PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

Title (Provisional)

Assessment of the impact of multi-cancer early detection test screening intervals on late-stage cancer at diagnosis and mortality using a state-transition model

Authors

Rous, Brian; Clarke, Christina A.; Hubbell, Earl; SASIENI, PETER

VERSION 1 - REVIEW

Reviewer	**1**
Name	**Carr, David**
Affiliation	**Wayne State University, Pathology**
Date	**25-Jun-2024**
COI	**None**

This is an important study modeling screening intervals for MCED. The methods are good and the authors provide a thorough estimate of screening performance across many cancers and many stages of cancer. The discussion should be reviewed for clarity and to better define what factors are being used to determine an optimal screening interval.

The authors performed a modelling study of multi-cancer screening performance across a range of screening intervals, and under two tumor growth rate scenarios, using previously published test performance characteristics for the GRAIL Galleri test. This study is an extension of prior work by these authors, and is important to inform our ongoing investigations of MCED tests as they are introduced into clinical practice.

The study design and methodology are good with feasible screening intervals chosen and clinically important outcome metrics modeled. The biggest items that caught my attention are in the discussion, which often feels disconnected from the stated objectives of the paper and specific items investigated in this study. I would suggest taking a close look at the discussion to fine-tune the messaging around the major take home of the paper (given the objective stated in the abstract) and remove extraneous discussion points not related to the questions addressed in this study.

Specific comments and suggestions are below.

Page 4, lines 8-9
“This study uses reliable datasets”
Remove the term ‘reliable datasets’ from this bullet point.

Page 5, lines 18-33
This paragraph highlights MCED tests. It would be helpful to mention that many different MCEDs are in development, testing a range of analytes, and that the optimal testing strategy including analytes tested and testing intervals are still undetermined.

Page 5, lines 47-51
By comparison, some tests detect invasive cancer signals and typically…
Seems like a reference is warranted, along with perhaps a mention of the cancer being referred to. Are the authors getting at lung cancer screening here?

Page 5, lines 8-11
"Thus screening intervals must be developed to maximize the benefits for the greatest number of people…"
I think this statement would apply to all nearly all medical practices. Perhaps it should say something like "maximize the benefits across a range of cancer types with different clinical features and growth rates"

Page 11 line 25-26
"screening with an MCED test have the potential to intercept a large fraction of all cancers before they reach a late stage"
"A large fraction" is too non-specific a term, particularly given the large number of interval cancers that occur with any screening strategy. Would be better to use an actual estimate from validation studies or clinical trial data – what % of cancers were diagnosed at stage 1 and what % at stage 2 (these should be mentioned separately and not combine into single number). This informs not only the need to improve MCED test performance, but also the importance of implementing MCED in addition to standard of care screening.

Page 11, line 40-42
"….modeling is required to select screening intervals that would then be investigated in clinical trials and real world evidence"
Remove "and real world evidence."

Page 11 lines 45-55 and page 12 lines 3-11
This paragraph compares the estimated number of deaths averted with a MCED strategy to total deaths from various combinations of common cancer types. It reads like an additional results paragraph ("To put our MCED test modeling results in context, we compared"). This is a rather confusing way of considering these numbers. I would suggest removing this paragraph.

Page 12, lines 33-37
"Due to these necessary modeling assumptions, the benefits of MCED screening on mortality in clinical practice may vary from those this model estimated."
This statement should be modified to emphasize that this is an upper bounds estimate. Thus, under these test performance assumptions, real-world benefits are predicted to be less than those estimated in the model

Page 13, lines 14-20
"As a consequence of this likely bias towards fast growing cancers…."

The authors highlight why overdiagnosis is less of a concern with MCED tests than standard screening. The authors should also discuss the flip side of this coin which is that limited stage I and II sensitivity may limit the efficacy of MCED screening overall. As has been reported in real-world studies, bias towards fast growing cancers means that many patients will have cancers detected at later stages when cure is no longer be possible which may limit the efficacy of MCED.

Page 13, lines 23-37
This paragraph feels out of place. This study doesn't address the question of MCED tests used instead of standard of care screening. The idea of physicians taking the opportunity to emphasize the benefit of screening is very speculative, and doesn't feel appropriate for this study or at this point in the discussion. I would suggest removing.

Page 13, lines 40-55; page 14, lines 4-9
Paragraph touching on two unrelated ideas. It begins discussing cost effectiveness. However, the authors don't mention their own finding that biennial screening is "more efficient" than annual screening in terms of # of deaths averted/100,00 tests. If the authors want to address cost-effectiveness they should discuss the implications of this finding. If they feel cost-effectiveness is "beyond the scope of this analysis", then this portion of the discussion can be removed.

The sentence "This analysis shows that an MCED that can detect a shared cancer signal across more than 50 cancer types with one simple blood sample (10) may detect cancer earlier than usual care" should be removed. This study did not show that, but rather used data from prior studies.

