Peer Review File

Article information: https://dx.doi.org/10.21037/ccts-23-3

Reviewer Comments

Reviewer A

With interest, I have read this manuscript providing a narrative review on lung cancer screening. However, I believe a major revision is required before this manuscript can be accepted. I have listed my major and minor concerns in the attachment.

Major concerns:

1. The abstract conclusion is vague and not strong enough. Progress is good, but what was the most important and most recent progress? What should the additional studies that are mentioned still focus on?

Reply 1. The most important progress has been stated in the background of the abstract: "Randomized clinical trials and meta-analyses have firmly established low-dose computed tomography (LDCT) screening is an effective intervention to reduce lung cancer mortality."

We have revised the Key Content and Findings as well as the Conclusion to show the most important and recent progress as well as the additional studies that need to be done as follows:

Key Content and Findings: There is a risk threshold below which there is no mortality reduction benefits from LDCT screening. An accurate lung cancer risk prediction tool such as the PLCOm2012 is more cost-effective and has significantly higher sensitivity and positive predictive value for identifying individuals who will be diagnosed with lung cancer compares to age and pack-years criteria. Screening of light and never smokers in the general population who do not currently qualify for LDCT screening will require development of an accurate risk assessment tool that includes other important risk factors such as cumulative ambient air pollution exposures. The use of geospatial mapping tools that take into account screening, diagnostic work-up and treatment resource capacities, disparity in access by social-economically deprived and underserved populations to guide screening program improvement need to be evaluated. An accurate, personalized screening LDCT management protocol is key to minimize potential harms from radiation exposure due to unnecessary imaging studies, and adverse events from biopsy or surgery for benign disease. AI/deep learning tools incorporating time-dependent changes in smoking behavior, age and CT findings are promising approaches to tailor individual screening intervals or diagnostic work-up referral. Improvement in the smoking cessation rate in a screening population who are older and nicotine dependent requires wider use of pharmacotherapy in addition to counselling. Primary care providers play an important role in providing smoking cessation pharmacotherapy and management of additional findings.

Conclusions: Tremendous progress has been made in lung cancer screening. Additional studies to optimize screening eligibility criteria, overcome barriers to screening uptake, and personalized screening protocol can further improve the benefits/harms trade-offs.

2. In the abstract, “Implementation” is mentioned as one of the focus points of the paper, however, there is limited information provided in key findings on things like invitations, resource capacity etc.

Reply 2: The implementation point is very broad. Due to word limitation, it is impossible to cover every point. We focused on contemporary issues that are important. Please see the revised abstract in Reply 1 above.

3. The paper below provides an overview of current evidence on LCS in the UK, Europe and North America. What is new/different here?

The Future of Lung Cancer Screening: Current Challenges and Research Priorities
Amna Burzic, Emma L. O’Dowd and David R. Baldwin Cancer Manag Res. 2022; 14: 637–645.
Published online 2022 Feb 16. doi: 10.2147/CMAR.S293877

Reply 3. In a review paper, overlap of topics discussed is inevitable. The title of the topics covered in our review may appear to be similar on the surface but there are important differences in the content. The differences are listed in the table below:

Burzic et al. Cancer Manag Res. 2022	Current Manuscript
Defining & Selecting the Eligible Population	In section 3.1, we explained what is risk-based screening. The concept that there is a risk threshold below which there is no mortality reduction benefit from LDCT screening is often not recognised resulting in unnecessary and misleading discussion regarding screening in never smokers where LDCT is used as a general health examination even in young people without risk factors. We compared the differences between a risk model-based approach versus a risk factor-based approach for determining screening eligibility. We showed the advantages of the former approach. We pointed out the knowledge gap in defining high risk in light and never smokers who are not eligible for screening based on current inclusion criteria for heavy smokers. This is a hot topic in screening especially in Asian countries or immigrants from Asian countries.
Recruitment and Participation	In addition to discussing the barriers to screening uptake, we offer a potential solution to use geospatial tools that take into account screening, diagnostic work-up and treatment resource capacity, disparity in access by social-economically deprived and underserved populations to guide screening program improvement. The potential

application of machine learning tool using routine laboratory data such as Complete Blood Count in addition to clinical information to determine if a LDCT is warranted or the use of a point of care blood test as a pre-screening tool to optimize screening eligibility assessment in underserved populations is also new.

