Peer Review File

Article information: https://dx.doi.org/10.21037/tlcr-23-256

Reviewer A

Comment 1: Although the authors collected data from a broad and representative population, they pooled together patients treated with surgery and radiotherapy. Even though surgery and radiotherapy (SBRT) have equivalent outcomes for stage I, the same cannot be said for stage II. Additionally, surgery is the preferred treatment for early-stage NSCLC, and patients that usually undergo radiotherapy instead of surgery have poor performance status, comorbidities, or any other characteristic that compromises overall survival. In such a way, these populations should be analyzed separately.

Reply 1: Thank you for your suggestion. Indeed a fair point is highlighted, which we already took into consideration in our method. We stratified our dataset for stage and type of therapy in the survival analysis as we agree on the non-equivalence in outcomes for stage II disease.

Other outcomes were explored for the logistic regression. We only stratified the logistic regression for stage as we wanted to evaluate the impact of several covariates, e.g. type of therapy on waiting times. For our descriptive statistics we used the pooled data to get a complete overview. We are aware that the stratification for type of therapy was not clearly enough described and therefore have made some changes to clarify this.

Changes in the text:

- P5 line 79: '*Analyses were stratified for stage and type of therapy*.'
- P8 line 130: 'The objective of this study was to: a) examine the association between extended time-to-treatment and overall survival *for stage and type of therapy* and b) identify factors associated with extended time-to-treatment.'
- P13 line 241: 'The association between extended time-to-treatment and OS was evaluated using multivariable Cox proportional hazards (PH) models per stage of early disease *and type of therapy*.'
- P13 line 251: '*Especially as treatment outcomes might not be equivalent for stage II compared to stage I.*'

Comment 2: Many parameters that may affect the results profoundly were not explored nor discussed, like the type of resection (R0 x R1 x Ru), treating institution (academic x non-academic), adjuvant systemic chemotherapy (yes x no), and histological grade. In this regard, patients whose treatment started after 33 days were more frequently submitted to EUS and EBUS, suggesting they had a larger tumor burden. Could the authors run these analyses or comment on that?

Reply 2: By using data from the Dutch cancer registry, we had access to high numbers of variables. Unfortunately, data on radicality of the resection as suggested and histological grade were not analysed because we only included variables that were expected to impact time-to-treatment. We still have tried to correct for overlapping information by stratifying for type of therapy and correcting for histology.

We discussed the role of the treating institution as a covariate in our analysis with the statistician of the Netherlands Comprehensive Cancer Organisation. As we had to deal with

potential clustering of patients within the several types of hospitals, we performed a multilevel regression analysis. In this analysis, we included hospital of diagnosis as a random-effect parameter. We focused on the hospital of diagnosis because of the diagnostic work-up. Also, hospital of diagnosis overlapped most of the time with hospital of treatment. Results showed a negligible clustering effect. Additionally, we stratified the survival analysis for type of hospital (general, teaching, academic) to further explore this. However, no differences were seen in outcomes and therefore we did not adjust for the type of hospital in further analysis. We added this to the methods, results and discussion.

We did not explore the role of adjuvant treatment as this was not within our research aim. Adjuvant treatment has been established as a standard for patients with completely resected NSCLC and we do believe that this might impact survival. However, from our point of view, this variable was less important as we focused on the diagnostic work-up in exploring the association between start of primary treatment and survival.

Last, results of the logistic regression analysis indeed showed significant higher odds ratios for diagnostic procedures as EUS, EBUS and transthoracic biopsy when it comes to longer waiting times. However, we explored the significance of these parameters in the association between time-to-treatment and survival with backward and forward stepwise regression (see description in methods page 13 line 246: 'Other covariates were selected by use of backward and forward stepwise regression, parameters with p-values >0.05 were eliminated from the model.'). None of these variables showed a significant impact on the association and therefore these were not included in the final analysis.