The paragraph then transitions to a discussion of reducing inequalities in cancer screening. The suggestion of improved outcomes in this group is completely speculative. Considering the Galleri test is currently only available if paid out-of-pocket and requires additional testing and procedures to investigate positive results, which present similar socioeconomic barriers to those currently affecting cancer screening and care, it's not at all clear that this would reduce inequalities in cancer mortality. This portion of the discussion should be removed.

Page 15, lines 47-53
I'd suggest revisiting and strengthening the last paragraph of the discussion. The following thoughts/comments come to mind when reading this final paragraph, but really apply generally to the whole discussion, which could be reviewed and edited for clarity and focus in messaging.

I'm not clear what is meant by "staggering annual" screening.

It feels like the authors are suggesting a "best screening interval" (my words not theirs), but never actually explain the criteria they use to determine an ideal screening interval.

- "even at a suboptimal interval, addition of an MCED test complements current guidelines-based cancer screenings". How is an optimal interval defined?
- Related is the reference to "acceptable lifetime risk," which as I read it is a risk of false positives. What would an unacceptable risk be? Given that you find "no consistent trend to overdiagnosis with differing screening intervals" wouldn't the determination of an optimal interval rest on something else?
- Does this mean that a 6 month interval is superior given it detects more cancers without a trend towards increased overdiagnosis? Does something like cost effectiveness come into play here? If that were the case, does the increased efficiency of biennial screening in terms of deaths averted/100,00 tests suggest that biennial is optimal?

Reviewer	**2**
Name	**Hammer, Anne**
Affiliation	**Aarhus University, Department of Clinical Medicine**
Date	**31-Jul-2024**
COI	**none**

Thanks for the invitation to review the paper entitled “Sensitivity of reductions in late-stage cancer and mortality to screening interval choice in a multi-cancer early detection state-transition model”

Overall, the paper is well written and clinically relevant. I have a few comments the authors should consider

a. Please revise the title as it is very long and hard to understand. I had to read it several times and it wasn’t until I had read the paper I somewhat understood.

b. Abstract does not accurately summarize the paper. Most readers will not be familiar with MCED, and so, there is a need to explain this a bit more in the abstract. Please explain the tumor growth scenarios in the methods as this will make it easier to read the results.

c. Page 3, please add to the last sentence the number of states contributing with data. Most readers will not know that this is not a nationwide database.

d. Please add more information on which cancer types that are detectable using the MCED. Readers should not need to read other papers to understand.

e. Page 5, last sentence: add “to reduce mortality”

f. How were the dwell times estimated?

g. It appears that specificity was estimated in another study and only briefly mentioned in the current paper. I would appreciate more information on sensitivity, FFP and NPV overall and if any differences were observed between cancer types.

h. What about assumptions regarding death after treatment? Does the model incorporate previous results on stage-specific survival for each cancer type for the estimation of the number of deaths averted? Supplementary Figures? This is not clear in the manuscript.

i. Page 13. You might want to add that transition depends on the histological subtype of the cancer. For example, non-endometrioid uterine cancers are more aggressive than the more common endometrioid uterine cancer.

j. Figures are very difficult to assess as they are small, text is overlapping the figures and there are no headings to the figures.

VERSION 1 - AUTHOR RESPONSE

Reviewer: 1 [PLEASE SEE ATTACHED FILE FOR FULL/ADDITIONAL COMMENTS FROM REVIEWER 1]
Dr. David Carr, Wayne State University

Comments to the Author:
This is an important study modeling screening intervals for MCED. The methods are good and the authors provide a thorough estimate of screening performance across many cancers and many stages of cancer. The discussion should be reviewed for clarity and to better define what factors are being used to determine an optimal screening interval.

The authors performed a modelling study of multi-cancer screening performance across a range of screening intervals, and under two tumor growth rate scenarios, using previously published test performance characteristics for the GRAIL Galleri test. This study is an extension of prior work by these authors, and is important to inform our ongoing investigations of MCED tests as they are introduced into clinical practice.

The study design and methodology are good with feasible screening intervals chosen and clinically important outcome metrics modeled. The biggest items that caught my attention are in the discussion, which often feels disconnected from the stated objectives of the paper and specific items investigated in this study. I would suggest taking a close look at the discussion to fine-tune the messaging around the major take home of the paper (given the objective stated in the abstract) and remove extraneous discussion points not related to the questions addressed in this study.

Specific comments and suggestions are below.

Page 4, lines 8-9

“This study uses reliable datasets”

Remove the term ‘reliable datasets’ from this bullet point.

> **Author Response:** We have deleted the words ‘reliable datasets’ from the bullet point, which now reads: *"This study uses performance estimates from a published case-control study and outcomes from the Surveillance, Epidemiology, and End Results (SEER) database, a widely used database for modeling studies."*

Page 5, lines 18-33

This paragraph highlights MCED tests. It would be helpful to mention that many different MCEDs are in development, testing a range of analytes, and that the optimal testing strategy including analytes tested and testing intervals are still undetermined.

