Peer Review File

Article Information: https://dx.doi.org/10.21037/tlcr-23-291

Reviewer A

The review is intended to be comprehensive. However unfortunately this does not seem to be the case. The review covers certain publications and not others. There are categories of markers that have been investigated some with substantial extent of validation. The authors should recognize that the ultimate evidence of mortality benefit is hard to come by and will take a long time. Nevertheless some markers with sufficient validation are available. Thus the authors/ assertion that the level of evidence supporting clinical efficacy is not sufficient to be translated to clinical practice seems too radical. The authors are encouraged to be more extensive in their review of biomarkers and to rank biomarkers according to the extent of their validation which would provide some objectivity into the state of biomarkers for lung cancer screening.

We thank the reviewer for their time and expertise and the opportunity to improve this manuscript. Our responses to specific comments are detailed in the table below (all page and line numbers were recorded with 'Show Markup' on):

No.	Comment	Answer	Change in text
1	The review covers certain publications and not others	In this review we targeted articles that were most progressed in terms of their impact assessment and aimed to highlight the most clinically efficacious biomarkers of lung cancer in updated literature with applications in lung cancer screening and in never-smokers. If there were any articles in particular that were inadvertently missed, we will gladly include them assuming inclusion criteria are met and/or they are relevant to other	N/A

		identified markers.	
2	Recognize that the ultimate evidence of mortality benefit is hard to come by and will take a long time. Nevertheless some markers with sufficient validation are available. The authors assertion that the level of evidence supporting clinical efficacy is not sufficient to be translated to clinical practice seems too radical.	When evaluating the level of evidence supporting each of the biomarkers discussed in this review, we employed the NHMRC Evidence Hierarchy which indicates the highest level of supporting evidence for a study is a systematic review of a number of randomized control trials. We believe that, while certain studies have extensive bodies of supporting evidence, this does not necessarily make them fit for clinical translation, as they are lacking higher level validation (as per NHMRC Evidence Hierarchy and Early Development Research Network Five Phase Approach) and further research is required before they can be translated into clinical practice. We do however recognise that this level of validation takes long time and as such have included a brief paragraph acknowledging this in the conclusion.	Please see "Conclusion" section page 20 line 490 onwards. Text: Several markers in various stages of development are currently available for LCS and LCINS, and further advancement in term of external validation and impact assessment is in progress. Randomised trials are considered to be the gold standard for external validation (160). However, proving ultimate evidence of mortality benefits is challenging and may take a significant amount of time. As a result, more time- and cost-effective models are increasingly being used to complement clinical decision-making with the aim of improving patient outcomes (161, 162). Such models have already been implemented to compared the effectiveness of certain biomarkers in LCS (163). To conclude, significant advancements have been made in the field of lung cancer biomarker research, with numerous biomarkers of lung cancer displaying varying levels of clinical efficacy and showing improvement in diagnostic accuracy over standard clinical workflow in LCS and LCINS. The priority now should be the validation of existing candidate markers in appropriate clinical contexts to integrate these into clinical practice. To do this, randomized controlled trials or similar methods of validation should be designed to test the efficacy of these biomarkers, This will positively impact lung cancer diagnosis and treatment, and help to reduced lung cancer mortality worldwide.

3	Be more extensive in their review of biomarkers	The articles included in this review were subject to screening detailed in the methods section, and articles that did not satisfy our inclusion criteria were excluded. If there are specific articles that we have inadvertently omitted, we would gratefully receive any recommendations of articles that meet our inclusion criteria for inclusion in this manuscript.	N/A
4	Rank biomarkers according to the extent of their validation to provide some objectivity	We agree an objective ranking system should be employed to assess articles included in this review. Using the Early Detection Research Networks's five-phase approach we have ranked each of the studies referenced and included their ranking in Table 2 (as per Reviewer comment 15).	Please see Table 2 in pages 10-11

