Author’s response to reviews

Title: Navigating the Sustainability Landscape: A Systematic Review of Sustainability Approaches in Healthcare

Authors:

Laura Lennox (l.lennox@imperial.ac.uk)

Lynne Maher (lynne.maher@middlemore.co.nz)

Julie Reed (julie.reed02@imperial.ac.uk)

Version: 1 Date: 15 Dec 2017

Author’s response to reviews:

Dear Editor and Reviewers,

We would like to thank you for your time in reviewing this paper. Your thoughtful questions, comments and suggestions for the manuscript have been greatly appreciated.

We have now responded to the comments and addressed suggestions in the updated manuscript. Our response has been organised to first address the editorial comments followed by each reviewer in turn. All changes within the manuscript have been marked as tracked changes.

We believe the changes to the paper have greatly added to its clarity, readability and interest for readers. We look forward to your decision in due course.

Best wishes,

Authors
Editors Comments:

1. Was the search conducted by an information scientist? Was it peer reviewed using PRESS?

Response:

- The search strategy was developed and conducted in collaboration with a medical librarian at Imperial College London. This process involved identifying and testing search terms and reviewing returned papers for relevancy. This was used to iteratively identify which search terms to include in our review. The paper was not peer-reviewed using PRESS. The involvement of a medical librarian during the conception and retrieval of papers was deemed to be sufficient by authors. The manuscript was however registered on the PROSPERO database which assesses the submission and publication of key information about the design and conduct of the systematic review. Although this is not peer-reviewed, this requires all reviews to be assessed to ensure that necessary baseline data have been provided.

We have now specified the involvement of the medical librarian in the manuscript text: “The selection of databases, search terms and search strategy was supported by a medical librarian to ensure an overall quality and coverage of the systematic review.” On page 7 line 3-5.

2. Was the search limited to English-language articles?

Response:

- Yes the search was limited to English language articles. We have now specified within the manuscript on page 7 line 18-19.

3. What is the potential impact of having a single person conduct screening, abstraction and coding?

Response: Thank you for this question and highlighting our need to address this limitation. Although double data extraction is recommended it is also recognised that this is often not possible in many cases due to time and resources issues. Unfortunately, in our case, it was not possible to have all authors conduct all steps within the review due to these pragmatic constraints. It is recommend that ‘if single data extraction is used, reviewers and readers need to be mindful of the possibility for more errors and the potential impact these errors may have on effect estimates’. (Buscemi, 2006) We recognise that this approach may have led to some bias in exclusion or inclusion of particular articles but we attempted to mitigate this where possible by involving multiple authors at particular stages.
Specifically, there was input from co-authors in 20% of screening, data extraction for 100% of articles was checked by a second author and coding for 25% of papers was verified by a second author. Discussions between all authors took place throughout each stage to discuss any questions or issues encountered. We believe this allowed us to mitigate some bias where possible but we also recognise that this may be seen as a limitation so we have now addressed this in our discussion section on page 24 lines 14-19.

“The use of one author to conduct of the majority of screening, data extraction and coding is also a limitation of this work. Although double data extraction is recommended in most systematic reviews, it is also recognised that this is often not possible in many cases due to time and resources constraints.[94] This may have resulted in bias in inclusion or exclusion or resulted in missing or erroneous information being collected. To address this limitation we involved multiple authors where possible in selection of the studies (20% screened by a second author) and coding of constructs (25% of studies). Data extraction was also checked against full text articles for all included papers.”

Reviewer #1 Comments:

1. A major limitation of the review, that is only briefly addressed in the paragraph on the limitations of the review in the Discussion section of the paper, is that the review does not include any assessment of the quality of evidence or the soundness of theory behind the described methods, constructs, etc.

Response:

- Thank you for highlighting this issue within the manuscript. We appreciate the opportunity to respond and describe our assessment process.

This work focused on mapping current sustainability approaches, identifying knowledge gaps and understanding if there was consensus on sustainability constructs across healthcare fields. From the outset, we did not want to make judgements on the approaches themselves or their constructs but aimed to provide readers with information on available methods, describing their proposed approaches and how they have been designed.