> **Author Response:** Thank you for bringing up this point, we have expanded this paragraph to include that other MCED tests are under development and highlighted that the optimal testing strategy is still undetermined.

The paragraph now reads: "*Multi-cancer early detection (MCED) tests are innovative new technologies that utilize cancer biology to screen for a broad set of cancer types with a single blood sample (10, 11). There are several MCED tests currently under development that utilize a variety of different analytes to detect a cancer signal (12). Cell-free DNA (cfDNA) is one such analyte that can be shed by tumors into the bloodstream and cancarry cancer-specific signals (13, 14). By analyzing circulating cfDNA, in combination with machine learning, an MCED test (Galleri®; GRAIL, Inc, Menlo Park, CA) has been developed to detect this shared cancer signal with high specificity (10). This test can complement, though not replace, existing single-cancer screenings, as well as expand categories of screenable cancers (11). The high specificity of this test decreases the likelihood of false positives and the subsequent extraneous triage and diagnostic testing. However, practical strategies for cancer screening using cfDNA, including the interval of screening tests, remain to be determined." Introduction, paragraph 2.*

Page 5, lines 47-51

By comparison, some tests detect invasive cancer signals and typically…

Seems like a reference is warranted, along with perhaps a mention of the cancer being referred to. Are the authors getting at lung cancer screening here?

Author Response: We agree with the reviewer that the paragraph was unclear. We have revised and reordered the paragraph to more clearly contrast the various types of screening effects (affecting pre-cancerous lesions, early-stage detection, potential for overdiagnosis).

The revised paragraph now reads: *"Most guideline-based single-cancer screenings are conducted every one to five years, depending on various factors, including the cancer growth rate (2, 15-19). By detecting precancerous lesions, some single-cancer screenings have the potential to reduce cancer incidence and can be performed at longer intervals based on the precancerous lesion growth rate (20, 21). By comparison, some tests, such as low-dose computerized tomography screening for lung cancer, detect invasive cancer signals and typically need to be conducted relatively frequently to most effectively detect cancer in early stages to reduce mortality (22). Tests for asymptomatic cancers may lead to overdiagnosis, such as in the case of screening for prostate cancer (23). Selecting an optimal screening interval must balance the possibility of improved and prolonged life due to earlier cancer detection against false positive test results and overdiagnosis, which could lead to unnecessary testing and treatment (24)." Introduction, paragraph 3.*

Page 5, lines 8-11

"Thus screening intervals must be developed to maximize the benefits for the greatest number of people…"

I think this statement would apply to all nearly all medical practices. Perhaps it should say something like "maximize the benefits across a range of cancer types with different clinical features and growth rates"

Author Response: We thank the reviewer for the suggestion. The sentence has been reworded accordingly to be more specific.

The sentence now reads: *"Thus, screening intervals must be developed to maximize the benefits across individuals who may develop a range of cancer types with different clinical features and growth rates, rather than optimizing for a single cancer type." Introduction, paragraph 5.*

Page 11 line 25-26

"screening with an MCED test have the potential to intercept a large fraction of all cancers before they reach a late stage"

"A large fraction" is too non-specific a term, particularly given the large number of interval cancers that occur with any screening strategy. Would be better to use an actual estimate from validation studies or clinical trial data – what % of cancers were diagnosed at stage 1 and what % at stage 2 (these should be mentioned separately and not combine into single number). This informs not only the need to improve MCED test performance, but also the importance of implementing MCED in addition to standard of care screening.

Author Response: We thank the reviewer for pointing this out - this is a modeling study using published sensitivity parameters from completed case-control validation studies. We have added to this paragraph 1) the proportion depleted from stages III-IV, as described in Table 1, and 2) the stage I-II breakdown. In addition, we have expanded our examination of non-shedding cancer cases later in the Discussion section, noting that even current limited performance can lead to a significant depletion in late-stage incidence in long-term screening.

Page 11, line 40-42

"….modeling is required to select screening intervals that would then be investigated in clinical trials and real world evidence"

Remove "and real world evidence."

Author Response: The text referred to has been deleted.

Page 11 lines 45-55 and page 12 lines 3-11

This paragraph compares the estimated number of deaths averted with a MCED strategy to total deaths from various combinations of common cancer types. It reads like an additional results paragraph ("To put our MCED test modeling results in context, we compared"). This is a rather confusing way of considering these numbers. I would suggest removing this paragraph.

Author Response: We appreciate the reviewer's insight. We agree that this paragraph may fit better within the results and have moved it to that section. The paragraph has also been simplified to focus on the illustrative comparison of the modeled results to hypothetical perfect prevention of deaths for groups of individual cancers in order to contextualize the scale of the effect on cancer deaths.

Page 12, lines 33-37

"Due to these necessary modeling assumptions, the benefits of MCED screening on mortality in clinical practice may vary from those this model estimated."

