PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Using the Theoretical Domains Framework and the Behaviour Change Wheel in an Overarching Synthesis of Systematic Reviews
AUTHORS	Richardson, Michelle; Khouja, Claire; Sutcliffe, Katy; Thomas, James

VERSION 1 - REVIEW

REVIEWER	Associate Professor Gillian S Gould The University of Newcastle, Australia Gillian S Gould is principally funded by NHMRC Australia, Cancer Institute New South Wales, Cancer Australia, Cure Cancer Australia, and NSW Ministry of Health. She collaborates with and has authored papers with Lou Atkins and Robert West from University College London; her PhD was supervised by Andy McEwen UCL, and she has attended courses at UCL, which may be competing interests.
REVIEW RETURNED	11-Sep-2018

GENERAL COMMENTS	See attached. I think this is a challenging paper to write as a worked example, but there could have been a more critical assessment of the use of TDF and BCW as an overview and more transparency about who did what and how in the team. It may be easier to write just as a straight overview, and make some reference to its novel aspects using this method. More details about the included reviews would have useful as context. BCW/TDF overview comments 1. Line 16 not certain what is meant by reported elsewhere – was this overview already published before – is this a secondary analysis then? 2. The first paragraph should not outline the methodology but give an intro to the problem? Is this about the methods of synthesis or about the topic of minor illness is unclear. If the former than let's hear about the state of the art of overviews of SRs and how they are usually synthesized in more detail, if latter we should start with 2nd paragraph 3. Line 31 reference does not follow numbering (Foot) 4. Table 1 should not be in the introduction if it is a result of the study 5. Aims: 'bring together the qual and quant findings'…..please say for what purpose – to inform practice, policy etc? Or just proof that this can be done?

	6. Methods – how were the reviews selected and why were these chosen? Where there any others that could have been included – was any sort of systematic search conducted? We need to know this in order to decide what sort of biases are operating here. The referencing makes it unclear which citations refer to the reviews that were actually included. Did the authors choose these reviews because they were their own work, or were they open to finding any other reviews? Or did they search and not find any – if so what data bases were used and what date limits – who did the search etc. This needs to be transparent – ie search terms etc. 7. Was there any quality rating of the included reviews (eg AMSTAR) – this should be done as if the reviews were those of the authors (unclear) then this introduces bias – how was this accounted for and what remedies were taken to account for bias – perhaps quality rating could have been done independently by a different team? 8. Page 6 who did the work of mapping? Which authors? Was there any independence of coding? Line 18 – what were the other online sources? They should be referenced and accessible. 9. P7 unclear who the stakeholders were and their role – how were they selected? 10. Most of the results are on the findings related to MA and the synthesis not on the findings about the use of the methodology – which is probably the aim of the study, is it not? 11. Page 8 – this is the first mention of the statistical significance of interventions and this being assessed – should be in the methods. Unclear from methods how the interventions were going to be synthesized for effect by the TDF/BCW. 12. Page 9 paragraph starting at line 10 – if these are a comprehensive frameworks why couldn't these be mapped onto them? Wouldn't the 'best place for the problem' be related to environment? Failed self-care – beliefs about capabilities? It would be good to know whether the people doing the analysis had prior expertise in the TDF – so when they say not clear about mapping onto TDF – is this a problem with lack of skill of researchers or a deficit in the model? 13. Table 3 – I do not see reference to BCW in this table only mapped to TDF. Clarify in title 14. P 11 line 45 error message 15. P 55/56 the statement that educational interventions are ineffective is unclear whether this is a general statement (if so reference) or pertinent to the findings of this review . Is symptom diagnosis a physical skill – I would have thought it was psychological. Surely it depends what it is. 16. Table 4 "persuade service-user from being overwhelmed by anxiety" is not clear how this could be achieved. "Pressure" is not a recognized part of the BCW – OK I see the note you mean coercion. It doesn't sound very friendly – could induce stigma/shame and make people feel bad or foolish. These suggestions should have more critical thought, as I don't think they are appropriate especially for vulnerable or high-priority populations. If I have missed the point here then probably not clear about what authors are trying to achieve. 17. I may have missed it but should be noted in the methods or introduction that the TDF and BCW have previously been mapped to each other. 18. P16 line 14 – I would be wary of claiming these as causal determinants as this would need sophisticated methodology. Probably best to say associations – we know behaviour is a 2-way street.

	19. Discussion should start with a summary of main findings. Again is this paper about the methodology or about MA as a topic? It seems like here you are discussing the challenges of the methodology. Isn't this then part of the results rather than the discussion? Line 39-40 some things didn't quite fit – but not clear why and this should be more critically analysed – is it a flaw in the theory or its application to this task? Line 50-51 'should explore data carefully' but these authors did and still wasn't clear. A comment needs to be made on the suitability for use of the method. 20. P17 line 5 'cross-triangulated' more could have been made of this earlier in the article as a reason for the study. Good to look at the different types of triangulation and be transparent about who did what when and why. The researchers should be indicated by initial at least for each part in the methods and what degree of independence that did and how consensus was reached ie 'researcher triangulation' then you have data triangulation etc. 21. P17 line 22-48 should come at the beginning of the discussion. 22. Line 52-53 this statement comes rather late and should be substantiated early (no previous reviews) 23. Strengths and weaknesses – a comment made here that this study was on their own reviews, if this is the case? Did any of the included papers in the reviews overlap? 24. What will be done with the findings 25. Going back to the abstract I think the difficulties with the methodology should have been indicated in the abstract or the article summary, and a more critical analysis of the findings related to the aim rather than the worked example