Changes in the text:

- P14 line 266: *'Analyses were also stratified for type of hospital (general, teaching, academic) to explore whether effects differ between these types of hospitals.'*
- P21 line 389: '*Stratification of the analyses for type of hospital showed that effects were similar, irrespective of type of hospital.*'
- P28 line 488: '*Next, information on parameters such as type of resection and histologic grade lacked. However, we only included variables that were expected to impact time-to-treatment. By stratifying data for type of therapy and correcting for histology, we have tried to caught some potential confounding. The role of adjuvant treatment was not explored because we believe that this parameter is of less relevance in the association between timing of primary treatment and survival.*'

Comment 3: It is not completely clear why patients submitted to mediastinoscopy were excluded. They should have been included in the analysis once confirmed that they did not have N2 disease.

Reply 3: We excluded these patients because the Dutch guidelines for the treatment of patients with lung cancer consider other timelines when receiving a mediastinoscopy (<49 days vs <35 days). Unfortunately, we only had information if a patient received mediastinoscopy yes or no. As the dataset did not allow us to include patients submitted to mediastinoscopy who did not have N2 disease, we excluded these patients on forehand. Sample sizes were too small to run multivariable analyses when stratifying the dataset for stage, type of therapy and mediastinoscopy.

Changes in the text:

- P9 line 170: ‘ Also, other timeliness are recommended by the Dutch guidelines when receiving a mediastinoscopy *compared to a normal diagnostic work-up* (<49 days *versus 35 days*).

Comment 4: Albeit the authors tried to explain why stage II patients that were pathologically downstaged to stage I had a worse overall survival following surgery delay, the explanation was not convincing enough, and the results are pretty counterintuitive. One would expect that stage II patients that turn out to be stage I to have improved outcomes. Is there any potential bias that might have contributed to these results, such as inadequate mediastinal lymphadenectomy?
Reply 4: There might indeed be potential bias that contributed to these results. Patients might have been misclassified due to incomplete mediastinal lymphadenectomy or because pathological evaluation in NSCLC does not include description of isolated tumor cells. We have made adjustments in the discussion section as requested.
Changes in the text:

- P21 line 398: ‘Potentially, this group with suspected radiological stage II disease, based on enlarged or FDG positive lymph nodes, is *misclassified as pathological stage I due to incomplete mediastinal lymfadenectomy or because pathological evaluation/staging in NSCLC does not include description of isolated tumor cells.*
- P22 line 406: ‘*Ideally, detection should be done rapidly to shorten the diagnostic interval and fasten treatment* [37].’

Minor comments: a) A native English speaker should revise the manuscript, b) The paragraphs are too long, which compromises engagement, c) I see no point in separating the discussion into different topics and, d) Figure legends can be improved. Add letters to point out what each graph represents.
Reply minor comments: Thank you for your valuable suggestions. We have modified our text as advised.
Changes in text:

- Minor comment a: Unfortunately we runned out of time and therefore the manuscript is not revised by a native English speaker. We were wondering if TLCR can offer some support with a native English speaker.
- Minor comment b: We have tried to improve the manuscript by adjustments in the method and discussion section.
- Minor comment c:
    - We removed the headlines in the discussion section on pages 21-30.
- Minor comment d:
    - P10 line 183: ‘*The flowchart presents the number of patients with non-small cell lung cancer in the Netherlands, 2014-2019, excluded with each exclusion criterium and the final number of patients eligible for analysis.*’
    - P16 line 303: ‘*The histograms present the distribution of stage I patients with non-small cell lung cancer in the Netherlands, 2014-2019, according to time-to-radiotherapy and time-to-surgery.*’
    - P16 line 307: ‘*The histograms present the distribution of stage II patients with*

non-small cell lung cancer in the Netherlands, 2014-2019, according to time-to-radiotherapy and time-to-surgery.'

- P17 line 337: 'The shading in these figures represent the confidence intervals. *A and B show the Kaplan Meier plots for earlier (Q1=blue) and latter (Q2-Q4=red) radiotherapy in stage I (left) and stage II (right). C and D show the Kaplan Meier plots for earlier (Q1=blue) and latter (Q2-Q4=red) surgery in stage I (left) and stage II (right.* The log rank test compares the survival distribution of the earlier *(Q1=blue)* and latter *(Q2-Q4=red)* treatment groups and show if there is a significant difference between the survival curves.'
- P18 line 345: '*This Kaplan Meier plot shows the difference in survival between the earlier (Q1) and latter (Q2-Q4) surgical treatment in clinical stage II patients that were pathologically staged I (A en B) and clinical stage II patients that were pathologically staged II (C and D).*'