While we explored the use of existing quality assessment tools for our studies, we found they were not fit for purpose as they did not take into account the exploratory nature of the paper, with many of the criteria not seen as helpful for assessing articles introducing sustainability methods. For example, many of the papers identifying sustainability approaches did not address the tool as the topic of interest or research question. Approaches were frequently introduced as mechanisms to conduct the research, therefore the data collected within each study was not an indication of the soundness of the approach but of the individual study. For example in the study
by Bray et al in 2009, the ‘sustainability pyramid model’ was developed before the study was conducted on the basis of the authors’ observations as participants in quality improvement (QI) collaboratives.

As we wanted to scope the literature for available methods, excluding articles based on inappropriate and unrelated research questions was considered counterproductive. We therefore chose to develop a quality assessment that responded to the needs of our article. We designed a data quality and extraction table that would allow us to capture the findings we believe were most relevant to readers. The data extraction form was developed to ensure the baseline quality expected- all articles had to clearly describe the proposed approach and give information about how it could be used in practice. Our data extraction form allowed us to understand if the included articles were providing sufficient information to readers. Here we list each criteria with the rationale for its inclusion and its benefit to readers:

1. Purpose of use (Why the tool has been designed),
2. Healthcare Setting (where has it been designed for?),
3. Level of Healthcare Use (What area of healthcare has it been designed for?),
4. Description of application (How should/can it be applied?),
5. Constructs identified (What has been deemed important to sustainability),
6. Scoring Mechanism (What data will be generated?),
7. Target user (Who can use the tool?),
8. Definition of Sustainability (How have the authors defined sustainability? Does this resonate with the readers views on sustainability?),
9. Theoretical underpinning (What underlying assumptions are there?)
10. Sustainability Perspective (Is sustainability being assessed or influenced?) and
11. Approach development details (How the method created? Do we think this is sound?)

We believe this criteria allowed us to assess that each paper had a minimum level of data to allow readers to make the final judgements on the appropriateness of a method for their settings. Details particularly around approach development may be used by teams to assess whether they believe the approach has enough merit to be used in their site.
Although we believe this approach has strengthened the relevancy and resonance for readers we also acknowledge that this may be seen as a limitation by others. Therefore we have addressed this within the main manuscript acknowledging this limitation and making it clear to readers that assessment of the individual tools has not been done using an existing tool. The first change is in the methods section where we describe why an established quality assessment checklist was not deemed appropriate for our study on page 8 lines 2-19.

“A quality assessment and data extraction form was developed for identified articles. Existing quality assessments were explored but it has been noted that available quality assessment approaches often fail to consider the rationale and context of studies.[22,23] Their use to determine the inclusion of qualitative studies is often not recommended as many existing tools do not capture the multiple meanings of “good quality” and “rightness” therefore studies should often not be excluded based on this quality assessment.[22,24] The available assessments were not sensitive to the inclusion criteria of our study which was exploratory in nature. We sought to provide an overview of available approaches for sustainability and designed our data extraction form to identify and describe the included articles. The aim of the data extraction was to report descriptions and study information not to ascertain validity of the approaches or their constructs.”

We have also acknowledged this as a potential limitation in the discussion section on page 25 line 12-21.

“Another key limitation of this work is that we did not use an existing quality assessment tool and cannot attribute value or accuracy of constructs from each approach. While the quality criteria set out in our data extraction allowed us to ensure each paper had a minimum level of data to adequately describe the approach, it did not assess quality of the approaches themselves. We extracted information on each of the approaches which others may wish to use to attribute validity to findings. Details, particularly those around sustainability approach development, may be used by readers to assess whether they believe the approach has enough merit to be used in their site. It is important to note that many approaches (24%) were informed by professional expertise, a technique that may be difficult to assess for quality but appears to be very significant in the creation of sustainability approaches.”

This comment also raised another important question which we have now highlighted as future work. ‘How can these approaches be assessed for accuracy or soundness?’ There is currently no clear standard for assessing the quality of these approaches but we believe this can only be assessed by reviewing their use and application in future studies. This will ascertain whether they have included the correct constructs and if they are able to impact sustainability processes. We see identifying and describing the available approaches for sustainability, as a significant first
step in aiding teams to understand what tools are available and describing how they can be applied in practice. Prior to this article these approaches remained dispersed across various fields with little evidence that they are being applied in practice. It is hoped this review will allow for a more structured approach to testing these tools in practice by providing a resource where available approaches are documented and their characteristics are described. Future work will explore the evidence of use of these approaches in practice which will help us to understand the true impact of these tools, unfortunately we cannot address it fully in this review.

We have now updated a sections within our discussion to highlight the need for assessment of these approaches to understand if and how they impact sustainability. Page 25 line 24 to page 26 line 5.

