PEER REVIEW HISTORY

This paper was submitted to a another journal from BMJ but declined for publication following peer review. The authors addressed the reviewers' comments and submitted the revised paper to BMJ Open. The paper was subsequently accepted for publication at BMJ Open.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A systematic review of clinician-directed nudges in healthcare contexts
AUTHORS	Last, Briana; Buttenheim, Alison; Timon, Carter; Mitra, Nandita; Beidas, Rinad

VERSION 1 – REVIEW

REVIEWER	Yoong, Serene Swinburne University of Technology, John Street, Hawthorn Vic Australia
REVIEW RETURNED	25-Aug-2020

GENERAL COMMENTS	Thank you for the opportunity to provide comments on this interesting review. Overall, it was a very informative review and interesting read but the authors should consider using the PRISMA checklist to inform headings and information included in section of the manuscript to help increase readability. Important information that allow others to judge the quality of the review and risk of bias of included studies have not been discussed in the main manuscript. Additionally, there seems to be a lack of references throughout. I hope these comments are helpful and constructive in helping improve readability and interpretation of results. The introduction is interesting and well written - it would be useful to discuss briefly the availability of studies for synthesis that meet the aims of the review and recent reviews published in the field. Please reframe your research aims in line with PICO. Please use the PRISMA checklist to inform the headings of your manuscript if not already as this may help guide the information included in each section. Some information in the section (i.e. relating to not undertaking meta-analysis) should be moved to the analysis section. Search strategy: A very pragmatic search strategy was employed. How have the authors identified relevant reviews and commentaries? Was there a systematic process used to identify this? Was screening/extraction/risk of bias assessment undertaken in duplicate by two authors? Please provide more detail of your systematic search of 'several databases' briefly in text on top of inclusion in the appendix. Please provide more detail about how the authors developed search terms for the database as this is

	unclear. As nudge strategies do not self-identify as nudge, this is a challenge to the field so more description here will be helpful. Was there are grey literature search? There is no section detailing the type of data extracted - please add this to the manuscript. Information about risk of bias was included in the appendix - please include in the main text and discuss in the results. The interpretation of your study outcomes are also influenced by the quality of the conduct studies. The text under the heading analysis does not provide detail about the type of synthesis undertaken in this review rather it talks about data extraction and how nudge strategies are categorised, please add a section about analysis. Please provide justification for selection of taxonomy by Patel et al. Only 2407/2486 articles were retrieved- please provide a reason for this. Please include the study designs in the results summary. Please include how many studies assessed nudge alone and nudge with other more intensive strategies. Please describe the control arm of studies in a little more detail as often usual care or control arms may employ some strategies that could be classed as nudge. It would be useful to look at the impact of studies by nudge only and nudge with other intervention strategies to specifically answer the research question. " Many of the studies (n = 21) did not report how many clinicians interacted with the nudges." - I'm not sure what this means - can you rephrase so it is clear what is being addressed by interacting with nudge? The Cochrane Handbook advises author not to use vote counting approaches as it fails to take into account sample weighting -rather effect sizes or visual estimations be provided? Please consider whether the data synthesis approach used here is appropriate in this instance. I suggest the authors remove reference to statistical significance and consider specifying the effect sizes reported within individual studies. Please include references in the discussion where appropriate. There are a lot of sweeping statements without references so it is unclear whether these are the authors perceptions or an established understanding in the field (e.g. Moreover, people often see the default option as an information signal). The discussion is interesting but it would be useful to talk about its implications for practice and policy more explicitly (i.e. what do your findings mean for existing clinical practice and what changes should occur as a result). Please contrast your findings with other reviews in the setting and discuss whether they align or do not align ,and what this means. New findings are discussed in the future directions section (e.g. several studies found defaults backfiring) - please include in results as no new results should be presented in the discussion. There are broad statements about what future research should be undertaken - (e.g. Research should attend to how clinicians make inferences about the intentions and beliefs of the choice architect when interacting with nudges). Some suggestions about how this research could be undertaken will be helpful given this is not immediately obvious (e.g. should this be measured prior to undertaking the intervention and mediation analysis undertaken) Can you add a conclusion section?