Reviewer B

Comment 1: Abstract (background) and Introduction

The authors mentioned that the published data behind time to treatment and overall outcomes in early-stage lung cancer has showed contradictory results. This statement is not totally accurate and it could be misleading. The majority of well-done studies agree that timely therapy in early-stage lung cancer, particularly those receiving surgery, is associated with better outcomes (overall mortality, upstaging, and recurrence). The conflicting results and controversy relies on the ideal cut off time associated with worse outcomes (> 6 weeks vs. > 8 weeks vs. > 12 weeks). The main reason behind the conflicting results is the variability in the definition used for "time to treatment": time between first abnormal CT chest, time between diagnosis and treatment, time between first referral to specialist to treatment, etc. Most. the national guidelines recommend treating early-stage lung cancer as soon as possible.

Reply 1: Thank you for your valuable comment. It was not our intention to formulate misleading statements. In first instance we decided to use the term 'contradictory' as the study of Aggondowati et al. (17) reported that longer time intervals have a positive impact on survival. However, we rediscussed this topic with involved pulmonologists and see that this result might be due to patients that harbor specific conditions different form others. Therefore we decided to rephrase our statement in term of 'varied outcomes' as previous findings are not completely on the same line. Additionally, we rephrased parts of our discussion to highlight the main reasons for the conflicting results and discuss the results of Aggondowati et al. (17) separately.

Changes in text:

- P4 line 58: 'Previous research reported *varied outcomes on* the impact of time-to-treatment on survival for early-stage NSCLC.'
- P5 line 70: '*Varied outcomes on* the relation between time-to-treatment and survival in early-stage non-small cell lung cancer (NSCLC) patients *are reported*.
- P7 line 120: 'Previous research reported *varied outcomes on* the impact of time-to-treatment on survival for early-stage NSCLC (13). Some studies showed that *longer waiting times, e.g. time-to-surgery,* decreased survival (14-19), while subset analysis presented that shorter time-to-treatment was associated with a higher risk of death (17). *Varying definitions of time-to-treatment might (partly) explain these different findings.*

- P23 line 428: '*We also found a significant decrease of overall survival in patients who waited longer for surgery [14, 17, 40-42]. One study suggests that mortality risk differs meaningfully across patients subgroups when stratified by stage and specific histologic subtype (ie, adenocarcinoma and squamous cell carcinoma) [42]. We did not repeat this multivariable analysis for due to small subgroups but further research should incorporate this stratification. The other results* are not in line with previous studies that found no effect [14, 18, 43], or significant worse survival for shorter time intervals [17]. *We believe that the main reason for these conflicting results is the variability in definitions, e.g. time-to-treatment and cutoff values for extended time-to-treatment. But also,* stratification for therapy and stage, the use of either clinical or pathological tumor *stage*, and the prognostic factors selected for multivariable analysis such as performance status and comorbidity. *Previous timeframes that have been evaluated vary from dichotomous to categorical, e.g. 0-7days, 8-14 days, 15-60 days, >61 days, but also <42 days vs ≥42 days, or 1-37 days vs >37 days (14,18,40).* We have chosen to use the time between first outpatient visit to start of treatment as a measure of time-to-treatment because it also includes any delay due to the diagnostic process. A sensitivity analysis investigating the effect of using a different definition of time-to-treatment did not change our conclusion, i.e. time from diagnosis to start of treatment gave similar results to time of first consultation at the pulmonologist to start treatment. Others used the time between diagnosis and treatment, leading to shorter time intervals, which can even become 0 days if the diagnosis is made during surgery. *Shorter time intervals are less likely to show any significant findings as chances on tumor progression are reduced.*

Comment 2: Reference 17, mentioned as example for negative association between time to treatment and overall outcomes in NSCLC analyzed time to treatment without adjusting for other confounder variables (multivariable analysis) make it prone to bias.