In the discussion: “Although consensus across approaches was found with 6 constructs included in over 75% of approaches we cannot be certain these are the ‘right’ constructs to assess sustainability. In order to understand the validity of these findings the approaches must be assessed in practice. Future work will explore if and how these approaches have been applied to ascertain if their constructs accurately represent sustainability in specific settings and if they fulfil their stated aims.”

2. The approach of categorising and counting different ‘perspectives’ seems overly simplistic.

Response:

- Thank you for this comment. We chose to outline the multiple theoretical perspectives underpinning each of the approaches as a mechanism to demonstrate the varied and complex views guiding the creation of the sustainability approaches. While more information could be added to describe the theories in detail we believe this would distract from the overall message which is that diverse perspectives exist and inform how sustainability is viewed.

This section could also be removed if it is deemed too simplistic but we believe it adds value in its current form in demonstrating the multiple ways of looking at sustainability in the current literature.

3. Similarly, the reference in the review to the findings as ‘evidence’ is questionable.

Response:

- Thank you for this comment. We appreciate the need to cautiously represent our findings so we have removed the term ‘evidence’ from multiple sections of the results on page 22 line 23 and page 23 lines 3 and 9.
4. The claim in the paper that the framework presented in Table 3 represents a 'consolidated' framework seems to be an over-statement of the findings of the review.

Response:

- In order to develop our framework we compared and contrasted sustainability constructs across all methods. This process involved combining constructs across published approaches that had different representations, language or classifications but which similar or overlapping in concepts in their definition. This process also involved breaking down large amounts of information to parse their underling concepts. In this way we believe this represents a consolidated framework (Similar work was conducted to develop the Consolidated Framework For Implementation Research (CFIR) by Damschorder et al.) While we believe our framework does represent a consolidation of the available sustainability constructs, we are willing to consider adjusting the name of the framework to an ‘overview’ or ‘summary’ framework if this is deemed to be more representative of the work presented.

5. The reason why I say above that the topic of sustainability is somewhat fraught is that it seems increasingly clear that there is little in healthcare that should be sustained, beyond a culture of learning and improvement. We are in a world of constant evolution and change, and one where everything about us can and should be improved. This may be one reason behind the difficulty of finding a suitable definition for 'sustainability in healthcare'.

The authors refer to this difficulty of definition in the first paragraph of page 5, and to a 'working definition' which includes ... (page 5, lines 8-10). But they do not provide an actual working definition for the review. On page 10 they discuss the definitions that are used in the papers included in the review, but there is no critical analysis of these definitions. Dictionary definitions of sustainability are informative in describing how the term can be used as either a process or as a noun, and this could provide a useful basis for assessing how the term is used in the scientific healthcare literature. The question about what should really be sustained is left hanging ...

Response:

- Thank you for this thought provoking comment and for your helpful suggestions regarding framing definitions.

We agree that with the evolution of healthcare practices, sustainability may not always be a desirable or feasible goal. Unfortunately, there are examples in the literature of ineffective practices continuing for long periods of time (The Drug Assistance Resistance Education program continued to be operated by thousands of police departments despite studies showing
that it has little effect on the prevention or reduction of drug use by youth (Frumkin & Reingold, 2004.) This is of course wasteful and inappropriate. While sustainability of every initiative is not always appropriate, we also recognise that sustainability is still required in order to not only understand the long term impact and effectiveness of many interventions and practices but to also promote spread and uptake of successful practices across new settings. Without sustainability of beneficial improvements we encounter many issues which we have outlined in our introduction such as variation in care, waste of scarce resources and declining staff morale.

We agree that stagnancy is not desirable but maintaining improved patient outcomes, processes and experience should be the focus of initiatives, understanding that adaption and changes are needed to evolve with new evidence and emerging needs of systems. We have addressed this within the manuscript by adding new text on sustainability as a process incorporating adaption and continuous improvement. Updated text on page 5 line 10-21 reads:

Sustainability is often viewed as an ‘outcome’ where health benefits, activities or workforce capacity are maintained.[8] Some have cautioned against this linear perspective on sustainability as it ‘does not take account of the recursive or reflexive character of sustainability and learning or of the continuous adjustments that shape the sustainability process’.[15] More recently the ability to adapt and continuously improve has also been recognized as a potential definition of sustainability.[16] This concept of sustainability as a ‘process’ rather than an ‘outcome’, incorporates concepts of adaptation, learning and continuous development.[17] This lens allows sustainability to be viewed as a change process that can be influenced by individuals throughout initiatives by continuing to develop and adapt in response to the needs of the system.[17–19]

For the purposes of the review we did not want to judge definitions as ‘right’ or ‘wrong’ but describe the definitions held within each of the approaches found. This will allow readers to judge if and how these definitions resonate with their settings or needs. We have now included an explicit definition of sustainability on page 5 line 21-24. We believe this will aid readers in understanding our perspective on what sustainability meant to us while the review was conducted.