	Biomarker Source	Advantages	Disadvantages	Development Phase	Reference
Nucleic Acid Markers					
DNA Markers					
Microsatellite instability/ loss of heterozygosity and plasma DNA	Sputum/Plasma	High sensitivity, minimally invasive	Moderate specificity	Phase 3	(37)
DNA methylation	Sputum	High sensitivity, minimally invasive	Low sensitivity	Phase 3	(38)
microRNA Markers					
microRNA-155	Plasma	High sensitivity and specificity, minimally invasive, detects non-small cell lung cancer (NSCLC) at early-stage	No large-scale validation	Phase 3	(39-43)
Multiple microRNA panels	Plasma/Serum	High sensitivity, stable in serum in plasma, specific to lung cancer, minimally invasive	No large-scale validation	Phase 4	(44-47)
Protein Markers					
Osteopontin	Plasma	Can be used for both risk and nodule stratification, minimally invasive	Not unique to lung cancer, evidence is still in early stage (control-case)	Phase 2	(48)
Carcinoembryonic antigen	Serum	Minimally invasive, high sensitivity	Low sensitivity	Phase 2	(49-51)
Combination proteins	Serum	Can be used for both risk and nodule stratification, cost-effective, minimally invasive	Overlap with other cancers and inflammatory diseases	Phase 2	(52, 53)
Complement fragments	Plasma	Minimally invasive	Discrepancies in results between studies, quantification of C4d is not standardised	Phase 3	(54-56)
T-Cell receptors	Peripheral blood	Stable in pathological subtypes, minimally invasive	Only studied in stage I lung cancer, limited to Asian population	Phase 3	(57)
Autoantibodies	Peripheral blood	Panel compensate for cancer heterogeneity, assessed in randomized control trial, minimally invasive	Moderate sensitivity	Phase 5	(58, 59)
Metabolites	Serum	Panel has high sensitivity and specificity, minimally invasive	Biomarkers are still in exploratory stage	Phase 4	(60-62)
Circulating lipids	Serum	High sensitivity and specificity, minimally invasive, detects NSCLC at early-stage	Confounded by systemic pathology, validated in female only cohort	Phase 3	(63)
Cellular Markers					
Circulating tumour cells	Bloodstream	High detection rate by CellCollector, minimally invasive	Confounded by systemic pathology, non-specific	Phase 4	(64, 65)
Circulating inflammatory cells	Bloodstream	Minimally invasive	Confounded by systemic pathology, non-specific	Phase 3	(66, 67)

Table 2: Summary of advantages and disadvantages of biomarkers for LCS and in never-smokers. The NIH Early Detection Research Network Five-Phase Approach was used to rank biomarker development.

Reviewer B

Overall, this a very-well written article that presents an important summary of the literature on biomarkers for early detection of lung cancer. Given the literature search yielded thousands of articles but fewer than 100 are included, it would be helpful to have a flow chart showing which studies were removed. As I point out below, measurements of circulating leukocyte DNA methylation in lung cancer cases (vs controls) have also been explored for early detection of lung cancer but these studies are not mentioned in this article.

The section "Nucleic acid markers of lung cancer" mentions "liquid biopsy", RNA and cfDNA markers in blood. It isn't initially clear that the DNA section after it is a follow-up of that section and that the prior section was an introduction; the flow of the sections could be improved.

Page 5, line 187. Smoking has been clearly shown to alter DNA methylation levels (so not "probably" but definitely). Moreover, other environmental exposures and medical conditions have also been associated with DNA methylation levels in blood, including air pollution, BMI, diabetes, inflammation, etc.