Reply 2: In the manuscript of Aggondowati et al. (17) it is mentioned that Kaplan Meier models were adjusted for covariates. See legenda Table 2 and Table 3:

- *'*, model was adjusted for age, sex, race, urban/rural, distance to hospital, facilty type, primary payer, Charlson/Deyo comorbidity score, histologic type, and treatment modalities.'*
- *'*, model was adjusted for age, sex, race, urban/rural, disdtance to hospital, facility type, primary payer, Charlson/Deyo comorbidity score, and histology type. Excluding patients who died within 1 months after diagnosis.'*

Also in the section 'outcome and predictors' they describe for which factors they corrected and where detailed information can be found: 'The covariates adjusted in the analysis were age at diagnosis, sex, race, urban/rural status, distance to the reporting hospital, primary payer, facility type, stage at diagnosis, histology, treatment type, and Charlson-Deyo comorbidity score because those variables can affect both decision on timing of treatment and patient survival. Detailed information about how variables in the NCDB are defined by the American College of Surgeons is provided elsewhere (19).' For this reason we believe that these models not prone to bias. Therefore we did not make any changes in the manuscript based on this reference. In case we misunderstand reviewer 2, we might be more than happy to reconsider this.

Comment 3: Methods - 1. Patient selection

Selecting patients only referred to pulmonologist creates selection bias. In reality, not all patients get to diagnosis through this pathways. Some patients are referred directly to thoracic surgery, or biopsy. Excluding patients that look for second opinion, or visited multiples hospitals also creates more selection bias. Finally, excluding patients with mediastinoscopy prior to surgery self select a population with shorter time to treatment. All these selection bias probably explain why the time to treatment period was so short in this study (average 47 days). This issue can be addressed by including all comers with early-stage lung cancer, and see if that changes the results.

Reply 3: Thank you for your fair point. Indeed we agree that we introduced some selection bias within our approach. However, in the conceptual phase, we thoroughly discussed our method with our team, involving statisticians from the Netherlands Comprehensive Cancer Organisation / Netherlands cancer registry and pulmonologists. We have decided to focus on outpatients and excluded patients who received a mediastinoscopy for several main reasons which we will clarify within this reply. As the Dutch national guidelines recommend a median time-to-treatment of 35 days from first clinical presentation at the general practitioner to start of treatment, we believe an average of 47 days is reasonable. Especially as the logistics in Dutch healthcare are well organized and all inhabitants have equal access to health care. We excluded patients that received a mediastinoscopy because for these patients other timeliness are recommended by the Dutch guidelines(<49 days vs <35 days). Unfortunately, there was only information available if a patient received of did not received a mediastinoscopy. Stratifying the dataset for stage, type of therapy and mediastinoscopy resulted in too small samples sizes to run multivariate analyses.

Further, we are aware that we reduced our sample size as we focused on a more 'routine' referral pathway according to the guidelines. Overall, patients with lung cancer can be considered inpatient and outpatient. Outpatients are most of the time referred by a general practitioner and do have a 'regular' diagnostic work-up. Inpatients enter the clinical pathway differently, e.g. emergency department, and have a wide range of underlying prognostic factors that harbor specific conditions. These may influence the association between time-to-treatment and survival. It is difficult to correct for all these underlying confounding factors. Therefore we focused on outpatients, comprising 55% of our study population, as we wanted to approach our dataset as homogenous as possible to reduce potential confounding. In case we would had found a significant effect, over all newly diagnosed patients with early-stage disease, it would be up for discussion if this was due to longer or shorter time intervals or due to underlying prognostic factors we did not adjust for. From our perspective this reason outweighs the selection bias. We have tried to elaborate a bit more on this reasoning in the limitations where we already highlight this shortcoming.

Changes in text:

- P10 line 170: 'Also, other timeliness are recommended by the Dutch guidelines when receiving a mediastinoscopy *compared to a normal diagnostic work-up* (<49 days *versus 35 days*).
- P26 line 501: '*It is difficult to correct for all these underlying factors, especially as most are not registered..*'

Comment 4: Methods - 2. Time to treatment definition

I would perform a sensitivity analysis with a different definition of time to treatment like "diagnosis to treatment" and see if the results remain the same.