“For the purposes of this work, ‘sustainability’ will refer to the general continuation and maintenance of a desirable feature of an initiative and its associated outcomes as well as the process taken to adapt and developed in response to emerging needs of the system. This definition as well as any additional domains found will be explored in the review.”

Although we have not stated which definition may be most desirable we have added some critical reflections on the potential risks of only assessing programme maintenance. We caution that if this single measure is used, ineffective programmes may be maintained and other more adaptive programmes may be judged to have failed. This text has been included on page 21 line 18 to page 22 line 7.
“Results have demonstrated that sustainability is most often defined and assessed as the maintenance of programme activities. Although multiple definitions were found (continuation of the health benefits from an initiative, capacity built in the workforce or community, further development or adaptation, and the ability to recover costs) there was a clear dependence on this one measure which has been previously observed in the literature.[92] It is important to note that while measuring continuation of programme activities is important to assessing sustainability, relying solely on this measure may risk other key sustainability variables being missed.[14,92] For example it may result in the continuation of ineffective or undesirable practices if health benefits are not taken into account. This was observed in with the Drug Assistance Resistance Education program in America which continued to be implemented in schools despite studies showing that it had little effect on prevention or reduction of drug use by students.[93] Using continuation of programme activities as the sole measure of sustainability also risks initiative being unfairly judge as failing to sustain if activities are adapted. If the definition is broadened, adaptation could also signify sustained improvement, especially if the adaptations contributed further to health benefits or cost recovery. These examples highlight the need for care consideration of what will be sustained and what evidence there is for sustainability to occur.[45]

6. The discussion of limitations should provide a more critical assessment of the methods of the review, and what can be learned from the review.

Response:

- Thank you for this comment. We have now expanded our limitations section to provide a more critical assessment of the methods for our review as well as key learning to be taken from the review.

We have highlighted that risks associated with one author to conducting the majority of the search and data extraction. Although double data extraction is recommended it is also recognised that this is often not possible in many cases due to time and resources issues. Unfortunately, in our case, it was not possible to have all authors conduct all steps within the review due to these pragmatic constraints. It is recommend that ‘if single data extraction is used, reviewers and readers need to be mindful of the possibility for more errors and the potential impact these errors may have on effect estimates’. (Buscemi, 2006) We recognise that this approach may have led to some bias in exclusion or inclusion of particular articles but we attempted to mitigate this where possible by involving multiple authors at particular stages. Specifically, there was input from co-authors in 20% of screening, data extraction for all articles was checked by a second author and coding for 25% of papers was verified by a second author. Discussions between all authors took place throughout each stage to discuss any questions or issues encountered. We believe this allowed us to mitigate some bias where possible but we also recognise that this may
be seen as a limitation so we have now addressed this in our discussion section on page 25 lines 1-8.

“The use of one author to conduct of the majority of screening, data extraction and coding is also a limitation of this work. Although double data extraction is recommended in most systematic reviews, it is also recognised that this is often not possible in many cases due to time and resources constraints.[94] This may have resulted in bias in inclusion or exclusion or resulted in missing or erroneous information being collected. To address this limitation we involved multiple authors where possible in selection of the studies (20% screened by a second author) and coding of constructs (25% of studies). Data extraction was also checked against full text articles for all included papers.”

We have also highlighted the possible limitations of not using a quality assessment framework on pages 25 line 12-21.

“Another key limitation of this work is that we did not use an existing quality assessment tool and cannot attribute value or accuracy of constructs from each approach. While the quality criteria set out in our data extraction form allowed us to ensure each paper had a minimum level of data to adequately describe the approach, it did not assess quality of the approaches themselves. We extracted information on each of the approaches which others may wish to use to attribute validity to findings. Details, particularly those around sustainability approach development, may be used by readers to assess whether they believe the approach has enough merit to be used in their site. It is important to note that many approaches (24%) were informed by professional expertise, a technique that may be difficult to assess for quality but appears to be very significant in the creation of sustainability approaches.”