This statement should be modified to emphasize that this is an upper bounds estimate. Thus, under these test performance assumptions, real-world benefits are predicted to be less than those estimated in the model

> **Author Response:** The sentence has been modified to reflect your suggestion and now reads: *"Due to these necessary modeling assumptions, real-world benefits are likely to be less than those estimated in the model." Discussion, paragraph 2.*

Page 13, lines 14-20

"As a consequence of this likely bias towards fast growing cancers...."

The authors highlight why overdiagnosis is less of a concern with MCED tests than standard screening.

The authors should also discuss the flip side of this coin which is that limited stage I and II sensitivity may limit the efficacy of MCED screening overall. As has been reported in real-world studies, bias towards fast growing cancers means that many patients will have cancers detected at later stages when cure is no longer be possible which may limit the efficacy of MCED.

> **Author Response:** We agree that this is an important consideration. Although we have explicitly modeled scenarios involving fast growing cancers, the fast growth rate does not directly imply detection at late stage. This is determined by the time spent in early stages and detectability of the cancer by MCED test while at those early stages, which is why modeling is required. Importantly, prevalence round screening does not have a substantial opportunity to detect cancers at early stages, as cancers are at whatever stages they happen to be at the time of the initial blood draw. Here, we modeled the effect of sustained screening when the pre-existing stage distribution has been swept out by the prevalent round screen, as this is most representative of long-term performance of cancer screening programs. As we expand on in the Discussion, even the current limited performance of MCED detection can have a significant effect on reducing late-stage cancer.
>
> We have also added statements to the Discussion addressing cfDNA non-shedding cancers:
>
> *"Cancers that shed cfDNA in a limited amount at early stages, cancers that do not shed, or cancers that grow rapidly may be diagnosed at late stage by usual care in the interval between MCED tests. If shedding onset only occurs at late stage, cancers may be found earlier by an MCED test, but still in a late stage where curative treatment is less likely to be possible. It is therefore necessary to model across cancer types and stages to account for these variations rather than using an average estimate of performance. Even current performance numbers provide an opportunity to reduce late stage cancer incidence (Figure S6)." Discussion, paragraph 4.*

Page 13, lines 23-37

This paragraph feels out of place. This study doesn't address the question of MCED tests used instead of standard of care screening. The idea of physicians taking the opportunity to emphasize the benefit of screening is very speculative, and doesn't feel appropriate for this study or at this point in the discussion. I would suggest removing.

> **Author Response:** In accordance with this suggestion, we have removed the last four sentences of this paragraph to be more focused with our Discussion. The sentence explaining that MCED screens are meant to complement existing recommended screening has been retained, as we feel that it is important to clarify for readers the potential relationship between MCED screening and usual care.

Page 13, lines 40-55; page 14, lines 4-9

Paragraph touching on two unrelated ideas. It begins discussing cost effectiveness. However, the authors don't mention their own finding that biennial screening is "more efficient" than annual screening in terms of # of deaths averted/100,00 tests. If the authors want to address cost-effectiveness they should discuss the implications of this finding. If they feel cost-effectiveness is "beyond the scope of this analysis", then this portion of the discussion can be removed.

The sentence "This analysis shows that an MCED that can detect a shared cancer signal across more than 50 cancer types with one simple blood sample (10) may detect cancer earlier than usual care" should be removed. This study did not show that, but rather used data from prior studies.

The paragraph then transitions to a discussion of reducing inequalities in cancer screening. The suggestion of improved outcomes in this group is completely speculative. Considering the Galleri test is currently only available if paid out-of-pocket and requires additional testing and procedures to investigate positive results, which present similar socioeconomic barriers to those currently affecting cancer screening and care, it's not at all clear that this would reduce inequalities in cancer mortality.

This portion of the discussion should be removed.

> **Author Response:** We appreciate the reviewer's perspective. This paragraph has been removed to focus the Discussion on the direct implications of the modeling analysis.

Page 15, lines 47-53

I'd suggest revisiting and strengthening the last paragraph of the discussion. The following thoughts/comments come to mind when reading this final paragraph, but really apply generally to the whole discussion, which could be reviewed and edited for clarity and focus in messaging.

- I'm not clear what is meant by "staggering annual" screening. It feels like the authors are suggesting a "best screening interval" (my words not theirs), but never actually explain the criteria they use to determine an ideal screening interval.

Author Response: The final paragraph has been revised to more thoroughly and plainly state the key takeaways from the analysis. The term 'staggering' has been removed to help with clarity.

The Conclusion now reads: "*In conclusion, annual MCED screening has a lifetime risk of false positive results comparable to the status quo of single-cancer screening and is predicted to result in downstaging of diagnosed cancers under a variety of hypothetical scenarios, including fast and aggressive tumor growth. Biennial screening was shown to be more efficient in terms of PPV, but with a noticeable decrease in potential reductions in late stage diagnoses due to fewer people screened. The optimal choice of screening interval will depend on assessments of real-world cancer survival and the costs of confirmatory testing after MCED screening. However, both annual and biennial MCED screening intervals have the potential to avert deaths associated with late-stage cancers when used in addition to current guideline-based cancer screening.*"

- "even at a suboptimal interval, addition of an MCED test complements current guidelines-based cancer screenings". How is an optimal interval defined? Related is the reference to "acceptable lifetime risk," which as I read it is a risk of false positives.