Reply 4: We have done a sensitivity analysis which led to comparable results, please see tables below (multivariable Cox proportional hazard model, which is added to the manuscript as supplementary material). The reason that we have chosen to use the time between first outpatient visit to start of treatment as a measure of time-to-treatment was because it also includes any delay due to the diagnostic process. This offered us insight in factors that are at risk to cause delay in waiting times.

Clinical stage I	**Clinical stage I RT (N=3989)**		**Clinical stage I Surgery (N=3438)**	
Time-to-treatment quartiles (days)	aHR (95% CI)	p-value	aHR[2] (95% CI)	p-value
Q1 (0-33 d)	Ref	-	Ref	-
Q2 (34-47 d)	1.03 (0.96-1.09)	0.4006	0.99 (0.88-1.13)	0.9586
Q3 (48-64 d)	1.01 (0.95-1.06)	0. 8516	0.97 (0.86-1.10)	0.6780
Q4 (64-180 d)	1.04 (0.99-1.08)	0.0617	0.91 (0.79-1.04)	0.1851

Clinical stage II	**Clinical stage II RT (N=747)**		**Clinical stage II Surgery (N=1325)**	
Time-to-treatment quartiles (days)	aHR[1] (95% CI)	p-value	aHR[2] (95% CI)	p-value
Q1 (0-33 d)	Ref	-	Ref	-
Q2 (34-47 d)	0.90 (0.80-1.02)	0.0977	1.23 (1.09-1.37)	0.0005
Q3 (48-64 d)	0.92 (0.83-1.00)	0.0614	1.14 (1.01-1.28)	0.0373
Q4 (64-180 d)	0.95 (0.88-1.03)	0.1986	1.22 (1.09-1.37)	0.0009

Changes in text:

- P24 line 444: '*A sensitivity analysis investigating the effect of using a different definition of time-to-treatment did not change our conclusion.*'

Comment 5: Methods - 3. Covariates

Even though patients include performance status as a co-variable, they didn't include a comorbidity index, chronic lung disease, and tobacco use in the covariate analysis. These factors have shown to influence significantly overall survival and timeliness to therapy in patients with lung cancer. The comorbidity index in the study population must have been very high because only 46% of patient with stage I received surgery (the preferred treatment for these patients). In addition, it is not clear if all the 109 hospitals were tertiary hospitals or small community centers. Studies have demonstrated better outcomes in patient managed in specialized cancer centers

Reply 5: We completely agree on this shortcoming. Unfortunately, comorbidity index, chronic lung disease and tobacco use were not registered on a national level by the Netherlands cancer registry. We strongly believe that comorbidities influence time-to-treatment as specific

conditions may take more or less time for diagnostic work-up, decision-making and initiation of treatment. However, we were not able to correct for these factors and tried to elaborate on this shortcoming in our limitations.

Regarding the comment that the comorbidity index in the study population must have been very high because only 46% of patient with stage I received surgery (the preferred treatment for these patients), we would like to mention that in contrast to some other countries, the Netherlands Cancer Registry also has information on the patients with clinical diagnosis (i.e. without pathologically confirmed lung cancer). Especially among the patients with SBRT, the proportion with a clinical diagnosis is high, leading to a relatively low proportion of patients receiving surgery.

Further, we understand that it is not clear which types of hospitals are included. The 109 hospitals enclosed all Dutch hospitals between 2014 and 2019 and included general, teaching and academic hospitals. We explored potential clustering of patients within the several types of hospitals. Therefore, we have done a multilevel regression analysis. In this analysis, we included hospital of diagnosis as a random-effect parameter. We focused on the hospital of diagnosis, which most of the time overlapped with hospital of treatment. Results showed a negligible clustering effect. Additionally, we stratified the survival analysis for type of hospital (general, teaching, academic) to further explore this. However, no differences were seen in outcomes and therefore we did not corrected for this variable in further analysis. Based on your comment, we have added the types of hospitals in the methodology section to be more clear.