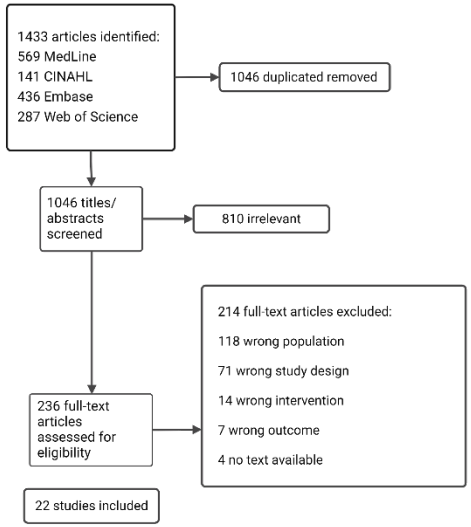
The DNA methylation section should include a discussion of studies that have used leukocyte DNA methylation for prediction of lung cancer. CpGs in circulating leukocytes can provide accurate measurements of smoking history (e.g., pack-year methylation scores have been developed) and DNA methylation levels at other CpGs may reflect other environmental exposures. Changes in leukocyte DNA methylation (i.e., not from circulating tumor cells) may be used as predictors of lung cancer risk and mortality and could be used to improve risk stratification for early detection of lung cancer.

Page 6, line 231. Typo for "markets" should be markers.

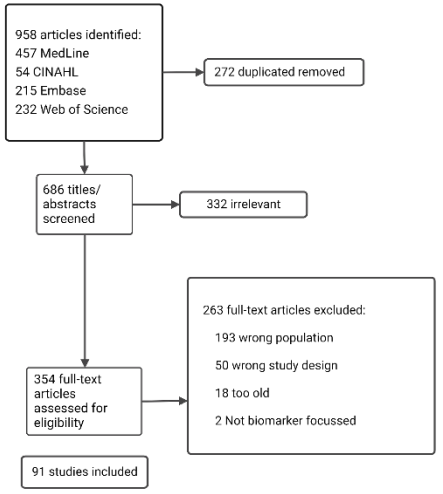
We thank the reviewer for their time and expertise and the opportunity to improve this manuscript. Our responses to specific comments are detailed in the table below (all page and line numbers were recorded with 'Show Markup' on):

No.	Comment	Answer	Change in text

5	Include a flow chart showing which studies were removed	We agree a flow chart will improve clarity around study exclusion. Flow charts for both literature searches have been included as supplementary figures.	Please see Supplementary Figures 1 and 2. Lung cancer screening (Supp Fig 1):



Lung cancer in never-smokers (Supp Fig 2):



6	Improve the flow of "Nucleic acid markers of lung cancer" section for leading in to subsequent paragraphs	The "Nucleic acid markers of lung cancer" section has been reworded to improve clarity.	Please see "Nucleic acid markers of lung cancer" section on page 12 line 197 onwards. Text: Gradual accumulation of genetic and epigenetic changes in the cell nucleus can be used to detect lung cancer formation, progression and metastasis . While the disease is primarily driven by somatic alterations, typically linked to smoking exposure, germline mutations could also predispose individuals to lung cancer development (69, 70). Emerging as promising biomarkers, nucleic acid markers for lung cancer such as cell-free DNA (cfDNA) and circulating RNA are significantly advancing lung cancer diagnosis through the immense potential of liquid biopsy detection methods. The advent of real-time polymerase chain reaction (PCR) and next-generation sequencing has enhanced the sensitivity and specificity of circulating nucleic acid analysis, making it a valuable asset in the arsenal of lung cancer detection methods (73, 74).
7	Page 5, line 187. Smoking has been clearly shown to alter DNA methylation levels (so not "probably" but definitely)	We acknowledge our statement is inaccurate and have amended it accordingly.	Please see "Nucleic acid markers of lung cancer" section on page 12 line 199. *Please see reviewer comment **6**.*
8	Mention other environmental exposures and medical conditions have also been associated with DNA methylation levels in blood	We agree this is an important factor to include and have added a paragraph describing the confounding effect other drivers of methylation may have on markers present in blood.	Please see "Methylation" section page 14 line 260-271 Text: DNA methylation has been linked to environmental exposures and comorbidities other than lung cancer, adding