Psychological Factors Anxiety associated with screening is transient and is related to how we explain the benefits and harms to potential participants. We chose not to discuss this.

Screening Process, Nodule Management, Screening Interval In section 3.3, in addition to discussing traditional nodule management protocols, we discussed the evidence by us and others for biennial screening in lower risk individuals. The 4-IN-THE-LUNG-RUN (4ITLR) trial in Europe is still in progress while this strategy is already in use in Canada and Italy. We also discussed the use of deep learning to personalize the screening interval and the knowledge gap in the management of non-solid lung nodules

Incidental Findings Management In section 3.5, we emphasized only findings that have clinical implication should be reported as investigations may add considerable cost to screening program and cause harm to the patient.

Investigation Following Referral In section 3.4, in addition to discussing the problem of biopsy or surgery for benign lesions, we discussed the over-reliance of PET imaging in the diagnostic workup of small lung nodules, the rational choice of biopsy methods and how to mitigate a high benign nodule resection rate.

Smoking cessation Instead of saying the optimal strategy for integrating smoking cessation remains unknown, we pointed out improvement in the smoking cessation rate in a screening population who are older and nicotine dependent requires wider use of pharmacotherapy in addition to counselling.

Primary care providers play an important role in providing smoking cessation pharmacotherapy and management of additional findings.

Implementation Challenges and Unanswered Questions Some are common and some are different. We provided a “road map” how to address these challenges in different sections of our paper.

4. In line 82, the authors mention issues that need to be addressed for a screening program to have a substantial public health impact, including the safety of the screening management protocol, however, this aspect is not discussed in the remainder of the paper or mentioned as something that was searched for but could not be found in the literature. The same applies to factors influencing adherence and uptake of screening programs.

Reply 4. Safety of the screening management protocols has been adequately discussed in sections 3.3 and 3.4. Screening uptake has been discussed under section 3.2. Screening adherence varies with jurisdictions. There is evidence to suggest the effect

of one-time screening may last for several years. Participants with positive screens are followed up clinically for two or more years. Missing the next routine surveillance screen may not be as great a problem as portrayed.

5. Moreover, other important aspects which would be expected to be discussed in the literature, and are a crucial part of the discussion of lung cancer screening include the risk of overtreatment, radiation exposure, and resource capacity (limited capacity of radiologists and hospital equipment)

Reply 5. These have been discussed in Sections 3.3, 3.4 and 3.5. Getting a program off the ground and expand is better than doing nothing. There is no perfect health care system. Resource capacity is an ongoing challenge for every disease or program, not just for lung screening.

6. The benefits of lung cancer screening using LDCT are compared to chest X-ray results several times, however, the benefits of screening should be determined by comparing results to no screening.

Reply 6: We have clarified this in section 1.1 "Reduction in lung cancer mortality by screening with low dose computed tomography (LDCT) versus no screening or chest x-ray in high-risk individuals who have ever smoked has been firmly established by meta-analyses and randomized controlled trials that included the pivotal National Lung Screening Trial (NLST) and the Dutch-Belgian NELSON trial (2–6)." The NELSON trial compares screening with no screening.

7. 326-329: Major downsides are mentioned only very briefly, a further, more detailed discussion on these would be valuable.

Reply 7. The downside of screening such as radiation exposure has been discussed in Section 3.3 under Nodule Management. Adverse events related to biopsy or surgery have been discussed in the Diagnostic Workup section (3.4). In section 3.4, we have stated that "Since the major downsides of screening are complications related to radiation exposure, biopsies or surgery for benign disease, best practices need to be evaluated using quality indicators such as percent of LDCT screens which recommended additional imaging studies, positive predictive values of a diagnostic work-up biopsy and resection rate of benign lesions (113)."