Changes in text:

- Page 9 line 154: 'From 2014 to 2019, the NCR recorded 23,428 patients diagnosed with clinical stage I or II lung cancer from 109 hospitals *(general, teaching and academic hospitals)* in The Netherlands.'
- Page 25 line 479: 'First, *detailed information on comorbidity index, chronic lung disease, tobacco use and lung function test are not available in the NCR, limiting the possibility to explore the impact of these variables.*'

Comment 6: 5. Stage of the disease

There is a significant limitation on the study findings by using clinical stage. In this study cohort there was up to 43% discordancy between clinical and pathological stage II disease (20% downstaged, and 23% upstaged). I would perform the analysis with pathological stage for all whom underwent surgery (all stage I, all stage II, and all stage I or II).

Reply 6: Thank you for your suggestion. The specific reason that we have chosen for cTNM rather than pTNNM, is because decisions on diagnostics and therapeutics, made during multidisciplinary team meetings, are based on cTNM. Timing of treatment is therefore impacted by cTNM and not pTNM as the pathological status is a result from treatment. However, it is common to use pTNM in survival analyses. In our study weperformed sensitivity analyses for pTNM subgroups in clinical stage II patients that underwent surgery. From our perspective, performing stratification of datasets in this order gives a close presentation of current clinical pathways. As you suggested, we also performed sensitivity analysis for pathological staging (all stage I, all stage II) of patients who underwent surgery. For this analysis we did not stratify for clinical stage beforehand. See results in table below. Results of pathological stage I approaches significance. This overlaps with results described in the manuscript as clinical stage

II patients that downgrade to pathological stage I show significant results.

	Stage T1a-T1b radiotherapy (N=1228)		**Stage T1a-T1b surgery (N=905)**		**Stage T1c radiotherapy (N=550)**		**Stage T1c surgery (N=483)**	
Time-to-treatment quartiles (days)	aHR[1] (95% CI)	p-value	aHR[2] (95% CI)	p-value	aHR[3] (95% CI)	p-value	aHR[4] (95% CI)	p-value
Q1 (0-33 d)	Ref	-	Ref	-	Ref	-	Ref	-
Q2 (34-47 d)	1.07 (0.74-1.55)	0.7182	0.96 (0.48-1.89)	0.9065	0.78 (0.44-1.39)	0.3960	1.72 (0.78-3.83)	0.1814
Q3 (48-64 d)	1.07 (0.70-1.62)	0.7623	1.41 (0.68-2.94)	0.3546	1.33 (0.76-2.32)	0.3253	1.21 (0.49-3.04)	0.6783
Q4 (64-180 d)	1.04 (0.72-1.49)	0.8455	0.97 (0.48-1.97)	0.9252	0.91 (0.54-1.52)	0.7059	**3.11 (1.38-7.01)**	**0.0061**

Clinical stage II	**Pathological stage I (N=3170)**		**Pathological stage II (N=1410)**	
Time-to-treatment quartiles (days)	aHR[1] (95% CI)	p-value	aHR[2] (95% CI)	p-value
Q1 (0-33 d)	Ref	-	Ref	-
Q2 (34-47 d)	1.25 (0.99-1.57)	0.0545	1.22 (0.86-1.72)	0.266
Q3 (48-64 d)	1.26 (0.98-1.61)	0.0691	1.24 (0.83-1.85)	0.305
Q4 (64-180 d)	1.04 (0.79-1.37)	0.0890	1.32 (0.86-2.03)	0.201

Comment 7: 6. Figures
I don't see the benefit of showing the kaplan Meier curves by type of treatment. Its expected the survival for people treated with radiation (not candidates for surgery) would be lower compared to those who were candidates for surgery
Reply 7: We included these figures in the manuscript as we thought it would be easier to interpret the results when supported by Kaplan Meier curves. However, we agree that these figures may be a little obvious. Therefore we removed the first two plots in Figure 4.
Changes in text:

- Please see Figure 4

Reviewer C
Comment 1: Is it ideal to have the analysis center on stage but also have the definition of stage change mid-way through the collection of data? I might argue that it would be preferable to focus on the size of the tumor and/or nodal definitions, irrespective of the AJCC edition. As it is, many patients who were stage IB in 7th edition were now stage II in 8th edition, but your analysis speaks of stage I vs. II as very distinct populations.
Reply 1: Thank you for your valuable comment. We had the same thoughts during the conceptual phase of our study. Therefore, we explored the same analysis on tumor size for stage T1a-T1b and T1c with a main focus on the last because this stage was added in the 8th edition. Please find in the table below our findings. A significant effect was found for time-to-treatment >65 days in surgical T1c patients. However, the relatively large CI suggests that the sample size of this subgroup is too small for a multivariate Cox proportional hazards model (n-483). Therefore we have chosen to stay with stage I and II in the first place. In our final Cox proportional hazard models, we have corrected for stage shift during time as we adjusted for incidence years.