REVIEWER	Sarah Denford University of Exeter
REVIEW RETURNED	18-Oct-2018

GENERAL COMMENTS	overall, I think the manuscript is an important and well written contribution to the literature. I have a few minor questions / issues for the authors. Background: Clearly describes the research aim and relevant literature. Methods: I was a bit confused as to where the three included reviews came from. You state in the introduction that there is only one published SR of self-care for MA; yet you apparently include three SRs of SC for MA in the present study. It would also be helpful to describe how you identified these reviews, how they were chosen etc. I appreciate that this is not the aim of your manuscript; however, you do present findings in a way that suggests that you can utilize the data to inform future interventions. However, there is no consideration of how these reviews were identified or selected, or whether any other similar reviews exist. It would be useful to know something about the aims, inclusion criteria, quality etc of the included studies. It also appears that two of the reviews were synthesized using the TDF (Table 1). Would the approach used still work for reviews that do not use the TDF when synthesizing the data? You mention that you amend the labels for the themes and sub themes for consistency. Did you check that the original authors were happy with your modified labels?

	Results: Was any information relating to the content of "educational" interventions? Knowledge of services (which didn't seem to be an issue for participants who completed one of the survey studies) is quite different to knowledge relating to whether or not a symptom needs urgent treatment. This appears to be the bigger issue for participants in the interview studies. I was unsure why "previous experience" is considered reinforcement? Could do with some clarification. I am also unclear why delayed prescribing was mapped onto reinforcement. To me, this is more about managing emotions (I feel better knowing that I have / can take my prescription if I need to). Discussion: Some discussion of the weighting given to the various types of evidence would be useful. You considered themes to be important if there was consistency, or if they were salient in survey respondents. Was this also the case for the qualitative data? Minor points: Error message on page 11.

REVIEWER	Colin Greaves University of Birmingham, UK
REVIEW RETURNED	24-Oct-2018

GENERAL COMMENTS	GENERAL COMMENTS This is an interesting methodological paper and I support its publication in BMJ Open. However, there at a number of mainly presentational issues that need to be addressed ... ABSTRACT The notation "n" in the abstract seems to refer to number of studies - standard notation in reviews is to use "k" for number of studies and "n" for number of participants. For clarity /readability by a wide audience it might be better in any case to say "(review 1, 20 studies)" Not clear what is meant by "Salient TDF domains were then integrated into BCW" - quite hard to follow for non-specialists in particular ARTICLE SUMMARY " ...allowing for specific recommendations to be made and tested." Suggest to remove "and tested" as no such testing was performed in the current study INTRO International readers may not understand what "NHS111" is /can you offer a few brief words of explanation? In general, the Intro is lacking in references: e.g. what is the evidence that minor ailments "often place an unnecessary strain on these overstretched services."? Is Table 1 cited correctly in the text (the Table doesn't seem to match what the text is saying) - maybe need to swap Table 1 and Table 2?

	It would be helpful to include a diagram of the BCW - if you don't already know what BCW is, the text alone does not describe sufficiently what is intended here "The COM-B system forms the hub of the BCW7 and, in conjunction with the next layer of the BCW, can be used to identify potentially relevant intervention functions, based on the salient TDF and COM-B domains.". A diagram would help to understand these complex concepts. In the Introduction or Discussion, you could refer to other methods used to synthesise data from different methodologies, particularly in the context of informing intervention design (e.g. triangulation protocol [1], Framework synthesis). Also, needs some text at the start to set up the problem /the rationale for doing this study ... for example along the lines of... In developing complex (especially behavioural) interventions, we often need to synthesise data of multiple types to identify barriers to change and potentially effective intervention components to overcome these barriers, as well as theoretical change process. This may mean synthesising data from both qualitative and quantitative systematic reviews, evidence based guidelines, individual trials, surveys and other studies. Doing this in a way that filters the evidence in a way that can inform choice of intervention components is desirable, but as yet, no clearly defined methods to achieve this complex task. THEN outline one or two existing approaches (but point out that the aforementioned process of "strategic filtering" /purposeful synthesis is not included. THIS IS THE KEY VALUE OF THIS PAPER, so worth flagging this up at the outset? METHODS It is not clear how "findings from the evaluations in review 3 were mapped onto the TDF and cross-tabulated, where possible, with the findings from the interview and survey studies". Maybe referring to one of the tables or using a new table to provide an example of how this mapping process might work - to be maximally useful and promote uptake of this innovative method, the paper needs to demonstrate HOW to perform this complex form of synthesis. I understand that this is challenging, but I'm not sure I could replicate this process from the current text. RESULTS I expected to see worked examples of how the various mapping processes worked as well as summary tables of the mapping (even if supplementary files). As it stands, there is something of a gap in the narrative of the paper - we go from a broad summary of the methods to summary level Results. There seems to be an explanatory /illustrative step missing here to complete the audit trail - how you we get from the data to the Results? I feel this is really important, so that others can replicate the methodology in future studies. It would be good to reserve judgement /discursive text for the Discussion (e.g. "This is surprising as the participants in the qualitative review requested information on symptom management ..." In general the Results section lacks structure and comes across as a list of observations (albeit gathered under two main headings). (As above) I cant see how the Results map onto the