8. 386-387: Providing suitable pharmacotherapy to increase quiting rates of smoking seems crucial to realise the potential benefits of lung cancer screening. Therefore, it is relevant that the authors indicate why such pharmacotherapy is not regularly provided even if this is inexpensive and effective as mentioned in line 382.

Reply 8. Pharmacotherapy using Varenicline is expensive and requires a prescription. In many countries or regions, even nicotine replacement therapy is not free. We have revised the section to read as follows: Cost and the need for a prescription may be the issue. Cytisine, an inexpensive partial agonist binding nicotine acetylcholine receptor nicotinic receptor similar to varenicline may be an option – achieving a 32% 12-month quit rate compared to 7% with counselling alone (144). Treatment of tobacco dependency and evaluation of the success of smoking cessation programs should consider the level of nicotine dependency which can be readily assessed by asking a simple question on time to first cigarette upon waking (125). Engagement of primary care providers to provide pharmacotherapy needs to be improved. Lobbying governments to provide free nicotine replacement therapy and other smoking cessation drugs should be a priority.

9. 395-397: The paper shows that a lot of research has already been done in this line to develop risk tools. What additional, needed research have the authors identified?

Reply 9: In section 3.1, we discussed the need to develop risk assessment tools for light and never smokers who do not meet current screening eligibility criteria for heavy smokers. In the Conclusion, we also pointed out that "to maximize the benefits of LDCT screening, research is needed to identify high risk individuals who are not currently eligible but would benefit from LDCT screening. The role of blood or breath biomarkers should be evaluated for potential application to identify high risk individuals for LDCT screening in randomized controlled studies (146)."

10. 397-398: The paper states that biomarkers such as blood tests need further investigation, while the conclusion states that these should be tested in RCTs. There is often some evidence indicating cost-effectiveness/health economic potential before an RCT is started. Please provide the already existing evidence warranting an RCT more explicitly.

Reply 10. The difference in opinion regarding RCT in lung screening tests is not new. Unfortunately, RCT is the gold standard to get reimbursed by health services providers because there is no comparator to allow a valid cost-effectiveness/health economic analysis with a single arm study. Potential is a commercial proposition.

11. 400-402: There are many studies on how to approach effective smoking cessation. Do you suggest comparing these studies/approaches with each other or conducting more specific empirical studies?

Reply 11: We agree with the reviewer that there are (too) many studies on how to approach effective smoking cessation. It is well established from Cochrane review and other studies that pharmacotherapy is important to improve the quit rate in this older, nicotine dependent population. We do not see a value of comparing studies/approaches with each other or conducting more specific empirical studies.

Minor concerns:

12. 49: Why is there a reference to indigenous people? Does the paper have a specific country in mind as a target audience?

Reply 12: Many countries e.g. USA, Canada, Australia, New Zealand, Brazil (to name a few), have indigenous people who are often neglected or underserved in the health care delivery system.

13. When the need for further evidence is discussed in the paper, there is always a focus on prospective RCTs, in particular in the context of AI, machine learning and optimising risk thresholds, performing RCTs may not be feasible. It would be interesting for the authors to discuss alternative study designs more suitable and more feasible in this context.

Reply 14. Please see Reply 10 above. Using AI as an example. The need for a prospective randomized clinical trial to validate the performance of AI algorithms become evident when we consider the inherent biases introduced by retrospective validation. Retrospective validation, while informative, is susceptible to bias due to the way individuals are followed based on recommendations from the standard of care rather than an AI algorithm. For example, if AI recommends an earlier imaging study or biopsy for a more aggressive tumor, this benefit cannot be observed in retrospective data because the patient did not have a visit at this earlier time point, and it is unknown whether the patient already had lung cancer at that time.

14. 67: The high incidence is not due to late diagnosis.

Reply14: Thank you for your comment. We have deleted the word incidence.

15. 74-77: Citations are necessary for every country, as far as I can find not all the listed countries have implemented programs. From the sources found in a quick search, screening has been implemented in the US and Canada, Poland, South Korea, Croatia, China, the Czech Republic, and Taiwan. As far as I can find, screening has not been implemented in South Africa, there was only a recommendation from the thoracic society of Southern Africa. The Netherlands does not have an implementation program, only an implementation trial.