Comment 2: Surgery is the standard of care for early stage NSCLC, and yet 54% of patients with stage I NSCLC underwent radiation rather than surgery. The fact that it is only 36% for stage II suggests that the difference isn't medical operability -- there shouldn't be a marked difference in resectability between stages, so this suggests that an inordinately large proportion of patients with stage I NSCLC received less than optimal treatment. Moreover, the discussion attends to the potential for upstaging and accuracy of clinical staging, but this cannot be assessed for the patients who underwent radiation rather than surgery. At the very least, this is a major limitation in interpreting the data. It arguably undermines a lot of the interpretation of the data here.

Reply 2: Fair point. However, the question is 'what is optimal treatment?' as this differs per patient, including patient preferences and individual characteristics. Indeed a large proportion of patients with stage I received less optimal treatment. However, the use of radiotherapy became more prominent over time in stage I NSCLC in the Netherlands over the past years. We woud like to refer to the manuscript of Evers et al. (Trends and variations in treatment of stage I-III non-small cell lung cancer from 2008 to 2018: A nationwide population-based study from the Netherlands). They describe that in stage I NSCLC, the rate of surgery decreased from 58% (2008) to 40% (2018) while radiotherapy use increased over time (from 31% to 52%), which mostly concerned stereotactic body radiotherapy. In stage II, 54% of patients received surgery, and use of radiotherapy alone increased form 18% to 25%. The likelihood of receiving radiotherapy instead of surgery was lower for patients with a 15-44 min drinving time to a radiotherapy facility, compared to patients with less than 15 min driving time. Also, we have made use of real world data which differs from studies that used powercalculations and strict inclusions criteria. Using real world data may lead to larger differences in proportions than expected in the first place.

Further, indeed our discussion on the potential for upstaging cannot be assessed for ptaients who underwent radiation rather than surgery. However, we did not aimed to perform subset analysis for radiation therapy as we did not find any significant results for these patients. This makes timing of treatment less urgent. We have tried to engage our findings to surgical patients. With following adjustment we hope to clarify this.

Changes in text:

- P23 line 425: '*We did not further assessed the potential for upstaging for patients with a pathological status who underwent radiation rather than surgery.*'

Comment 3: In the discussion, it would be helpful for the background to include some benchmark numbers for typical time to treatment in other studies.

Reply 3: Thankyou for your suggestion. We were not aware that we did not report any other benchmarks at all. Therefore we have made some modifications as suggested.

Changes in text:

- Page 23 line 439: '*Previous timeframes that have been evaluated vary from dichotomous to categorical, e.g. 0-7days, 8-14 days, 15-60 days, >61 days, but also <42 days vs ≥42 days, or 1-37 days vs >37 days (14,18,40).*'

Comment 4: Also, while there is some discussion of fact that the first visit with a pulmonologist was used, that seems like a rather idiosyncratic choice. In many health care

systems, a surgeon would initiate the workup, and in fact there was a very large attrition of patients from the cohort because they were not first seen by a pulmonologist. In addition, the interval of the workup could arguably start when an imaging-based abnormality was first identified. If it took weeks to be seen by a pulmonologist, why is that interval not clinically relevant? That may not be a concern in the Netherlands, but if practice patterns are very different across different health care systems, I would be concerned about the generalizability of the observations.