	The limitation section needs to note limitation of the study (not the limitations of the literature being synthesised). For example, heterogenous outcomes is not a limitation of the review, rather the authors should consider noting the strengths/limitations of their search strategy, analysis methods, any subjectivity in coding according to taxonomy. While the authors mention publication and heterogeneity, no formal test (e.g. I2 or funnel plots) were undertaken to examine this, so I would not spend a lot of time on this. The length of discussion could be reduced to include important methods of undertaking the review in the main manuscript.

REVIEWER	Meeker, Daniella University of Southern California, Preventive Medicine
REVIEW RETURNED	25-Aug-2020

GENERAL COMMENTS	This is a well-written paper that applies existing taxonomies in a literature review of clinician nudges. The authors apply best practices in systematic reviews and base their methods on prior related research. The review gives excellent context, and results are clearly presented and discussed. Some clarifications would be helpful, as described below. The points made in the final paragraph should be emphasized throughout to ensure the reader is aware of the search strategy. Abstract: Objective - "A systematic review methodology was used to collect and consolidate results from current papers *citing* nudge strategies and to determine whether nudges are effective..." P5l44 - suggest a minor edit, something along the lines of “before health systems rush to embrace nudges.. *an understanding of how quality improvement interventions specifically designed with these theories in mind is warranted*” P6L5: This review endeavored to answer the following primary research question: Are interventions *inspired by nudges and related theories* effective in improving clinical decisions? P5L19 - Study Population. I suggest an explicit statement here indicating that QI studies that do not reference nudges or include related concepts in their titles are excluded. (If I have correctly interpreted) Methods - Search Strategy - If I understand correctly, there were two search methods. (1) Snowball Search (as described) and (2) Title Search (As in appendix A -- aka Supplement 1?). A short paragraph describing Appendix A/Supplement 1 would be helpful, as those details do not come across clearly. P9L24 - “*Interventions* on the bottom of the ladder tend to be more passive, offering decision makers information and reminders… *interventions* are more assertive and reduce the decision to a limited set of choices or by creating default options” p11L3 -typo (e.g., displaying clinicians’ pre-commitment *letter sin* their own examination rooms) P16L17 - Discussion. “*In addition to alert fatigue*, clinicians may experience nudge fatigue and begin to ignore decision support embedded in the EHR.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Thank you for the opportunity to provide comments on this interesting review. Overall, it was a very informative review and interesting read but the authors should consider using the PRISMA checklist to inform headings and information included in section of the manuscript to help increase readability. Important information that allow others to judge the quality of the review and risk of bias of included studies have not been discussed in the main manuscript. Additionally, there seems to be a lack of references throughout. I hope these comments are helpful and constructive in helping improve readability and interpretation of results.

1. The introduction is interesting and well written — it would be useful to discuss briefly the availability of studies for synthesis that meet the aims of the review and recent reviews published in the field.
    a. **We have now discussed other reviews published in the field and the availability of studies (p 7-9).**
2. Please reframe your research aims in line with PICO.
    a. **We have reframed our research aim in line with PICO guidelines (p. 9)**
3. Please use the PRISMA checklist to inform the headings of your manuscript if not already as this may help guide the information included in each section.
    a. **We now use the PRISMA checklist to inform our headings.**
4. Some information in the section (i.e. relating to not undertaking meta-analysis) should be moved to the analysis section.
    a. **We have moved our discussion of why we did not do a meta-analysis to the "Data Synthesis" section on p. 14-15.**
5. Search strategy: A very pragmatic search strategy was employed. How have the authors identified relevant reviews and commentaries? Was there a systematic process used to identify this?
    a. **We now describe the referred/convenience sampled article we used to generate articles. We also provide a reference to describe snowball sampling. The process was initially pragmatic, but then "snowballed" systematically (p. 15).**
6. Was screening/extraction/risk of bias assessment undertaken in duplicate by two authors? Please provide more detail of your systematic search of 'several databases' briefly in text on top of inclusion in the appendix.
    a. **We have provided more detail on the several databases we used as well as the way we conducted screening/extraction/risk of bias assessment (see Appendix A on p 61-65). For specific information about multiple coders see p. 12-13.**
7. Please provide more detail about how the authors developed search terms for the database as this is unclear. As nudge strategies do not self-identify as nudge, this is a challenge to the field so more description here will be helpful.
    a. **Search terms were discussed with a leading researcher applying nudges in healthcare contexts (the founding director of the first healthcare nudge unit at Penn Medicine), Dr. Mitesh Patel, with a medical librarian, and with an expert in applying behavioral economics to healthcare, Dr. Alison Buttenheim (as described on p. 60). That some nudge studies do not self-identify as nudges is a challenge, though we also included search terms such as "default," "opt-out," "pre-commitment," and coders reviewed this literature suggesting that our search was more**