Reply 15. We have corrected the list of countries that have organized screening program and countries that have formal commitment to implementing one. The source is from the Lung Cancer Policy Network (reference #7).

16. 92: The authors provide limited information in the Methods section on the

approach to searching the literature. Ideally, more information would be provided on how the search was approached. What were the search terms "related to" the keywords?

Reply16: The requested information can be found in Table 1, entitled "The search strategy summary".

17. Figure 1: Please indicate the data source underlying Figure 1.

Reply 17. The figure is from Reference #8. This has been added to the legend

18. 116: The value of Table 2 is very limited. It does not add to the explanation in the text and all PLCOm2012 predictors are also mentioned in the text.

Reply18: Table 2 is removed.

19. 125-129: "PLCOm2012 has …higher…more life years gained, has smaller number needed to screen…" Please specify the comparator. PLCOm2012 provides more life years gained than what? Additionally, sensitivity is an input to a cost-effectiveness analysis rather than a result of these analyses as implied in the current text.

R19: We have now clarified PLCOm2012 has statistically significantly higher sensitivity and positive predictive values (PPV) for identifying individual who will be diagnosed with lung cancer, (line 143-145).

20. 131: Please specify what the cost-effectiveness ratio was for the previous USPSTF recommendations for easy comparison and also state whether these ratios were gathered from the same cost-effectiveness model.

Reply 20. After consultation with my colleague, Rafael Meza, senior author of the study, to avoid confusion, we have changed the sentence to: A recent cost-effectiveness analysis compared risk-based strategies with the USPSTF 2013 and 2021 recommendations and showed that PLCOm2012 was cost-effective down to a risk threshold of 1.2%/6 years with an incremental cost-effectiveness being more cost-effective than the USPSTF 2013 or 2021 recommendations, which were both under the analysis' efficiency frontier and strongly dominated by other strategies that result in more QALYs for similar costs, while having a similar level of screening coverage (person ever screened 21.7% with PLCOm2012 1.2%/6 years versus 22.6% with USPSTF) (33).

21. 133-135: It is not clear what is meant here, as the screening criteria is not the screening tool. What is meant by "more sensitive screening the same number of participants"?

Reply 21: We have now clarified that: Head-to-head comparison between USPSTF2013 and PLCOm2012 showed the PLCOm2012 prediction tool was 15.8 per cent more sensitive in identifying individuals with lung cancer screening the same number of participants.

22. 137-138: Please provide a citation for this statement, because the fact that women typically accumulate fewer pack years does not necessarily mean that the risk is being underestimated. It could just mean that fewer women are included.

Reply 22: Reference 35. Pasquinelli MM, Tammemägi MC, Kovitz KL, Durham ML, Deliu Z, Guzman A, Rygalski K, Liu L, Koshy M, Finn P, Feldman LE. Addressing Sex Disparities in Lung Cancer Screening Eligibility: USPSTF vs PLCOm2012 Criteria. Chest. 2022 Jan;161(1):248-256. PMID: 34252436.
Fewer women would be eligible for screening because they would not meet the 30 pack-years criteria.

23. 148: If the PLCOm2012 prediction model is that good, it would be relevant to know why this model is (still) not being used in implemented screening programs.

Reply 23: This tool is currently used in the LCS program in Canada and UK. We have now added this to the text (line 138).
There is a perception (misperception) that a risk model would take up too much time. However, study in UK and our experience in Canada showed that eligibility assessment using PLCOm2012 takes an average of 5 to 10 minutes; most of this time is to take a detailed smoking history. The assessment is web based and can be done over the telephone. Considering committing someone to up to 27 years of screening, it seems prudent to obtain an accurate risk assessment.

24. 174-175: Please provide a citation for this claim.

Reply 24: We do not seem to see a claim there. All the statements had a reference to support.

25. 175-177: It is not clear how this sentence relates to the surrounding sentences. Why did the incidence increase? Was this atier the implementation of screening (so overdiagnosis)?