  Author Response: We thank the reviewer for pointing out this ambiguity. The intent of the Conclusion is to convey that the risk of a false positive result in the case of annual MCED screening is not appreciably greater than what is expected from usual care. As well, while biennial screening is less efficient in terms of reductions in late-stage diagnoses, it is still predicted to have a positive impact relative to usual care alone. The conclusion paragraph has been revised to more clearly convey these points to the reader.

- What would an unacceptable risk be? Given that you find "no consistent trend to overdiagnosis with differing screening intervals" wouldn't the determination of an optimal interval rest on something else?

  Author Response: We understand that the usage of 'acceptable' in this case may be misleadingly subjective. Our original goal was to contrast the benefits of MCED screening at different intervals with the usual care. We believe that our modeling results show that the potential harms of MCED screening are not significantly different from the status quo. We have revised the conclusion paragraph to remove these terms and add greater detail.

- Does this mean that a 6 month interval is superior given it detects more cancers without a trend towards increased overdiagnosis? Does something like cost effectiveness come into play here? If that were the case, does the increased efficiency of biennial screening in terms of deaths averted/100,00 tests suggest that biennial is optimal?

  Author Response: We agree with the reviewer that cost-effectiveness must be considered when optimizing a screening interval and that comprehensive consideration is out of scope for this paper. In our Discussion, we focus on annual and biennial screening intervals because they represent a reasonable range between the extremes of 6 month and 3-year intervals also modeled in our analysis. We hope it is now clear that our intended contribution is to understand the interplay of false positives, number of tests, and deaths and not to solve all needed inputs to a full understanding of an

optimal screening program so as to inform the design of future clinical trials and models.

Reviewer: 2
Dr. Anne Hammer, Aarhus University, Gødstrup Hospital

Comments to the Author:
Thanks for the invitation to review the paper entitled “Sensitivity of reductions in late-stage cancer and mortality to screening interval choice in a multi-cancer early detection state-transition model”

Overall, the paper is well written and clinically relevant. I have a few comments the authors should consider

a. Please revise the title as it is very long and hard to understand. I had to read it several times and it wasn’t until I had read the paper I somewhat understood.

Author Response: We thank the reviewer for providing this helpful feedback. The title has been revised and now reads:

“Assessment of the impact of multi-cancer early detection test screening intervals on late-stage cancer at diagnosis and mortality using a state-transition model”

b. Abstract does not accurately summarize the paper. Most readers will not be familiar with MCED, and so, there is a need to explain this a bit more in the abstract. Please explain the tumor growth scenarios in the methods as this will make it easier to read the results.

Author Response: We appreciate the reviewer for pointing this out. The Abstract has been updated to introduce MCED tests and provide details on the tumor growth scenarios.

c. Page 3, please add to the last sentence the number of states contributing with data. Most readers will not know that this is not a nationwide database.

Author Response: We have specified on page 3 as well as in the Methods section that the SEER18 database covers 14 US states, constituting 28% of the US population.

d. Please add more information on which cancer types that are detectable using the MCED. Readers should not need to read other papers to understand.

Author Response: We thank the reviewer for calling attention to this. We modeled the totality of cancer incidence, as almost all cancers can be detected to some extent by this assay (see Klein et al, 2021; doi: 10.1016/j.annonc.2021.05.806). A full list of organ-specific cancer groupings that we modeled as potentially being affected by the MCED test has been added to the first paragraph of the Methods section.

Detailed sensitivity information for each of these cancer groups is found in Supplementary Figure S1 for reference. Additionally, a table has been added to the Supplemental Information (Supplemental Table S1) that details the ICD-I-3 definitions and histological types of cancer reported by the MCED test and used to collect SEER data.

e. Page 5, last sentence: add “to reduce mortality”

Author Response: This suggested addition has been made.

f. How were the dwell times estimated?

Author Response: Estimation of dwell times was detailed in the previous analysis of Hubbell, et al. 2021 (doi: 10.1158/1055-9965.EPI-20-1134). A brief description of the methodology has been added to the Supplemental Information:

“Dwell times were estimated from a group of experts as noted in Hubbell et al (2). Due to uncertainty in these opinions, sensitivity analysis was done by examining multiple scenarios. Recent data suggests results from the fast aggressive scenario closely resembles detection rates in one biobank analysis (3). Briefly, preclinical sojourn time, defined as the total time before diagnosis of an invasive cancer in the course of usual care, was divided into dwell time per cancer type and stage. Two tumor growth scenarios were used: fast and aggressively fast. Dwell time for each cancer type and stage was approximated with an exponential distribution based on multiple previously published models (4-6). The stage at which usual care would diagnose a cancer is subject to a competing risk (discovery by usual care) and assigned a shorter dwell time distribution. Because cancer is a progressive disease, both tumor growth scenarios assumed that later stages spanned less time than earlier stages.” Supplementary Information, paragraph 4.

g. It appears that specificity was estimated in another study and only briefly mentioned in the current paper. I would appreciate more information on sensitivity, FFP and NPV overall and if any differences were observed between cancer types.