We have also expanded on the lessons to be taken from the review in the Strengths and Limitations section on page 23 line 21 to page 24 line 5.

“This is the first review to consolidate available approaches for sustainability across diverse healthcare settings. We believe this work represents a significant contribution to the field in organising and describing sustainability approaches which have until now remained isolated across healthcare fields and disciplines.[14] This review provides not only a resource for identifying available sustainability approaches but also outlines the aims, applications and constructs in each approach so readers can determine if one may be fit for their setting. This work has demonstrated that although many approaches were developed within specific interventions and settings, similar constructs for sustainability were found indicating general
learning can be gathered from across settings to inform sustainability processes and research. Additionally, this paper provides a consolidated summary of all constructs deemed to be important across approaches to serve as a sustainability knowledge base that is useful beyond specific settings or interventions.”

7. The claim in the conclusion that this review 'provides a valuable resource to researchers, healthcare professionals, and improvement practitioners ...' is questionable. There is no clear case made for why or how this review represents a significant advance in the understanding of sustainability.

Response: Thank you for highlighting the need for a change in this language. We hope the review will be a valuable resource but we do not yet know if it will be. We have changed the manuscript to reflect this and removed the statement. We have now added the paragraph outlined in the previous comment highlighting the contribution we believe the paper is making to the understanding of sustainability approaches. page 23 line 21 to page 24 line 5.

8. The title of Table 1 would more appropriately refer to the list of papers included in the review, rather than the list of methods.

Response:
- This has been changed on page 9 and page 33.

Reviewer #2 Comments:

1. In the title, is "methods" that best term to describe what you've studied? The purpose of the study, you write, is "to describe what methods are available to assess and influence sustainability in healthcare...". You find these to include frameworks, models, tools, strategies, a checklist and a process. You go on to identify common constructs across the studies that you have reviewed. From your Table 3, the constructs appear to be the variables that may be used to assess and influence sustainability; i.e., workload, complexity, staff involvement, urgency, etc. You also refer to "perspectives, applications, and constructs with these methods". I suggest that you just use the terminology of "constructs" in the paper rather than "methods". For many readers of IS, the term "methods" will be assumed to mean research methods. This is demonstrably not what you mean. Adjust the title, language in the abstract and text, and table and figure headers.
Response:

- Thank you for your suggestion and highlighting the confusion our previous terminology may have caused readers. We agree the term methods can be seen as misleading. We have discussed the use of an alternative terms and while the term ‘constructs’ was helpful we believe it does not accurately represent our search for constructs for sustainability within ‘structured approaches’ to influence sustainability.

We have decided to change our language to ‘approaches to influence sustainability’. We believe the term ‘approaches’ is broad enough to encompass the ‘models, tools, and frameworks’ included in our review without causing the confusion of methodology used. We have updated this language throughout the manuscript.

2. The authors communicate a strong bias towards sustainability in the abstract and text that sustainability is good. Is this appropriate, logically? That is, is long-running sustained use of an intervention in healthcare always a good thing? What if the intervention isn't optimal? Or what if other researchers--or indeed the same researchers--create a better intervention; wouldn't we want to see the first one displaced rather than sustained? What if a newer intervention is just as effective but less costly? We wouldn't want sustained use then, right? Or what if the host healthcare system or its patients change, as they do all the time, and a previously implemented intervention is no longer matching the needs of the organization or of its patients? My point is that sustainability can occur, perhaps as predicted by constructs identified in your study, but no longer be desirable. A less normative orientation to the topic in the manuscript would serve your purposes better.

Response:

- Thank you for this thought provoking comment. We agree that with the evolution of healthcare practices, sustainability may not always be a desirable or feasible goal. Unfortunately, there are examples in the literature of ineffective practices continuing for long periods of time despite studies showing that they have little effect on desired outcomes (Frumkin & Reingold, 2004.) This is of course wasteful and inappropriate. While sustainability of every initiative is not always appropriate, we also recognise that sustainability is still required in order to not only understand the long term impact and effectiveness of many interventions and practices but to also promote spread and uptake of successful practices across new settings. Without sustainability of beneficial improvements we encounter many issues which we have outlined in our introduction such as variation in care, waste of scarce resources and declining staff morale.