	Methods described, or how the synthesis worked. Some of it is clearly a narrative synthesis of the findings, but this seems to be based on a more detailed mapping process which is not illustrated. I realise that this sounds rather negative, but I am actually very supportive of publication - I think this is an innovative and interesting study, but it needs to be written and structured much more clearly to make it a high quality paper. The standard to aim for (as with all science) is to ensure that the Methods that have been used /analyses applied are replicable by people reading the article. Example tables showing the mapping of data source /data onto elements of the TDF would help to illustrate how the analysis led to the findings. Supplementary data files could also be used to present more of (or the whole of) the mapping process? It may also help to provide Table 3 at the start of the section to provide the reader with a structure for processing the text from each domain. NB: Is the word "barriers" the right word to head the column on Table 3 - maybe it is more about "influences" or determinants as the influences can be both positive or negative (e.g. both barriers and enablers are referred to elsewhere in the article)? P11 line 51-53: Grammar error? "... data from the views studies, highlighted the importance of ..." Similarly, Table 4 could come at the start of the 'Step 2' section to ground the reader in what is being discussed (in what is otherwise a difficult read for the non-specialist reader). Again "barriers" might be better phrased as "influences" or "determinants". As above, a figure outlining what the BCW is seems necessary here for de-coding the table /the findings DISCUSSION The sub-heading structure seems appropriate (e.g. Summary of the principal methodological findings). However, the text included under the first heading does not always seem to fit the heading - this could be shortened considerably - what are the key points you are trying to make here? Some very long sentences here that need breaking down /simplifying. E.g. "The ‘limited roles’ sub-theme (categorised within the Environmental context and resources domain) might also have been coded within the Social influences (professional role and identity) TDF domain, but after extensive deliberation and some uncertainty about the definition (‘A coherent set of behaviours and displayed personal qualities of an individual in a social or work setting’) the broader domain of Environmental context and resources was chosen (including any circumstance of a person’s situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behaviour)." I have read this 3 times and still don't get what it is trying to say. Strengths and limitations: Again, the text is quite long. Instead of saying "In the absence of evaluation data, themes and sub-themes were identified as

	important based on concordance of findings across both the views and surveys, where at least a quarter of survey participants identified the theme as important. The threshold of 25%, used to determine the relevance of themes in the survey data, was arbitrary.", why not just say "The threshold of 25%, used to determine the relevance of themes in the survey data, was arbitrary."? Other limitations should include the limited number and diversity of studies in each review and the lack of any evidence from trials targeting the specific influences /determinants identified. In addition, given that the analysis here is complex (and primarily a narrative synthesis) are there any limitations in relation to what level of confidence we can place in the findings - how do we know if the data is "saturated" or if the inferences being drawn are based on sufficiently powered statistical analyses? - how do we know how robust the findings of this type of data synthesis might be? For example, in quantitative review studies you could do a statistical power calculation, or run a publication bias test (for reviews), in qualitative studies you could use a quality assessment framework (e.g. Lincoln & Guba), but here there are (as yet) no clear criteria to allow researchers to judge the relative strength or weaknesses /limitations of the findings. Further work may be needed to define criteria delimiting the requirements for trustworthiness /confidence is outputs when using this type of synthesis? CONCLUSIONS These are nicely put. However, I wasn't sure about the last sentence. Possibly "The theoretical scaffold provides a means to COLLATE /SYNTHESISE the evidence" - but, I don't know what you mean by "giving longevity to the research" and cant see how the evidence presented "facilitates the generalisation of the findings to similar contexts." Maybe it is just a matter of phrasing, but please consider clarifying or editing this sentence. 1. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ. 2010;341:c4587.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Associate Professor Gillian S Gould

Institution and Country: The University of Newcastle, Australia

Please state any competing interests or state ‘None declared’: Gillian S Gould is principally funded by NHMRC Australia, Cancer Institute New South Wales, Cancer Australia, Cure Cancer Australia, and NSW Ministry of Health. She collaborates with and has authored papers with Lou Atkins and Robert West from University College London; her PhD was supervised by Andy McEwen UCL, and she has attended courses at UCL, which may be competing interests.

Please leave your comments for the authors below

See attached. I think this is a challenging paper to write as a worked example, but there could have been a more critical assessment of the use of TDF and BCW as an overview and more transparency about who did what and how in the team. It may be easier to write just as a straight overview, and make some reference to its novel aspects using this method. More details about the included reviews would have useful as context.

Please see attached file.

BCW/TDF overview comments

1. Line 16 not certain what is meant by reported elsewhere – was this overview already

published before – is this a secondary analysis then?

This is a revised version of one of four reviews funded by the Department of Health and Social Care and published on the EPPI Centre website [https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3728]. The submitted manuscript focuses on the methods used in the overarching synthesis of the three systematic reviews. It has evolved a lot since the original publication, including explicit mapping of the BCTs in review 3 and a focus on the synthesis methods (rather than the substantive findings). It has its own introduction, method, results and discussion sections that provide the background and rationale to the methods work and illustrate how the methods were applied.

In the revised manuscript, we have included the following statement, in the second paragraph of the introduction, to clarify the origin of this work:

Here we report an overarching synthesis of three interconnected systematic reviews undertaken by our team, including: syntheses of service-user views in interviews (review 1) and surveys (review 2), and evaluations (review 3) of a range of interventions and services, using the Theoretical Domains Framework (TDF)1 and the Behaviour Change Wheel (BCW)2. These reviews, summarised in Table 1 below (and reported in full elsewhere)a, explore the self-care of minor ailments (MAs) in the UK3.

Footnote (a) clarifies that this paper reports an extended version of the overarching synthesis (review 4) that focuses in-depth on the methods that were developed for that review.