Author Response: Sensitivity did vary between stages and cancer types, as depicted in Supplementary Figure S1. False positives (FPR) is an overall rate and not per cancer - the number of individuals receiving a FPR only depends on the number of tests and not the number of cancers modeled or tested. We have added a brief comment in the high-level modeling summary describing this important property:

“False positives occur at a rate depending on the number of tests performed, and do not depend on the number of cancer types modeled or tested for.” Methods, paragraph 2

In this paper, we are interested in modeling the effects of different screening intervals on late-stage cancer diagnosis and on mortality. Interval cancers or false negatives are found clinically at the same stage and therefore have no effect on mortality because patients are not worse off than with usual care. Negative predictive value is therefore not directly reported in this paper. While these statistics are beyond the scope of this

analysis, they are discussed in further detail in Klein, et al. 2021 (doi: 10.1016/j.annonc.2021.05.806.)

h. What about assumptions regarding death after treatment? Does the model incorporate previous results on stage-specific survival for each cancer type for the estimation of the number of deaths averted? Supplementary Figures? This is not clear in the manuscript.

Author Response: We thank the reviewer for noting this omission. This study builds upon an earlier paper and thus we have provided a high-level summary of the approach in this manuscript. A sentence has been added to methods to further clarify this point.

"The impact of early cancer detection by MCED on mortality was modeled by substituting the hazard of death appropriate for the earlier stage at screen-detection (accounting for lead time) for the hazard of death appropriate for the later stage at which clinical diagnosis would have occurred in the absence of screening. We estimated the overall impact on mortality by combining the separate hazards for each cancer type and stage." Methods, paragraph 2.

For further reference, we have also added Supplementary Figure S6 depicting stage-specific survival for each cancer type modeled in this analysis.

i. Page 13. You might want to add that transition depends on the histological subtype of the cancer. For example, non-endometrioid uterine cancers are more aggressive than the more common endometrioid uterine cancer.

Author Response: We appreciate the reviewer's suggestion. We have added a new statement regarding histological subtypes to the discussion, which reads:

"Not all cancers will progress sequentially through stages I to IV and some may skip stages. For example, some fraction of cancer cases may become metastatic early, and transition from stage I to stage IV. In particular, certain histological subtypes may be more or less aggressive than average and thus impact estimations of cancer stage shifting or mortality effects due to MCED screening. Complex distributions of dwell times are also possible." Discussion, paragraph 7.

j. Figures are very difficult to assess as they are small, text is overlapping the figures and there are no headings to the figures.

Author Response: The overlapping text that appears in the figures is a result of the pdf compilation process, with watermarks overlayed on the figure text that do not appear in the original high-resolution figures. Headings have been added to multi-panel figures for clarity and the legends modified slightly.

VERSION 2 - REVIEW

Reviewer 1
Name Carr, David
Affiliation Wayne State University, Pathology
Date 18-Dec-2024
COI

The authors have addressed the majority of items raised by both reviewers. This has improved the messaging and scope of the manuscript. There are still a few items that I would raise and a few specific suggestions/edits that I include.

Page 3, lines 53-55, page 4, lines 3-6
The final sentence of the results section of abstract is a run-on sentence.

Page 6, lines 18-20
remove “utilize cancer biology to”; I’m not sure what this term is supposed to mean

Page 6, lines 32-35
“The high specificity of this test decreases the likelihood of false positives and the subsequent extraneous triage and diagnostic testing…” Decreases compared to what? I think I understand the point but this should be reworded. If used in addition to standard screening MCED will only increase the likelihood of false positives and extraneous testing for a given population. I presume what you are intending to say is that MCED has lower false positive rates than other currently accepted screening modalities.

Page 6, lines 52-55
“Tests for asymptomatic indolent cancers may lead to overdiagnosis, such as in the case of screening for prostate cancer”
Again, I think I understand the point, but should be reworded. All screening is by definition for asymptomatic cancers and I don’t think anyone is explicitly testing for indolent cancer. I would reword slightly to state that any screening modality will lead to some overdiagnosis with the exact rates dependent on the frequency and detection of slow growing cancers.

Page 13 lines 9-25
I still don’t understand the purpose of this paragraph comparing deaths averted from MCED to some “hypothetical perfect screening technology” that would eliminate all deaths from particular cancers. This paragraph should be removed.

Page 14. Lines 31-33
I think there is one major potential harm of MCED that remains under discussed: the potential for patients to forgo established screening in favor of MCED. One recent paper addressed this in regards to colon cancer and found that there is a potential to increase CRC deaths if too many patients choose blood-based screening over colonoscopy (https://doi.org/10.7326/ANNALS-24-00910). This possibility should be mentioned when acknowledging the limitations of idealized modelling assumptions and possible downsides of MCED.