Stagnancy is not desirable but maintaining improved patient outcomes, processes and experience should be the focus of initiatives, understanding that adaption and changes are needed to evolve
with new evidence and emerging needs of systems. We have addressed this within the manuscript by adding new text on sustainability as a process incorporating adaptation and continuous improvement. Updated text on page 5 line 10-21 reads:

“Sustainability is often viewed as an ‘outcome’ where health benefits, activities or workforce capacity are maintained.[8] Some have cautioned against this linear perspective on sustainability as it ‘does not take account of the recursive or reflexive character of sustainability and learning or of the continuous adjustments that shape the sustainability process’. [15] More recently the ability to adapt and continuously improve has also been recognized as a potential definition of sustainability.[16] This concept of sustainability as a ‘process’ rather than an ‘outcome’, incorporates concepts of adaptation, learning and continuous development.[17] This lens allows sustainability to be viewed as a change process that can be influenced by individuals throughout initiatives by continuing to develop and adapt in response to the needs of the system.[17–19]”

We have now included an explicit definition of sustainability which we believe will aid readers in understanding our perspective on what sustainability meant to us while the review was conducted. This updated definition accounts for the need to be aware of the system and emerging changes. Page 5 line 21-24.

“For the purposes of this work, ‘sustainability’ will refer to the general continuation and maintenance of a desirable feature of an initiative and its associated outcomes as well as the process taken to adapt and developed in response to emerging needs of the system. This definition as well as any additional domains found will be explored in the review.”

Thank you for your comments on when sustainability may not be desired. To address this in the manuscript we have attempted to provide a more critical narrative around the importance of understanding the need to sustain and how this is defined. We caution that if a sustainability measure is used such as continuation of a programme, ineffective programmes may be maintained and other more adaptive programmes may be judged to have failed. This text has been included on page 21 line 18 to page 22 line 7.

“Results have demonstrated that sustainability is most often defined and assessed as the maintenance of programme activities. Although multiple definitions were found (continuation of the health benefits from an initiative, capacity built in the workforce or community, further development or adaptation, and the ability to recover costs) there was a clear dependence on this one measure which has been previously observed in the literature.[92] It is important to note that while measuring continuation of programme activities is important to assessing sustainability, relying solely on this measure may risk other key sustainability variables being missed.[14,92] For example it may result in the continuation of ineffective or undesirable practices if health
benefits are not taken into account. This was observed in with the Drug Assistance Resistance Education program in America which continued to be implemented in schools despite studies showing that it had little effect on prevention or reduction of drug use by students.[93] Using continuation of programme activities as the sole measure of sustainability also risks initiative being unfairly judge as failing to sustain if activities are adapted. If the definition is broadened, adaptation could also signify sustained improvement, especially if the adaptations contributed further to health benefits or cost recovery. These examples highlight the need for care consideration of what will be sustained and what evidence there is for sustainability to occur.[45]

3. Your study is about to identify 40 constructs across 62 publications. How can you claim that "relatively little is known about how to influence sustainability"? Revise this statement. Perhaps it is a problem of disorganization that you wish to call out?

Response:
- Thank you for highlighting this issue. We have adapted the text on page 5 line 1-2.

“Despite recognition of this challenge and considerable research conducted in this area, relatively little is known about how to translate this evidence into action to support the long term impact of improvement efforts.[12]”

Reviewer # 3 Comments:

1. The working definition of sustainability is provided (continuation of the health benefits from an initiative, continuation of initiative activities and capacity built in the workforce or community), but it is not clear how this definition was used in guiding the synthesis.

Response:
- Thank you for this comment. For the purposes of the review we do not want to judge definitions as ‘right’ or ‘wrong’ but to describe the definitions held within each of the approaches found. This was to allow readers to judge if and how these definitions resonate with their settings or needs. We have now included an explicit definition of sustainability on page 5 line 21-24. We believe this will aid readers in understanding our perspective on what sustainability meant to us while the review was conducted. We allowed this definition to provide a template for findings but did not let it limit the domains found as we looked for all definitions included in each approach.

“For the purposes of this work, ‘sustainability’ will refer to the general continuation and maintenance of a desirable feature of an initiative and its associated outcomes as well as the
process taken to adapt and developed in response to emerging needs of the system. These This definitions as well as any additional domains found will be explored in the review.”

2. In the Methods, under Search and Information sources, please indicate the search dates.

Response:
- Search dates have now been added to the manuscript on page 7 line 6-7.