			a layer of complexity to its potential use as a biomarker for lung cancer detection. Research has found correlations between methylation patterns and exposure to traffic-related pollutants, polycyclic aromatic hydrocarbons, and particulate matter rich in metals, affecting genes associated with immune responses and other processes (89-91). Moreover, methylation has been tied to medical conditions such as osteoporosis, obesity and chronic obstructive pulmonary disease (COPD), with studies showing distinct methylation profiles in individuals with these diseases compared to healthy controls (92-94). Obesity, in particular, has been associated with alterations in DNA methylation, influencing the likelihood of developing diseases like type 2 diabetes (93). Factors such as the intrauterine environment, physical activity, and diet can also impact both obesity and DNA methylation (95). Therefore, when investigating methylation as a potential biomarker for lung cancer, it is crucial to consider these additional influences.
9	Mention circulating leukocyte DNA methylation as biomarker for lung cancer in the DNA methylation section	The studies on circulating leukocyte DNA methylation we have identified fall outside our inclusion criteria (not performed in cohorts undergoing CT screening). If there was a specific article the reviewer wanted to see included we would gladly accept any recommendations.	N/A
10	Page 6, line 231. Typo for “markets” should be markers.	This mistake has been corrected.	Please see page 13 line 247.

			Text: Among epigenetic changes, DNA hypomethylation and hypermethylation of specific 5'-C-phosphate-G-3 (CpG)-rich regions in the promoter region of tumour suppressor genes are early events in carcinogenesis, making them ***markers*** of interest for early lung cancer detection.

Reviewer C

The paper reports the current study of the biomarkers of lung cancer in never smokers. The reviewer thinks this article is very interesting for lung cancer research, however, it needs to clear and complete in this article. To render the manuscript suitable for publication to Translational Lung Cancer Research, several corrections should be made before the paper should be accepted.
Specific Comments:
-The authors are requested to briefly more discuss indoor radon exposure-indued lung cancer in never smokers. It would be helpful if the authors give example or scenario to support its description. Clarification of this point in text is needed.
- In this research, please add more detail about the biomarkers of lung cancer from indoor radon exposure in never smokers in the manuscript.

We thank the reviewer for their time and expertise and the opportunity to improve this manuscript. Our responses to specific comments are detailed in the table below (all page and line numbers were recorded with 'Show Markup' on):

No.	Comment	Answer	Change in text
11	Briefly more discuss indoor radon exposure-indued lung cancer in never smokers, give example.	We agree referencing indoor radon exposure will improve clarity. When discussing LCINS in the introduction, an additional sentence has been added describing radon induced lung cancer.	Please see "Biomarkers of lung cancer in never-smokers" section page 8 line 150-153 onwards. Text: Radon exposure is considered the leading cause of LCINS – and the second leading cause of lung cancer overall – with individuals becoming exposed to high levels of radon when living or working in buildings with poor ventilation in areas of high environmental radon (29, 30).
12	Add more detail about the biomarkers of lung cancer from	We agree more information should be provided for this. More detail regarding	Please see "Carcinoembryonic antigen and combination markers" section page 16 line 328

	indoor radon exposure in never smokers	biomarkers of radon induced LCINS has been supplied.	onwards. Text: A South-East Asian study investigating the relationship between serum biomarkers and residential radon levels in never and former (>15 years) smokers, described a significant increase in serum CEA and cytokeratin 19 fragment (CYFRA21-1) in individuals with lung cancer, compared to heathy controls with high and low radon exposure. Interestingly, an increase in CEA (p=0.009) and CYFRA21-1 (p=0.0031) was also observed in healthy controls with high radon exposure when compared to low, potentially indicating high serum CEA as a biomarker for lung cancer development in never-smokers. Receiver operating characteristic analyses of CEA and CYFRA21-1 for diagnosing lung cancer illustrate high specificity (98% and 94% respectively) but inadequate sensitivity (57.3% and 58.6 respectively), which similarly has been reported in other studies investigating CEA as a biomarker of NSCLC and mutational status .