Page 15, lines 42-48

"The purpose of this model was to evaluate the sensitivity of MCED screening to differing schedules of screening following usual practice by modeling the sensitivity under the assumption of ideal practice, including screening adherence and diagnostic follow-up."
This sentence is unclear.

Page 16, lines 10-15
"Screening individuals aged 75-79 (or those with otherwise reduced life-expectancy) may be considered less cost-effective or not to have a favorable benefit:harms ratio"
It feels out of place to specifically discuss ages 75-79 in a single sentence without any context. Given that routine MCED is yet to be established clinically screening any age may not have a favorable benefit:harms ratio. Perhaps just remove this sentence.

Page 16, lines 38-40
"These data suggest that a 6-month screening interval would be too short, and a 3-year interval too long."
This is an important but underdeveloped statement in the discussion. Why is 6 months too short? And why is 3 years too long? Perhaps the authors feel a bit limited in making too strong a point from a modeling study, but this is the crux of the issue for both clinical trial design and clinical implementation.

I'd suggest this be moved to after the following few sentences about the prospective cohort study which found a cancer signal up to 3 years before diagnosis (seems like this is a big reason why 3 years is too long). I'm assuming here a bit, but 6 months is likely too short because of cost to the system and the fact that it is a clinically unfeasible time point to expect patients to come in for a blood draw. For me it would help the flow of the paper to understand more explicitly why annual and biennial strategies were chosen. The methods only state that "Annual and biennial screening intervals were modeled for most analyses, though 6-month intervals from 0 to 3 years were examined and are shown for some figures." I feel like important information is contained in these decisions, and wish it were discussed more explicitly.

VERSION 2 - AUTHOR RESPONSE

Reviewer: 1
Dr. David Carr, Wayne State University
Comments to the Author:
Thanks for the opportunity to review this resubmission. The authors have incorporated the majority of reviewed comments and submitted an improved manuscript. I offer a few remaining comments and suggested edits.

** ** **

The authors have addressed the majority of items raised by both reviewers. This has improved the messaging and scope of the manuscript. There are still a few items that I would raise and a few specific suggestions/edits that I include.

We sincerely appreciate the reviewer's response to our revisions and are grateful for their insightful feedback that we agree will help enhance the clarity and impact of our manuscript. Below are our responses to their comments.

Page 3, lines 53-55, page 4, lines 3-6
The final sentence of the results section of abstract is a run-on sentence.

The abstract Results section has been updated.

Page 6, lines 18-20
remove "utilize cancer biology to"; I'm not sure what this term is supposed to mean

This term relates to utilizing circulating tumor DNA, a facet of tumor biology, as an analyte to screen for cancer. Given that this is not a focus of our study, we have removed the phrase as requested.

Page 6, lines 32-35
"The high specificity of this test decreases the likelihood of false positives and the subsequent extraneous triage and diagnostic testing…" Decreases compared to what? I think I understand the point but this should be reworded. If used in addition to standard screening MCED will only increase the likelihood of false positives and extraneous testing for a given population. I presume what you are intending to say is that MCED has lower false positive rates than other currently accepted screening modalities.

This interpretation is correct; we thank the reviewer for pointing out that this could be better communicated. We have revised the sentence to emphasize that the high specificity of MCED screening tests allows for more types of cancers to be screened without a proportional increase in the rate of false positives currently seen with recommended screening.

The statement now reads: "*Owing to their high specificity, MCED tests are unlikely to significantly increase the overall rate of false positives already seen with accepted single-cancer screening modalities*." Introduction, paragraph 2.

Page 6, lines 52-55
"Tests for asymptomatic indolent cancers may lead to overdiagnosis, such as in the case of screening for prostate cancer"
Again, I think I understand the point, but should be reworded. All screening is by definition for asymptomatic cancers and I don't think anyone is explicitly testing for indolent cancer. I would reword slightly to state that any screening modality will lead to some overdiagnosis with the exact rates dependent on the frequency and detection of slow growing cancers.

We agree with the reviewer's observation and have incorporated their suggested language. The statement now reads: "*The degree by which a population-level cancer screening program contributes to overdiagnosis depends on the sensitivity of the test to indolent cancers, the incidence of slow-growing cancers in the population, and the upper age of screening*." Introduction, paragraph 3.

Page 13 lines 9-25
I still don't understand the purpose of this paragraph comparing deaths averted from MCED to some "hypothetical perfect screening technology" that would eliminate all deaths from particular cancers. This paragraph should be removed.

> It was our intention to illustrate the maximal benefit indicated by our model in terms the reader would readily understand, ie, the potential number of deaths averted via earlier diagnosis through regular MCED screening. Given that the comparison of MCED screening to a hypothetical perfect screening technology may be a distraction, we have removed them to allow the results to speak for themselves.