Reply 25: The incidence of stages 0-I lung cancer would increase if there is overdiagnosis, i.e. finding pre-invasive or indolent lung cancers that would not result in death of the patient would inflate the incidence rate without a corresponding decrease in mortality. These nodules, most of them non-solid (ground glass), would not be found if someone has not done a CT scan and then a biopsy and/or surgical resection).

26. 181-183: The risk assessment tool is only part of a screening/treatment strategy that as a whole may be cost-effective. The cut-off point used in combination with the risk assessment tool should result in a cost-effective application of lung cancer screening in a specific healthcare setting.

Reply 26: It has nothing to do with cut-off point if the risk assessment tool is inaccurate.

27. 202: Unclear, is it meant that the number of positive cases did not align with the lung cancer burden across the US?

Reply 27: It means more screen detected lung cancer cases are found from areas with a lower burden of lung cancer while we should be finding proportionately more lung cancers from high disease burden areas. This is a direct quote from the publication.

28. 203-204: It would be good to relate this information of differences in screening uptake with preference research looking into the barriers of participation.

Reply 28: Barriers to screening uptake is more than personal preference. As discussed in our paper, there are practical issues such as socio-economic and cultural barriers. For example, if a person is poor, he/she cannot afford to take time off from work, or pay for parking to go to the screening examination. Lack of an attached primary care provider is another major reason.

29. 256-257: Please provide a quantification of the benefits induced by the PanCan protocol as used in Canada.

Reply 29: If two thirds of the people have biennial screenings instead of annual screening, and if a CT scan costs $200 (depending on the jurisdiction), the minimum cost savings would be $200 x the number of screens saved plus other overhead costs. This does not include personal expenses such as travel time, time off work, parking etc. This level of health economic analysis does not seem to belong to a review article. The details can be found in references 30 and 32.

30. 294-296: One screening trial is being cited, while the sentence seems to refer to multiple trials “in screening trials”. Please specify in which trial this was the case.

Reply 30: We have now clarified the benign resection rate can be as high as 34% in one screening trial (reference 99). This figure was from a well-known academic centre in the US.

31. 336-338: The perspective on incidental findings can be extended with the perspective of multi-disease screening using LDCT.

Reply 31: Management of additional (incidental) findings on a screening LDCT is a subject on its own. With the word limit, we can only point out that "only findings that have clinical implication should be reported as investigations may add considerable cost to screening program and cause harm to the patient (117). When reported, information should be provided to the health care provider regarding the recommended next step." The perspective of multi-disease screening using LDCT is appealing but its benefit or cost-effectiveness has not been firmly established.

32. Table1: Please distinguish between exclusion and inclusion criteria

Reply 32: This has been modified in the text.

33. It would be good to include more precise search criteria

Reply 33: We have now included more precise search criteria in the text.

Textual comments:
49 & 139: peoples -> people: We have changed to people in the text

186: Please define CEA and SNP: the definitions are now spelled out

189: Please be consistent in using the same full stop for decimals such as in line 167. Agree
222: was -> were: corrected
229: define CBC: Complete Blood Count
259: define MILD: Multicentric Italian Lung Detection
262: define 4ITLR: 4-IN-THE-LUNG-RUN

More textual improvements can be made. Will leave to the discretion of the editor.

Reviewer B
I think this is an excellent and important review. I loved reading it. I thought it was extensive, complete and well written. The reference list was lenghty but given this being a review paper, I believe it is more than worth allowing for all of the references cited. The figure and tables were appropriate.

I have no specific edits for the manuscript itself. I have only one problem with the paper that I believe must be corrected - its title. When I first accepted reviewing this paper, its title communicated to me that this was going to be a review of the failure to integrate tobacco cessation with lung cancer screening. In fact, only 2 pages of this lengthy review is devoted to that topic. Essentially the same space as the authors use for each of the many other issues related to lung cancer screening discussed in this

paper. I am not asking for a more extensive review of tobacco cessation. The space allotted to this issue is appropriate. What I am asking for is a change in the paper's title - i might suggest 2 more accurate title options -
1. Lung Cancer Screening - Contemporary Issues. or
2. Contemporary Issues Related to Lung Cancer Screening

Reply 1. The title was given to us. Lung Cancer Screening - Contemporary Issues seems good to us if the editor agrees with the change.