expansive. As we describe in response to the Editor, we believe our review's focus on self-identified nudges is a strength because it enables us to evaluate nudges that are specifically guided by behavioral economic theory.

8. Was there are grey literature search?
    a. **We did not do a grey literature search.**
9. There is no section detailing the type of data extracted - please add this to the manuscript.
    a. **We now include this on p. 13.**
10. Information about risk of bias was included in the appendix - please include in the main text and discuss in the results. The interpretation of your study outcomes are also influenced by the quality of the conduct studies.
    a. **We have now included the risk of bias table in the main text (see Table 5) and interpret the study outcomes using the risk of bias tool (p. 17)**
11. The text under the heading analysis does not provide detail about the type of synthesis undertaken in this review rather it talks about data extraction and how nudge strategies are categorised, please add a section about analysis.
    a. **We have now added more detail on the type of synthesis undertaken (p.15)** 12. Please provide justification for selection of taxonomy by Patel et al.
    a. **We have now added more justification for the selection of the nudge taxonomy by Patel et al. Given that Patel's taxonomy is derived from an often-used public health taxonomy, we felt it had sufficient utility for the purposes of our work.**
13. Only 2407/2486 articles were retrieved- please provide a reason for this.
    a. **This was a typo that we have corrected.**
14. Please include the study designs in the results summary.
    a. **We will now include the study design in the results summary (p. 16).**
15. Please include how many studies assessed nudge alone and nudge with other more intensive strategies.
    a. **We have now delineated studies that assess nudges alone and nudges along with other strategies (see Tables 3 and 4).**
16. Please describe the control arm of studies in a little more detail as often usual care or control arms may employ some strategies that could be classed as nudge.
    a. **We have now included more detail about the control arms of the nudge studies (p. 16).**
17. It would be useful to look at the impact of studies by nudge only and nudge with other intervention strategies to specifically answer the research question.
    a. **Most nudge studies have some component of a non-nudge intervention embedded in the arm, which is why we suggest this is a limitation of the literature. For example, some studies that employed peer comparison letters educated providers about clinical guidelines in advance of sending the letters. This non-nudge was embedded within the nudge intervention. One way to directly study the effect of the nudge component of the intervention is to send placebo letters to the control arm with the clinical guidelines.**
    However, only one study that we reviewed did so. There is usually an educational component to all interventions we studied, which is why we mention it but cannot analyze/isolate the effects of the nudges.
18. "Many of the studies (n = 21) did not report how many clinicians interacted with the nudges." - I'm not sure what this means - can you rephrase so it is clear what is being addressed by interacting with nudge?
    a. **We have rephrased the sentence. In short, oftentimes the sample of clinicians was not reported (p. 16).**