Page 14. Lines 31-33
I think there is one major potential harm of MCED that remains under discussed: the potential for patients to forgo established screening in favor of MCED. One recent paper addressed this in regards to colon cancer and found that there is a potential to increase CRC deaths if too many patients choose blood-based screening over colonoscopy (https://doi.org/10.7326/ANNALS-24-00910). This possibility should be mentioned when acknowledging the limitations of idealized modelling assumptions and possible downsides of MCED.

> We agree that this topic is of interest for the optimal implementation of MCED screening. As with other considerations we discuss, use of MCED as an alternative to USPSTF-recommended screening is outside of the ideal practice assumed by our model in order to specifically isolate the impact of screening intervals on mortality. The clinical best practice would be for MCED screening to be performed in addition to all current recommended screening.
>
> We have added the following sentence to our discussion of limitations of our study assumptions: "*Individuals may also elect against recommendations and warnings otherwise to substitute MCED screening for recommended single-cancer screening, thereby constraining potential mortality benefits.*" Discussion, paragraph 2.
>
> We have also added the following to our discussion of potential harms associated with MCED screening: "*Because standard-of-care screening can identify early-stage cancers that MCED tests are less likely to detect, the incidence of malignant cancers that progressed from more indolent lesions may increase among individuals who replace single-cancer screening with MCED screening alone. To minimize this potential harm, MCED screening is intended to be performed in addition to USPSTF guideline-recommended screening practices, which were assumed to occur as part of our model.*" Discussion, paragraph 4.

Page 15, lines 42-48
"The purpose of this model was to evaluate the sensitivity of MCED screening to differing schedules of screening following usual practice by modeling the sensitivity under the assumption of ideal practice, including screening adherence and diagnostic follow-up."
This sentence is unclear.

We have revised this sentence to better describe the aim of our paper. The text now reads: "*The purpose of this model was to evaluate how sensitive the projected mortality benefits of MCED screening are to differing schedules of screening. Our modeling followed standard practice by assuming ideal screening practice, including screening adherence and diagnostic follow-up, in order to isolate the impact of screening schedules from other factors that would otherwise influence screening effectiveness.*" Discussion, paragraph 5.

Page 16, lines 10-15
"Screening individuals aged 75-79 (or those with otherwise reduced life-expectancy) may be considered less cost-effective or not to have a favorable benefit:harms ratio"
It feels out of place to specifically discuss ages 75-79 in a single sentence without any context. Given that routine MCED is yet to be established clinically screening any age may not have a favorable benefit:harms ratio. Perhaps just remove this sentence.

We have removed the sentence as suggested.

Page 16, lines 38-40
"These data suggest that a 6-month screening interval would be too short, and a 3-year interval too long."
This is an important but underdeveloped statement in the discussion. Why is 6 months too short? And why is 3 years too long? Perhaps the authors feel a bit limited in making too strong a point from a modeling study, but this is the crux of the issue for both clinical trial design and clinical implementation.

I'd suggest this be moved to after the following few sentences about the prospective cohort study which found a cancer signal up to 3 years before diagnosis (seems like this is a big reason why 3 years is too long). I'm assuming here a bit, but 6 months is likely too short because of cost to the system and the fact that it is a clinically unfeasible time point to expect patients to come in for a blood draw. For me it would help the flow of the paper to understand more explicitly why annual and biennial strategies were chosen. The methods only state that "Annual and biennial screening intervals were modeled for most analyses, though 6-month intervals from 0 to 3 years were examined and are shown for some figures." I feel like important information is contained in these decisions, and wish it were discussed more explicitly.

We have moved that statement forward as suggested and added several statements discussing the implications of the longest and shortest screening intervals. These points frame our conclusion that annual or biennial screening intervals are more realistic starting points for future MCED screening trials and programs.

The expanded discussion now reads*: "Though tumor growth rates for cfDNA-shedding cancers are incompletely understood, our analysis and recent studies suggest that a 3-year screening interval may be too long and allow excessive interval cancers. In a prospective cohort study of the MCED test using blood samples collected from participants diagnosed with cancer within 3 years of blood draw, a cancer signal was detected up to 3 years before diagnosis, with test positive rate increasing progressively with shorter preclinical timescales [47]. Retroactive assessment of plasma samples in two large prospective biobank studies suggests that preclinical detectability of cancer*

signals resembles the tumor growth rates examined here [48]. Additionally, while the shortest interval of 6 months would have the greatest impact on mortality, this benefit may be outweighed by the cost and procedural burden on healthcare providers and patients. The effect of screening saturates as fewer newly detectable cancers arise in the interval between screens, leading to a maximum number of lives that can be saved [25]. Even continuous MCED screening cannot find cancers that do not shed significant levels of ctDNA by the time of clinical diagnosis. The results of the present analysis suggest that although the annual and biennial intervals between these two extremes are expected to have noticeable differences in expected mortality, they may be optimal for the design of future MCED screening programs." Discussion, paragraph 6.