Reply 4: Thank you very much for your critical point of view. We understand that our choice to focus on the first visit with a pulmonologist might seem idiosyncratic for other healthcare systems. However, the routine work-up in the Netherlands is that patients suspected for lung cancer are referred by a general practitioner to a pulmonologist in secondary care. Contrary to other countries, Dutch lung cancer patients are primarily seen by pulmonologists as these manage the lung cancer pathways. Still, at the starting of this study, we discussed this topic thoroughly with our multidisciplinary team involved in lung cancer care. Ideally, we also wanted to include the diagnostic work-up before seeing a pulmonologist. We agree that this interval is clinically very relevant. However, this starting point in primary care is not registered by the Netherlands cancer registry. Therefore, we have chosen for a different method. Based on our discussions, we concluded that on national level there are two major groups of lung cancer patients: inpatient and outpatient. Outpatients follow the routine workup in secondary care, mostly starting with a first visit with a pulmonologist. Inpatients enter the clinical pathway differently because they harbor specific conditions different from others, e.g. emergency department because of severe symptoms. Inpatients have a broader range of underlying prognostic factors that may influence the association between time-to-treatment and survival. We are not able to correct for all these underlying confounding factors as most are not registered by the Netherlands cancer registry. We have tried to approach the dataset as homogeneous as possible and to give a close representation of current Dutch lung cancer pathways. Therefore we have chosen to focus on outpatients with a first visit with a pulmonologist. Still, practice patterns might be different from other countries. As our findings were obtained in a country with universal public healthcare, these can only be generalized to countries with a similar healthcare status. We recognize this as a shortcoming of our study and elabarated on this in the limitations.

Changes in text:

- Page 26 line 485: '*As well as information on time intervals before visiting a pulmonologist. We believe that the time interval from primary care to secondary care is clinically very relevant. Availability of such data would have put us in a better position to explain variations in time-to-treatment.*'
- Page 26 line 502: '*Also, Dutch practice patterns might differ from other countries as lung cancer patients are primarily seen by pulmonologists. As our findings were obtained in a country with a healthcare system equally accessible to all inhabitants, these can only be generalized to countries with a similar healthcare system.*'

Comment 5: Neoadjuvant chemoimmunotherapy has now become a standard of care for at least a subset of patients with stage II NSCLC. Moreover, there is also a place for at least limited biomarker testing to look for EGFR and ALK, arguably PD-L1 as well, in the initial

workup. This is likely to change and further slow the workup, making these results less relevant to current and future care. There is a passing mention of this around line 498, but I would submit that it should receive far more discussion and be covered earlier. This is extremely important for the implications of this work.

Reply 5: Thank you for highlighting this very relevant topic in early stage disease. We agree that further elaboration of this should be done earlier in the discussion section. We have reviewed recent literature and tried to discuss the potential role of neoadjuvant systemic therapy on time-to-treatment and survival in early stage NSCLC.

Changes text:

- P22 line 402: 'We suspect that some of these pathologic stage I cancers may actually be pathological stage II with lymph node micrometastasis not detected in the resected material because of current clinical methods or distant micrometastasis. *Lymph node micrometastasis are usually found in clinical N0/N1 NSCLC and with tumors smaller than 3 cm [30], which is associated with a poor prognosis [31-36].* Ideally, detection should be done rapidly to shorten the diagnostic interval en fasten treatment [37,38]. This hypothesis is supported by the fact that the survival curves of patients with clinical stage II/pathological stage I who were treated within 33 days was better than that of those patients who were treated later. Their survival curves became more similar to those of clinical stage II/pathological stage II cancers. *Early detection of occult disease could help to identify those patients who might benefit from neoadjuvant therapy. Results of a previous study indicate a place for limited biomarker testing (EGFR, ALK and PD-L1) in the initial workup of early-stage NSCLC [38]. During the study period, however, biomarker testing in early stages was only performed in clinical trials.'*

Comment 6: I think there may be a missing word in the beginning of the sentence starting in line 439. In any event, I don't understand the sentence.

Reply 6: Fair point. We adjusted the sentence.

Changes in text:

- Page 23 line 435: '*We believe that the main reason for these conflicting results* is the variability in definitions, e.g. time-to-treatment and cutoff values for extended time-to-treatment. *But also,* stratification for therapy and stage, the use of either clinical or pathological tumor stage, and the prognostic factors selected for multivariable analysis such as performance status and comorbidity.'

We would like to thank the reviewers again for their valuable suggestions. If you have any further additions or questions, please feel free to contact us.