2. The first paragraph should not outline the methodology but give an intro to the

problem? Is this about the methods of synthesis or about the topic of minor illness is

unclear. If the former than let's hear about the state of the art of overviews of SRs

and how they are usually synthesized in more detail, if latter we should start with 2nd

paragraph

Thank you. We agree that the methods and substantive findings were confused in the first submitted draft. We have reframed the paper to focus on the methods aspect of the work and clarified the objective accordingly. Specifically, underneath the heading 'objective', the following is stated:

Synthesis that can filter the evidence from multiple sources to inform choice of intervention components is highly desirable yet, at present, there are no clearly defined methods. Here, we demonstrate how using the Theoretical Domains Framework (TDF) and the Behaviour Change Wheel (BCW) made it possible to bring together the findings from a series of three interconnected systematic reviews on the self-care of minor ailments (MAs) to inform the choice of intervention components.

3. Line 31 reference does not follow numbering (Foot)

Thank you. We have checked the reference formatting of the revised manuscript.

4. Table 1 should not be in the introduction if it is a result of the study

Table 1 provides an overview of the reviews that were synthesised in the overarching synthesis and, therefore, are not the findings of the current manuscript. We have added text to explain Table 1 and believe that it provides the necessary context and structure for the reader.

Specifically, it is noted that:

Here we report an overarching synthesis of three interconnected systematic reviews undertaken by our team, including: syntheses of service-user views in interviews (review 1) and surveys (review 2), and evaluations (review 3) of a range of interventions and services, using the Theoretical Domains Framework (TDF)1 and the Behaviour Change Wheel (BCW)2. These reviews, summarised in Table 1 below (and reported in full elsewhere) , explore the self-care of minor ailments (MAs) in the UK3.

5. Aims: 'bring together the qual and quant findings'…..please say for what purpose –

to inform practice, policy etc? Or just proof that this can be done?

Thank you. We have now clarified that the purpose is to bring together qualitative and quantitative findings to inform intervention design. Specifically, the following aim is stated:

This research sought to apply the TDF, the COM-B system of behaviour change, and associated BCW, as tools for bringing together the quantitative and qualitative findings from three systematic reviews on self-care for MAs, to inform the choice of intervention components.

6. Methods – how were the reviews selected and why were these chosen? Where there any others that could have been included – was any sort of systematic search conducted? We need to know this in order to decide what sort of biases are operating here. The referencing makes it unclear which citations refer to the reviews that were actually included. Did the authors choose these reviews because they were their own work, or were they open to finding any other reviews? Or did they search and not find any – if so what data bases were used and what date limits – who did the search etc. This needs to be transparent – ie search terms etc.

This manuscript is a revised version of one of four interconnected systematic reviews (three individual systematic reviews and one overarching synthesis bringing the three reviews together) conducted by ourselves, funded by the Department of Health and Social Care, and published on the EPPI Centre website [https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3728]. The submitted manuscript focuses on the methods used in the overarching synthesis of the three individual systematic reviews. The reviews were conducted because a comprehensive search of the literature (conducted by an experienced information specialist) indicated that only one published systematic review had examined self-care for minor ailments in the UK. In this review by Paudyal et al. (2013), 31 UK pharmacy-based minor-ailment schemes (PMAS) were examined. For this reason, we did not include pharmacy services in our synthesis, but refer to this comprehensive work in our review.

We have now clarified the origin of the reviews and highlighted the single existing review in the literature that focused on community pharmacy. Specifically, the following is noted at the start of the Methods section, underneath the sub-heading 'Reviews included in the overarching synthesis':

We conducted three syntheses of service-user views in interviews (review 1, 20 studies) and surveys (review 2, 13 studies) that sought to explore the factors that may influence self-care for MAs, and evaluations (review 3, 21 studies)3 of the effectiveness of behavioural interventions and services that support self-care for MAs. These formed the basis of the overarching synthesis reported here. These reviews filled an evidence gap identified by a comprehensive search of the literature in 2015, restricted to 2000 onwards (see Richardson et al. 3 for full details), which identified only one systematic review that had examined self-care for MAs in the UK. In this review by Paudyal et al.4, 31 UK pharmacy-based MA schemes were synthesised.

Readers are directed to the EPPI Centre report for the full details due to the recommended journal space limitations.

7. Was there any quality rating of the included reviews (eg AMSTAR) – this should be done as if the reviews were those of the authors (unclear) then this introduces bias –

how was this accounted for and what remedies were taken to account for bias –

perhaps quality rating could have been done independently by a different team?

Thank you for this interesting comment. We didn't evaluate the quality of the reviews as they were our own. We have included this as a limitation in the discussion section (underneath the sub-heading 'Strengths and weaknesses in relation to other studies') and noted that criteria to assess the quality of outputs when using this type of synthesis would be beneficial, but to our knowledge, are not currently available in the literature. We have included the quality appraisal of the primary studies in Table 1.

8. Page 6 who did the work of mapping? Which authors? Was there any independence

of coding? Line 18 – what were the other online sources? They should be referenced

and accessible.

Thank you. We have now re-written the methods section (see page 6) including three new headings:

How interventions were mapped (step 1)

Triangulation of findings using the TDF and COM-B system of the BCW (step 2)

BCTs and intervention strategies (step 3)

Michelle Richardson completed the coding which was then checked by Claire Khouja and by another reviewer (Katy Sutcliffe) in the few instances where there were disagreements.

9. P7 unclear who the stakeholders were and their role – how were they selected?

Thank you. The stakeholder group was organised by colleagues at the Department of Health and Social Care. Since we were unable to contact all of the stakeholders to obtain their permission to report their name in the manuscript, we have chosen to detail the organisations that they represented.