19. The Cochrane Handbook advises author not to use vote counting approaches as it fails to take into account sample weighting -rather effect sizes or visual estimations be provided? Please consider whether the data synthesis approach used here is appropriate in this instance. I suggest the authors remove reference to statistical significance and consider specifying the effect sizes reported within individual studies.
    a. **The Cochrane Handbook suggests that a vote counting approach is acceptable when a meta-analysis is not feasible. Both systematic reviews of nudge strategies in health settings also employed a vote-counting approach. We cannot compare effect sizes given the significant heterogeneity in behavioral outcomes.**
20. Please include references in the discussion where appropriate. There are a lot of sweeping statements without references so it is unclear whether these are the authors perceptions or an established understanding in the field (e.g. Moreover, people often see the default option as an information signal).
    a. **We now include more references throughout.**
21. The discussion is interesting but it would be useful to talk about its implications for practice and policy more explicitly (i.e. what do your findings mean for existing clinical practice and what changes should occur as a result).
    a. **We now discuss the implications for clinical practice in the Future Research section (p. 2223).**
22. Please contrast your findings with other reviews in the setting and discuss whether they align or do not align, and what this means.
    a. **We now discuss how our findings relate to other reviews.**
23. New findings are discussed in the future directions section (e.g. several studies found defaults backfiring) - please include in results as no new results should be presented in the discussion.
    a. **The studies discussed in the section are not healthcare nudges and, therefore, were not eligible for the systematic review (p. 20).**
24. There are broad statements about what future research should be undertaken - (e.g. Research should attend to how clinicians make inferences about the intentions and beliefs of the choice architect when interacting with nudges). Some suggestions about how this research could be undertaken will be helpful given this is not immediately obvious (e.g. should this be measured prior to undertaking the intervention and mediation analysis undertaken)
    a. **We now make some suggestions for each future research item (p. 22-23).**
25. Can you add a conclusion section?
    a. **We now have a brief conclusion section (p. 23-24).**
26. The limitation section needs to note limitation of the study (not the limitations of the literature being synthesised). For example, heterogenous outcomes is not a limitation of the review, rather the authors should consider noting the strengths/limitations of their search strategy, analysis methods, any subjectivity in coding according to taxonomy.
    a. **The limitations section (p. 21-22) now outlines the limitations of our study. We have kept some of the limitations we initially included given that the limitations of the literature do have some bearing on the review. If studies are heterogeneous, for example, it hinders our ability to definitively answer the research question.**
27. While the authors mention publication and heterogeneity, no formal test (e.g. I2 or funnel plots) were undertaken to examine this, so I would not spend a lot of time on this.
    a. **While we mention heterogeneity, we spend less time on this in our discussion.**

28. The length of discussion could be reduced to include important methods of undertaking the review in the main manuscript.
    a. **We have now reduced the discussion.**

Reviewer: 2

This is a well-written paper that applies existing taxonomies in a literature review of clinician nudges. The authors apply best practices in systematic reviews and base their methods on prior related research. The review gives excellent context, and results are clearly presented and discussed. Some clarifications would be helpful, as described below. The points made in final paragraph should be emphasized throughout to ensure the reader is aware of the search strategy.

1. Abstract: Objective - "A systematic review methodology was used to collect and consolidate results from current papers *citing* nudge strategies and to determine whether nudges are effective..."
    a. **We have revised the objective to make it clear.**
2. P5l44 - suggest a minor edit, something along the lines of “before health systems rush to embrace nudges.. *an understanding of how quality improvement interventions specifically designed with these theories in mind is warranted*”
    a. **We have incorporated this feedback.**
3. P6L5: This review endeavored to answer the following primary research question: Are interventions *inspired by nudges and related theories* effective in improving clinical decisions?
    a. **We have now clarified the primary research question to make it inclusive of behavioral economics theory.**
4. P5L19 - Study Population. I suggest explicit statement here indicating that QI studies that do not reference nudges or include related concepts in their titles are excluded. (If I have correctly interpreted)
    a. **Thank you. We have now clarified study inclusion (p. 10).**
5. Methods - Search Strategy - If I understand correctly, there were two search methods. (1) Snowball Search (as described) and (2) Title Search (As in appendix A -- aka Supplement 1?). A short paragraph describing Appendix A/Supplement 1 would be helpful, as those details do not come across clearly.
    a. **We have clarified our methods to describe the two search methods (p. 11-12).**
6. P9L24 - “*Interventions* on the bottom of the ladder tend to be more passive, offering decision makers information and reminders… *interventions* are more assertive and reduce the decision to a limited set of choices or by creating default options”
    a. **Thank you. We have incorporated the reviewer’s suggestion.**
7. p11L3 -typo (e.g., displaying clinicians’ pre-commitment *letter sin* their own examination rooms) **a. The typo is fixed.**
8. P16L17 - Discussion. “*In addition to alert fatigue*, clinicians may experience nudge fatigue and begin to ignore decision support embedded in the EHR.
    a. **We have added this modifier.**