3. In the paragraph Data collection process and study selection, it is stated that ‘20% of the full text articles for inclusion. Any differences in selected articles were discussed and inclusion and exclusion criteria were refined to reflect these discussions’. This is not a common approach for a systematic review to refine the inclusion/exclusion criteria along the way, so could the author provide more details about how this modified the selection criteria?

Response:
- Thank you for this question and we apologise for the lack of clarity. The strategy of conducting an initial search and modifying inclusion and exclusion was recommended by a medical librarian as ‘sustainability approaches’ did not have a concrete definition and required iterative discussions to determine what was considered an approach. This strategy allowed authors to discuss and refine what would qualify as a ‘sustainability approach’ and what would not. For example, early on we found a measurement tool which claimed to aid in sustainability of initiatives but after discussion this was excluded as its main objective was for measurement and not to primarily influence sustainability. (Mattke, 2008) This process allowed us to refine inclusion criteria to specify inclusion of only structured approaches which aimed to influence sustainability directly. The final inclusion criteria was then applied to all retrieved papers.

4. More details about the Quality Assessment should be provided. What tool was used? How was the quality of studies considered in interpreting the results?

Response:
- Thank you for this comment and your questions relating to the quality assessment conducted for the review.

This work focused on mapping current sustainability approaches, identifying knowledge gaps and understanding if there was consensus on sustainability constructs across healthcare fields.
From the outset, we did not want to make judgements on the approaches themselves or their constructs but aimed to provide readers with information on available methods, describing their proposed approaches and how they have been designed.

While we explored the use of existing quality assessment tools for our studies, we found they were not fit for purpose as they did not take into account the exploratory nature of the paper, with many of the criteria not seen as helpful for assessing articles introducing sustainability methods. For example, many of the papers identifying sustainability approaches did not address the tool as the topic of interest or research question. Approaches were frequently introduced as mechanisms to conduct the research, therefore the data collected within each study was not an indication of the soundness of the approach but of the individual study. For example in the study by Bray et al in 2009, the ‘sustainability pyramid model’ was developed before the study was conducted on the basis of the authors’ observations as participants in quality improvement (QI) collaboratives.

As we wanted to scope the literature for available methods, excluding articles based on inappropriate and unrelated research questions was considered counterproductive. We therefore chose to develop a quality assessment that responded to the needs of our article. We designed a data quality and extraction table that would allow us to capture the findings we believe were most relevant to readers. The data extraction form was developed to ensure the baseline quality expected— all articles had to clearly describe the proposed approach and give information about how it could be used in practice. Our data extraction form allowed us to understand if the included articles were providing sufficient information to readers. Here we list each criteria with the rationale for its inclusion and its benefit to readers:

1. Purpose of use (Why the tool has been designed),

2. Healthcare Setting (where has it been designed for?),

3. Level of Healthcare Use (What area of healthcare has it been designed for?),

4. Description of application (How should/can it be applied?),

5. Constructs identified (What has been deemed important to sustainability),

6. Scoring Mechanism (What data will be generated?),

7. Target user (Who can use the tool?),

8. Definition of Sustainability (How have the authors defined sustainability? Does this resonate with the readers views on sustainability?),
9. Theoretical underpinning (What underlying assumptions are there?)

10. Sustainability Perspective (Is sustainability being assessed or influenced?) and

11. Approach development details (How the method created? Do we think this is sound?)

Although we believe this approach has strengthened the relevancy and resonance for readers we also acknowledge that this may be seen as a limitation by others. Therefore we have addressed this within the main manuscript-acknowledging this limitation and making it clear to readers that assessment of the individual tools has not been done using an existing tool. The first change is in the methods section where we describe why an established quality assessment checklist was not deemed appropriate for our study on page 8 lines 2-19.

“A quality assessment and data extraction form was developed for identified articles. Existing quality assessments were explored but it has been noted that available quality assessment approaches often fail to consider the rationale and context of studies.\[22,23\] Their use to determine the inclusion of qualitative studies is often not recommended as many existing tools do not capture the multiple meanings of “good quality” and “rightness” therefore studies should often not be excluded based on this quality assessment.\[22,24\] The available assessments were not sensitive to the inclusion criteria of our study which was exploratory in nature. We sought to provide an overview of available approaches for sustainability and designed our data extraction form to identify and describe the included articles. The aim of the data extraction was to report descriptions and study information not to ascertain validity of the approaches or their constructs.”