The following section is included under the sub-heading: Patient and Public Involvement

Stakeholder involvement has been an important aspect of this project, from the early planning stages through to analysis and write-up of the study findings. Stakeholders (n=5) (including representatives from the Department of Health and Social Care (DHSC) policy teams, the DHSC Policy Research Programme, NHS England, and the Economic and Social Research Council) provided feedback on the study protocol and helped to inform the scope of the research topics and research questions. They also met to review the study findings and to prioritise the behaviour change approaches for the self-care of MAs, drawing on the APEASE criteria. Whilst patients and the public were not involved in

the design or conduct of the review, patient views were central to the review and its findings as they were the focus of the analyses of service-user views (reviews 1 and 2). There were no participants as this work comprises a synthesis of reviews.

10. Most of the results are on the findings related to MA and the synthesis not on the

findings about the use of the methodology – which is probably the aim of the study,

is it not?

Thank you.

The results section has now been revised to reflect the methods focus of the paper and includes three main sections that map onto the methods (see, page 9):

- How interventions were mapped: identification of BCTs and TDF domains (step 1)
- Triangulation of findings using the TDF and COM-B system of the BCW (step 2)
- BCTs and intervention strategies (step 3)

In addition to mapping the presence or absence of determinants characterised using the TDF, Table 3 now shows the mapping of the determinants onto the COM-B model and highlights where the interventions (identified in review 3) target these.

A new column has been added to Table 4 entitled 'existing interventions' (to highlight if and how the determinant(s) are currently targeted by the interventions in review 3).

Supplementary tables (s1 to s3) have also been included to outline the mapping process not reported in the manuscript. Supplementary table 1 details the studies in the synthesis of evaluations (including information on behaviour, target, context, intervention provider and BCTs); Supplementary table 2 reports the BCTs identified in the interventions (with examples) and mapped onto the theoretical domains framework (TDF); and Supplementary Table 3 describes the choice of intervention functions.

11. Page 8 – this is the first mention of the statistical significance of interventions and

this being assessed – should be in the methods. Unclear from methods how the

interventions were going to be synthesized for effect by the TDF/BCW.

Thank you. We have now revised the methods section to clarify these issues. Specifically, the following is noted underneath the heading 'Triangulation of findings using the TDF and COM-B system of the BCW (step 2)':

Statistical information (direction of effect sizes and associated p-values) of the interventions is discussed, where relevant, and reported in supplementary Table 1 (see Richardson et al. (2018)5 for full details, such as effect sizes and confidence intervals).

We had planned to conduct meta-analysis, where feasible, but there were insufficient robust data for meta-analyses. Data were, therefore, synthesised narratively. The synthesis methods for review 3 are reported in the manuscript (see Table 1 which provides a summary of all three reviews).

Although not the focus of this paper, we have added the effectiveness data (in terms of direction of effect and whether statistically significant) to supplementary Table 1, which characterises the intervention studies. We also direct readers to the original review for more details, such as effect sizes and confidence intervals, and have added the following statement to the strengths and limitations section:

The BCT coding in review 3 was hampered by the quality of the intervention descriptions and it was not possible to statistically analyse the effectiveness of individual BCTs or of different combinations.

12. Page 9 paragraph starting at line 10 – if these are a comprehensive frameworks why couldn't these be mapped onto them? Wouldn't the 'best place for the problem' be related to environment? Failed self-care – beliefs about capabilities? It would be good to know whether the people doing the analysis had prior expertise in the TDF – so when they say not clear about mapping onto TDF – is this a problem with lack of skill of researchers or a deficit in the model?

Thank you. We have re-written the methods section and clarified explicitly the coding methods adopted (see pages 6 and 7). Michelle Richardson has experience in using the TDF, and all the reviewers have experience in using a priori frameworks to code data inductively. On reflection we agree that the determinants (previously coded as miscellaneous) could be mapped onto the TDF.

Most of them reflected peoples understanding of the qualifications and experience of various health providers and beliefs about the consequences of seeing a health provider and, therefore, have been amalgamated within the 'qualifications and experience' code (including confidence in service/self-care, failed self-care, best place, prefer to see GP, and expected advice only).

The remaining items originally coded as miscellaneous have been mapped onto the 'appropriate use of health services' code within the Social influences category (including did not want to bother GP/wanted to see nurse, and wanted a second opinion from pharmacy).

13. Table 3 – I do not see reference to BCW in this table only mapped to TDF. Clarify in Title

Thank you. Both the TDF and COM-B are included in the revised title of Table 3.

14. P 11 line 45 error message

Thank you. We have removed this typo.

15. P 55/56 the statement that educational interventions are ineffective is unclear whether this is a general statement (if so reference) or pertinent to the findings of this review . Is symptom diagnosis a physical skill – I would have thought it was psychological. Surely it depends what it is.

Thank you. We were referring to the findings of review 3 in the syntheses. We have now clarified this by adding review 3 in brackets after the statement.

We suggest that the management of minor ailments (which encompasses a cluster of behaviours) requires both psychological and physical capability. The point we were making is that the existing educational interventions (generally ineffective, review 3) primarily target the acquisition of knowledge or theory, indicating that psychological capability alone, is insufficient to promote self-care. According to the BCW, physical capability is targeted when skill development is emphasised, for example, the use of a symptom diary to support self-diagnosis and symptom management.