Reviewer C
This is an interesting topic. Lung cancer screening detected pulmonary nodules are quite common these days. How to balance the benefits and potential harms of a screening program is worth discussion. Following are my suggestions for this paper.

Major
1. The "Key Content and Findings" of the abstract needs to be rewritten. Please simplify the language and avoid repetition. The description of the risk prediction tool is difficult to understand. In addition, I do suggest the authors to ask some native speakers to help improve the language.

Reply 1: We followed the journal abstract format. We have thoroughly revised the abstract highlighting the key points we made in the table under Reply 3 to Reviewer #1's comments.

2. Please add more contents to the "Background" part to make the readers understand what is the scientific problem you are going to discuss.

Reply 2: We have switched some of the material from 1.2 to 1.1 and clarify the purpose of this review.

3. In paragraph 1.2, according to the title, a comparison of the present screening program and an efficient ideal program should be made. But the authors only listed the requirements for an effective screening program to meet, no gaps mentioned at all.

Reply 3. Section 1.2 addressed the contemporary issues relevant to health professionals and policy makers regarding lung cancer screening and the optimal approach to integrating smoking cessation with screening. Knowledge gaps are discussed in the different sections of the main text.

4. Line 159-Line 198: Lung cancer risk of non-smokers are discussed in these two paragraphs. Why will the authors talk about overdiagnosis in between ("Despite these tantalizing findings, there is concern about overdiagnosis when screening is done as an annual health examination or health check-ups especially among younger women "). Please reorganize this part and make it more logical.

Reply 4. The section has been revised as follows:
However, a study in Taiwan showed a 6-fold increase in the incidence of stages 0-I lung cancer from 2004 to 2018 with no significant change in the incidence of stages II-IV lung cancer (46). The Shanghai Cancer Surveillance Study in China from 2002 to 2017 also showed a sharp rise in early-stage adenocarcinoma, particularly among young women with no significant decline in late-stage disease and mortality (47). These studies raised concern about overdiagnosis especially in younger females without identifiable lung cancer risk factor(s) (46, 47). However, these studies did not discuss LDCT screening as a general health examination that includes younger individuals with no known lung cancer risk factor is different than screening based on risk factors such as family history of lung cancer, genetics, outdoor and indoor exposures (48). In addition, overdiagnosis related to the relatively high proportion of adenocarcinomas in-situ (Stage 0) can be reduced by improvement in our understanding of the molecular biology of nonsolid (ground-glass) nodules and refinement of the lung nodule management protocol (48).

Minor
1. In Table 1, no exclusion criteria are included, please add this content, and specify the selection process.

Reply 1: We have included the exclusion criteria in the table.

2. Line 102-104 The purpose of using an example here is to prove age is not enough to predict lung cancer risk. However, the comparison is made between LDCT and chest x-ray. Please explain this. And after this comparison, the authors give a figure of smokers' lung cancer mortality rate. But the authors haven't mentioned screening beneficial before. Please keep the content coherent and logical.

Reply 2. In section 1.1, we have clarified that: Reduction in lung cancer mortality by screening with low dose computed tomography (LDCT) versus no screening or chest x-ray in high-risk individuals who have ever smoked has been firmly established by meta-analyses and randomized controlled trials that included the pivotal National Lung Screening Trial (NLST) and the Dutch-Belgian NELSON trial (2–6).

3. Please add a simple title for Figure 1.

Reply 3. The title of Figure 1 was in the legend. We have changed it to the title.

4. Line259-273; Line 275-287 In these two paragraphs, authors discuss about the follow-up time interval and the GGN size threshold for screening. I suggest change the "3.3 Lung Nodule Management Protocols" into "3.3 Lung Nodule Management"

Reply 4. We have changed the title to 3.3 Lung Nodule Management