We have also acknowledged this as a potential limitation in the discussion section on page 25 line 7-16.

“Another key limitation of this work is that we did not use an existing quality assessment tool and cannot attribute value or accuracy of constructs from each approach. While the quality criteria set out in our data extraction allowed us to ensure each paper had a minimum level of data to adequately describe the approach, it did not assess quality of the approaches themselves. We extracted information on each of the approaches which others may wish to use to attribute validity to findings. Details, particularly those around sustainability approach development, may be used by readers to assess whether they believe the approach has enough merit to be used in their site. It is important to note that many approaches (24%) were informed by professional expertise, a technique that may be difficult to assess for quality but appears to be very significant in the creation of sustainability approaches.”
This comment also raised another important question which we have now highlighted as future work. ‘How can these approaches be assessed for accuracy or soundness?’ There is currently no clear standard for assessing the quality of these approaches but we believe this can only be assessed by reviewing their use and application in future studies. This will ascertain whether they have included the correct constructs and if they are able to impact sustainability processes. We see identifying and describing the available approaches for sustainability, as a significant first step in aiding teams to understand what tools are available and describing how they can be applied in practice. Prior to this article these approaches remained dispersed across various fields with little evidence that they are being applied in practice. It is hoped this review will allow for a more structured approach to testing these tools in practice by providing a resource where available approaches are documented and their characteristics are described. Future work will explore the evidence of use of these approaches in practice which will help us to understand the true impact of these tools, unfortunately we cannot address it fully in this review.

We have now updated a sections within our discussion to highlight the need for assessment of these approaches to understand if and how they impact sustainability. Page 25 lines 18-25.

In the discussion: “Although consensus across approaches was found with 6 constructs included in over 75% of approaches we cannot be certain these are the ‘right’ constructs to assess sustainability. In order to understand the validity of these findings the approaches must be assessed in practice. Future work will explore if and how these approaches have been applied to ascertain if their constructs accurately represent sustainability in specific settings and if they fulfil their stated aims.”

5. In the Results section, the number of studies retained (62) is large and gathers very different types of literature, mostly frameworks and models, for which constructs could be identified. However, it is not clear how the sustainability constructs were identified for other types of methods (e.g. tools, checklists). Were the authors deducing the underlying constructs when these were not explicitly mentioned and if so, on which basis was this done?

Response:

- Thank you for your question. In order to identify constructs across approaches we preformed the same analysis for all included articles regardless of type.

As per our inclusion criteria, all approaches included in the review identified a structured approach to sustainability and contained sustainability constructs for assessment or reflection. This meant that although the format for presenting included constructs was different (some were represented as factors, others as questions or statements) all constructs were described in the articles themselves. Therefore we only recorded observed constructs and did not deduce unmentioned constructs.
In order to consolidate findings across approaches, we combining constructs that had different labels, heading or representations which were redundant or overlapping in definition. This also required us to parse and separate larger concepts into multiple construct codes. A simple example would be in Edwards et al.’s Checklist, a question such as ‘Is project perceived to add “value” (i.e., people can see something in it for them)?’ This was coded as multiple constructs - ‘Incentives’ and ‘Demonstrating effectiveness’.

6. The authors propose a Consolidated Framework for Sustainability Constructs based on all the different constructs identified in the literature, grouped under 6 themes. Was inter-rated reliability verified?

Response:
- Yes inter-rater reliability was verified. This has been described on page 14 line 20-23.

“To assess coding clarity and reliability a second coder independently coded 25% of the all included articles and the inter-rater reliability score (kappa Coefficient) was calculated between two coders using the NVivoPro coding comparison function. The test showed a high level of agreement between coders with an inter-rater reliability score of 0.94. These results were discussed and the small discrepancies between coders were used to refine codes and revise the definitions each of the constructs.”

7. On p. 4, Please correct: ‘Sustaining worthwhile changes poses a challenge to those undertaking an improvement initiative’.

Response:
- Thank you for highlighting this error. Text has been updated. Page 4 line 6.

8. On p.5 ‘lack of consensus on how to define sustain sustainability’

Response:
- Text corrected on page 5 line 3.

9. On p. 11, ‘The sustainability methods come in a variety of forms: Frameworks/Conceptual Frameworks (32), models (16), tools (8), guidance strategies (4) checklists (1), and Processes (1).’ Please use either capital or lowercase letters.
Response:

- Text corrected on page 12 line 24-26.