16. Table 4 "persuade service-user from being overwhelmed by anxiety" is not clear how this could be achieved. "Pressure" is not a recognized part of the BCW – OK I see the note you mean coercion. It doesn't sound very friendly – could induce stigma/shame and make people feel bad or foolish. These suggestions should have more critical thought, as I don't think they are appropriate especially for vulnerable or highpriority populations. If I have missed the point here then probably not clear about

what authors are trying to achieve.

Thank you. We had changed the wording from coercion to pressure considering feedback from the stakeholder group. On reflection, we agree that this intervention function is inappropriate (because of the reasons you outline above) and have, therefore, removed it from the suggested list of strategies. This concurs with the feedback from the stakeholder group that coercion was an inappropriate strategy.

17. I may have missed it but should be noted in the methods or introduction that the

TDF and BCW have previously been mapped to each other.

Thank you we have now revised the methods section and include references to previous mappings. Specifically, it is stated that:

We drew on the links between the TDF and COM-B, identified by a group of experts in a consensus exercise (shown in figure 1 in the Behaviour Change Wheel guide6). (step 2)

BCTs linked to the intervention functions were identified using BCTv17 and the results from an expert consensus exercise that mapped BCTs onto intervention functions (see table 3.3 in the Behaviour Change Wheel guide). (step 3)

18. P16 line 14 – I would be wary of claiming these as causal determinants as this would

need sophisticated methodology. Probably best to say associations – we know

behaviour is a 2-way street.

Agreed. We have removed the 'causal' prefix to the word determinants. We have also added the following text to highlight this issue as a limitation:

Without longitudinal modelling studies or intervention designs it is difficult to establish which determinants are most important and whether the links between the theoretical assessments and BCTs are valid.

19. Discussion should start with a summary of main findings. Again is this paper about

the methodology or about MA as a topic? It seems like here you are discussing the

challenges of the methodology. Isn't this then part of the results rather than the

discussion? Line 39-40 some things didn't quite fit – but not clear why and this

should be more critically analysed – is it a flaw in the theory or its application to this

task? Line 50-51 'should explore data carefully' but these authors did and still wasn't

clear. A comment needs to be made on the suitability for use of the method.

The discussion section has been revised in light of the major amendments made to the methods and results sections (see p23).

In the original draft of this manuscript, most of the conceptual issues related to the synthesis in the primary reviews (reviews 1 and 2). The revised version of the manuscript is focused on the overarching synthesis, where there were fewer issues.

Nonetheless, as noted previously in response #12, the determinants previously coded as miscellaneous in the surveys (review 2) were successfully mapped onto the TDF when revisiting the data.

A few of the other determinants were re-coded on revisiting the data. These amendments are denoted in footnote e:

Some determinants identified in the interviews and surveys originally mapped onto environmental context and resources TDF domain were re-coded onto Memory, attention and decision making (information overload) or beliefs about consequences (conflicts of interest and treatment expectations)..

This highlights how the mapping process is supported by an iterative rather than sequential approach to analyses. For example, in some cases, it was through considering the COM-B domains, that it became evident which TDF domain was appropriate.

This has now been clarified in the discussion section:

The syntheses required considerable time which was supported by an iterative rather than sequential approach to analyses. In applying this method, the researcher should, therefore, expect to revisit and revise their coding and understanding of the topic several times, as knowledge evolves and emerges throughout the process of synthesis.

Given that these tools are generic rather than tailored to a specific context our understanding is that modifications are welcomed to improve applicability to a particular context. Given the infancy of the

approach, it's important to document potential problems of application so that this information can be fed into subsequent development of the tools and guidance. This has been clarified in the discussion section:

Users, therefore, should expect to apply the framework with some flexibility, as acknowledged by its creators. For example, combining or mapping BCTs onto multiple domains, where there is considerable overlap; adding domains or BCTs, where the model does not account for them; and documenting any such amendments to support the ongoing development and refinement of this method.

This issue also feeds into the problem of assessing the quality of this type of synthesis (raised in #7 above and by reviewer 3). We have, therefore, also noted the following in the strength and limitation section of the discussion:

Criteria to assess the quality of outputs when using this type of synthesis would be beneficial, but to our knowledge, are not currently available.

In terms of the suitability of the methods, it seems likely that this approach is most useful when the TDF/BCW/BCT taxonomy are used to explore a behavioural problem within separate syntheses, before being brought together. Whether this approach can be used among reviews that did not use these tools to originally synthesise their data is not known, but is likely to be influenced by the complexity of the behaviour, and the transferability of the findings onto the TDF/BCW system.

We have added this point in the strengths and limitations section of the discussion:

Both of the views reviews (interviews and surveys) were already synthesised using the TDF, whether this approach to synthesis would work if other models had been used is unclear.

20. P17 line 5 'cross-triangulated' more could have been made of this earlier in the

article as a reason for the study. Good to look at the different types of triangulation

and be transparent about who did what when and why. The researchers should be

indicated by initial at least for each part in the methods and what degree of

independence that did and how consensus was reached ie 'researcher triangulation'

then you have data triangulation etc.

Thank you. We agree and have briefly outlined the methods of triangulation and framework synthesis at the start of the introduction:

In developing complex (especially behavioural) interventions, the synthesis of multiple types of information, including data from individual trials, surveys and interviews, and systematic reviews of qualitative and quantitative evidence, and other studies, is often helpful, yet there are few established methods of synthesis. By addressing the same question from more than one perspective or technique, the findings from different methods can be compared and contrasted. This process is known as triangulation8. If the findings across the different methods are similar, or reinforce one another, then the findings can be considered more robust than those from each method alone. Framework synthesis utilises an a priori 'framework'9 and offers a highly structured approach to triangulation. In the context of intervention design, synthesis that can filter the evidence from multiple sources in a way that can inform the choice of intervention components is highly desirable yet, at present, there are no clearly defined methods for this.