Reviewer D

This manuscript is a narrative review highlighting current developments in the detection and diagnosis of lung cancer in screening and non-smoker populations. The manuscript covers a timely and clinically significant topic and, overall, does a good job of providing a balanced review of state-of-the-art developments.

The manuscript also has some noteworthy weaknesses that, if addressed, would strengthen the value of the study.

First, studies not covering body-fluid biomarkers for LCS or circulating biomarkers in never-smokers were excluded. Given the broad title of the manuscript ("Biomarkers of lung cancer for screening and in never-smokers"), it is unclear why these articles (e.g., biomarkers based on nasal epithelial specimens) are omitted. From the perspective of this reviewer, this omission diminishes the value of the review.

Second, while the database searches seem comprehensive, the Body is not explicit on how the final search terms yielded 1433 LCS and 686 LCINS articles. None of the queries in Supplementary Table 1 yielded these exact numbers. Moreover, given the significant number of articles that were removed from the original search results, a figure clarifying the number of articles that were excluded for various reasons should be provided. Third, while Table 2 provides a nice summary of the covered biomarkers, it would be more informative if information about where these biomarkers are in development (e.g., see the Early Detection Research Network five-phased approach https://edrn.nci.nih.gov/about-edrn/five-phase-approach-and-prospective-specimen-collection-retrospective-blinded-evaluation-study-design/) were provided.

We thank the reviewer for their time and expertise and the opportunity to improve this manuscript. Our responses to specific comments are detailed in the table below (all page and line numbers were recorded with 'Show Markup' on):

No.	Comment	Answer	Change in text
13	Studies not covering bodily fluids were excluded, and it is unclear why articles covering other biomarker sources are omitted.	As described in the methods, we searched for markers of LCS and never-smokers with high levels of evidence. Bodily fluids were targeted as our populations of interest (screening cohorts and never-smokers) would not have undergone more invasive interventions than blood, urine or sputum collections. As such we did not think to include nasal epithelia (or other	Please see "Body" section page 12 line 194-195. Text: Although we searched for studies without discrimination based on sample type (Figure 2), our review yielded blood- and sputum-borne markers only.

		minimally invasive non-body fluid source) as a targetable source of biomarkers. However, we have since been unable to find any studies investigating minimally-invasive non-body fluid biomarkers in our target cohorts Of note, we did not come across any studies that investigated markers in samples besides blood and sputum. We acknowledge this is not made clear in the text and have since mentioned our study did not identify markers with sufficient evidence from samples besides blood and sputum.	
14	Explain how 1433 LCS and 686 LCINS articles were reached, and include a figure clarifying article exclusions and reasons.	We have now clarified this in the manuscript. These numbers were achieved following de-duplication of articles once these were collated following the database searches.	Please see "Body" section page 12 line 187-193, and Supplementary Figures 1 and 2. Text: Our literature search yielded 1433 and 958 articles for LCS and LCINS, respectively. For LCS, our literature search yielded 1433 articles: 569 from MEDLINE, 141 from CINAHL, 436 from Embase and 287 from Web of Science. For LCINS, our literature resulted on 686 articles: 457 from MEDLINE, 54 from CINAHL, 215 from Embase and 232 from Web of Science. After removing duplicates, articles that do not cover the topic of body-fluid biomarkers for LCS or circulating biomarkers in never-smokers were excluded. Following screening,

			these were reduced to 22 and 91 total studies, respectively (Supplementary Figure 1 and 2).
15	Given the significant number of articles that were removed from the original search results, a figure clarifying the number of articles that were excluded for various reasons should be provided.	We agree that this is unclear. We have added flow charts as supplementary figures indicating the exclusion of articles and reasons for exclusion to aid clarity.	Please see Supplementary Figures 1 and 2. *Please see reviewer comment **5**.*
16	Include development phase of biomarkers listed in table 2.	We thank the reviewer for this suggestion and agree that this will enhance our analysis. We have employed this to add information to Table 2.	Please see Table 2 in section pages 10-11. *Please see reviewer comment **4** above.*