result, we regretfully could not perform a direct comparison with our study in terms of lung cancer detection rates. We acknowledge the limitations in this regard and have made note of it in our discussion

4.3 In the fifth paragraph, the diagnostic method (Table 5) is a statistically significant variable in the characteristics of lung cancer patients according to age. But the recommendations for clinicians discussed by the authors in this paragraph is very general, and I think it is an decision or consideration that is now taken for granted in

clinical practice. Or perhaps some new and more in-depth thinking could be done.

Response:

Thank you for your valuable comments. We have carefully considered your feedback and incorporated a discussion on the appropriate treatment in the late 70s compared to the early 70s group.

(see discussion page 379-395)

"Our data show promising results of LC treatment in the older population. In this study, 81% of the patients in the SDLC group received treatment appropriate to their cancer stage. Pham et al. showed that 73.7% of patients aged 70s with LC were presented for treatment discussion at a multidisciplinary meeting (15). Walter et al.(26) showed that 79.8% of the patients with LC aged 75–84 years received any tumor-directed treatment. A ganti et al. showed that 75% of patients with LC aged >65 years received any tumor-directed treatment (27). Despite concerns regarding treatment toxicity, shorter expected life expectancy, treatment intolerablity, and nihilism in an older population, the proportion of SDLC that received cancer treatment in our study was comparable to those in previous studies (15, 26, 27). Even though this could not guarantee survival gain from screening, at least, results show that a significant number of patients actively pursued aggressive treatment after screening.

In this study, 59% of the early 70s group underwent surgery, whereas only 20% of the late 70s group received surgical treatment. The lower rate of surgical interventions in the late 70s group suggests that the possibility of achieving complete recovery from LC may be lower in this group. However, in older patients, stereotactic

ablative radiation therapy has been reported to have comparable treatment efficacy to surgery for early-stage LC (28-30). Therefore, it would be inappropriate to conclude solely based on the lower surgical rate that older patients who underwent screening have lower treatment rates. “

4.4 The seventh and ninth paragraphs are more like restatements of the results section, which I personally believe should be accompanied by appropriate explanations, clinical implications, etc.

Response:

Thank you for your valuable comment, we decided to remove the paragraphs.

Minor problem

1. Apart from the content of the full text, there are still some places that need to be optimized. The overall scientific writing needs to be improved and it is highly recommended that a native English language editor be chosen for optimization.

Response:

Thank you for your comments. We acknowledge that the scientific writing quality of the initial manuscript was not up to the desired standard. We understand the importance of clear and accurate communication in scientific research, and we have taken your suggestion into consideration. We have engaged the services of a native English language editor to optimize the manuscript and ensure its readability and clarity.

2. Throughout, there are multiple errors in the use of punctuation, redundant or missing spaces, complete sentences missing full stops, inconsistent line spacing and other low-level errors. These details should be checked and eliminated prior to formal submission.

Response:

Thank you for your comments. Thank you for bringing these errors to our attention. We apologize for any inconsistencies or errors in punctuation, spacing, and other low-level details in the initial manuscript. We have taken your feedback seriously and have conducted a comprehensive review of the manuscript to eliminate these errors.

3. Some abbreviations, such as 95% confidence interval and relative risk ratio, do not need to be repeated with both the full spelling and abbreviation after the first occurrence.

Response:

Thank you for your comments. We deleted unnecessary repetition.

LDCT (methods, discussion page ___ line___)

LCS (discussion, highlightbox page ____ line ____)

SDLC (discussion)

NSDLC (discussion)

4. The X-axis heading in Figure 2a is partially obscured.

Response:

Thank you for your kind comments. We have taken immediate action to rectify this issue, and the revised version of Figure 2a with clear and visible X-axis heading

has been included in the manuscript.

5. eTable 1 has two titles.

Response:

We appreciate your feedback and have made the decision to remove eTable 1.