We have also revised the methods section to be more transparent about the process of triangulation including who did what, when and why (as discussed in response to #8, above).

21. P17 line 22-48 should come at the beginning of the discussion.

The discussion section has been revised considering the major amendments to the methods and results section (starts p23).

22. Line 52-53 this statement comes rather late and should be substantiated early (no

previous reviews)

Thank you. We have now included this as a point in the methods section as discussed in response to #6 (above).

23. Strengths and weaknesses – a comment made here that this study was on their own

reviews, if this is the case? Did any of the included papers in the reviews overlap?

Thank you. We have added this point as a limitation:

Limitations include the limited number and diversity of studies in each review; the limited evidence from evaluations targeting the specific determinants identified in the interviews (review 1) and surveys (review 2); and possible biases in the primary reviews, which were conducted by ourselves.

There was one overlap, but no data duplication as one study used both interview and survey methods. This is noted in the methods section using a footnote (see footnote c).

Footnote c: One study was included in both the interview and the survey synthesis

24. What will be done with the findings

The series of reviews were conducted to feed into policy thinking and development at the Department of Health and Social Care.

25. Going back to the abstract I think the difficulties with the methodology should have

been indicated in the abstract or the article summary, and a more critical analysis of

the findings related to the aim rather than the worked example

Thank you. Please see our response to #19 above.

Reviewer: 2

Reviewer Name: Sarah Denford

Institution and Country: University of Exeter

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

overall, I think the manuscript is an important and well written contribution to the literature. I have a few minor questions / issues for the authors.

Background:

Clearly describes the research aim and relevant literature.

Methods:

I was a bit confused as to where the three included reviews came from. You state in the introduction that there is only one published SR of self-care for MA; yet you apparently include three SRs of SC for MA in the present study. It would also be helpful to describe how you identified these reviews, how they were chosen etc. I appreciate that this is not the aim of your manuscript; however, you do present findings in a way that suggests that you can utilize the data to inform future interventions. However, there is no consideration of how these reviews were identified or selected, or whether any other similar reviews exist. It would be useful to know something about the aims, inclusion criteria, quality etc of the included studies. It also appears that two of the reviews were synthesized using the TDF (Table 1). Would the approach used still work for reviews that do not use the TDF when synthesizing the data?

Many of these comments replicate those of reviewer 1. Please see #1 and #6, reviewer 1 above. In addition, in Table 1 (in the introduction), which summarises the reviews included in the overarching synthesis, we have added columns detailing the study focus, type of data, conceptual synthesis, data synthesis and quality of the primary studies in the reviews. For more detailed information, readers are directed to the report published on the EPPI Centre website (due to journal space limitations and the methods focus of the current review).

In terms of the suitability of the methods, it seems likely that this approach is most useful when the TDF/BCW/BCT taxonomy are used to explore a behavioural problem within separate syntheses, before being brought together. Whether this approach can be used among reviews that did not use these tools to originally synthesise their data is not known but is likely to be influenced by the complexity of the behaviour under examination, and the transferability of the findings onto the TDF/BCW system.

We have added this point in the strengths and limitation section of the discussion:

Both of the views reviews (interviews and surveys) were already synthesised using the TDF, whether this approach to synthesis would work if other models had been used is unclear.

You mention that you amend the labels for the themes and sub themes for consistency. Did you check that the original authors were happy with your modified labels?

Given the infancy of the approach and the generalisation of theory from the general to the specific context, the creators of the tools encourage modifications and revisions, where necessary. This is an ongoing process that helps to feed into the evolution of the tools. Our understanding is that feedback is invited and highly encouraged. In any case, these revisions were made in the original syntheses of the reviews, not the overarching synthesis discussed here. Please see response, on this issue, to reviewer 1 above (#19).

Results:

Was any information relating to the content of "educational" interventions? Knowledge of services (which didn't seem to be an issue for participants who completed one of the survey studies) is quite different to knowledge relating to whether or not a symptom needs urgent treatment. This appears to be the bigger issue for participants in the interview studies.

Thank you. We have explicitly mapped the content of the evaluations for BCTs (see supplementary tables 1 and 2), TDF and COM-B domains (see Table 3).

We have acknowledged that one of the limitations of the review is the diversity of studies and behaviour. We also note that:

All the reviews targeted multiple behaviours as is common in this field of research to support self-care. We were unable, therefore, to map the behaviour change theory to specific behaviours as has been done in other studies.

I was unsure why "previous experience" is considered reinforcement? Could do with some clarification.

Thank you. This has now been clarified.

In the qualitative review, it was suggested that previous prescriptions can reinforce the need for a consultation with a GP, and that those who had attended A&E in the past were more likely to attend again. We assumed that this was due to habit formation for the behaviour. Consistent with this, data from the surveys indicated that past-experience of the service or treatment was important to A&E attenders (weight mean 63%), and fairly important to pharmacy attenders (38%). Re-attendance was also reported for the GP (weighted mean 26%) and urgent care centres/walk-in centres (weighted mean 21%) albeit to a lesser extent. The label of the sub-theme has been clarified in Table 3: Non-self-care due to habituated behavioural patterns in consulting.

I am also unclear why delayed prescribing was mapped onto reinforcement. To me, this is more about managing emotions (I feel better knowing that I have / can take my prescription if I need to).

Delayed antibiotic prescribing was mapped onto the reinforcement domain as it was assumed that it worked through weakening a habitual association between minor symptoms and the need to see a GP, as assessed by subsequent consultation rates. We agree that having negative emotions (fear of serious illness) may be reduced by having a prescription if needed, but the outcomes assessed in the interventions are about subsequent consultation rates, which may include consultations for new episodes of illness.

Your comment highlights the importance of exploring the relationship between the TDF domains. This was outside of the scope of this review but has been added as a point for future research:

Furthermore, whilst exploring the inter-relationships between the TDF domains was outside the scope of the review, greater consideration of the inter-relationships between theoretical domains may be warranted. For example, the recursive relationship between environmental and resource factors, and individual perceptions and behaviour, in the decision to self-care.

Discussion:

Some discussion of the weighting given to the various types of evidence would be useful. You considered themes to be important if there was consistency, or if they were salient in survey respondents. Was this also the case for the qualitative data?

We didn't weight the qualitative data; instead emphasis was placed on consistency between the methods. The threshold of 25%, used to determine the relevance of themes in the survey data was arbitrary and this is noted as a limitation in the methods section. We have commented on triangulation methods and noted how consistency can provide more confidence in the findings than each method alone.

Minor points:

Error message on page 11.

Thank you. We have now corrected this typo.

Reviewer: 3

Reviewer Name: Colin Greaves

Institution and Country: University of Birmingham, UK

Please state any competing interests or state 'None declared': none declared

Please leave your comments for the authors below

GENERAL COMMENTS

This is an interesting methodological paper and I support its publication in BMJ Open. However, there at a number of mainly presentational issues that need to be addressed ...

ABSTRACT

The notation "n" in the abstract seems to refer to number of studies - standard notation in reviews is to use "k" for number of studies and "n" for number of participants. For clarity /readability by a wide audience it might be better in any case to say "(review 1, 20 studies)"

Thank you. We agree and have revised the abstract as suggested.

Not clear what is meant by "Salient TDF domains were then integrated into BCW" - quite hard to follow for non-specialists in particular

Considering the manuscript restructuring this section has been reorganised and the sentence (Salient TDF domains were then integrated into BCW) removed from the abstract. We hope that this reorganisation is more accessible to the reader.

ARTICLE SUMMARY

" ...allowing for specific recommendations to be made and tested." Suggest to remove "and tested" as no such testing was performed in the current study

The word ‘tested’ is used to convey that the suggested strategies need to be evaluated as the review does not do this.

INTRO

International readers may not understand what "NHS111" is /can you offer a few brief words of explanation?

Thank you. We have now clarified that this is a telephone triage service in the UK.

In general, the Intro is lacking in references: e.g. what is the evidence that minor ailments "often place an unnecessary strain on these overstretched services."?

Thank you. We have added more references to support the claims made in the introduction section as suggested.

The following reference is used to support the evidence that minor ailments often place an unnecessary strain on health services:

Fielding S, Porteous T, Ferguson J, et al. Estimating the burden of minor ailment consultations in general practices and emergency departments through retrospective review of routine data in North East Scotland. Family practice 2015;32:165-72.

Is Table 1 cited correctly in the text (the Table doesn't seem to match what the text is saying) - maybe need to swap Table 1 and Table 2?

Thank you for highlighting this issue. This was a typo. We have now added text to explain Table 1, and cited Table 2 correctly:

Here, we report an overarching synthesis of three interconnected systematic reviews, undertaken by our team. These were syntheses of service-user views in interviews (review 1) and surveys (review 2), and evaluations (review 3) of a range of interventions and services. The overarching synthesis used the Theoretical Domains Framework (TDF)1 and the Behaviour Change Wheel (BCW)2. The

reviews, summarised in Table 1 below, (and reported in full elsewhere) explored the self-care of minor ailments (MAs) in the UK3.

BCTs were then mapped onto the TDF framework (see labels and definitions, Error! Reference source not found.).

It would be helpful to include a diagram of the BCW - if you don't already know what BCW is, the text alone does not describe sufficiently what is intended here "The COM-B system forms the hub of the BCW7 and, in conjunction with the next layer of the BCW, can be used to identify potentially relevant intervention functions, based on the salient TDF and COM-B domains.". A diagram would help to understand these complex concepts.

Thank you. We have added the diagram of the BCW (in the introduction) and agree that this addition facilitates comprehension of the model and manuscript.

In the Introduction or Discussion, you could refer to other methods used to synthesise data from different methodologies, particularly in the context of informing intervention design (e.g. triangulation protocol [1], Framework synthesis). Also, needs some text at the start to set up the problem /the rationale for doing this study ... for example along the lines of... In developing complex (especially behavioural) interventions, we often need to synthesise data of multiple types to identify barriers to change and potentially effective intervention components to overcome these barriers, as well as theoretical change process. This may mean synthesising data from both qualitative and quantitative systematic reviews, evidence based guidelines, individual trials, surveys and other studies. Doing this in a way that filters the evidence in a way that can inform choice of intervention components is desirable, but as yet, no clearly defined methods to achieve this complex task.

THEN outline one or two existing approaches (but point out that the aforementioned process of "strategic filtering" /purposeful synthesis is not included. THIS IS THE KEY VALUE OF THIS PAPER, so worth flagging this up at the outset?

Thank you. This is extremely helpful, and we have revised this section as suggested.

The opening paragraph of the revised introduction sets up the problem in the context of intervention design and briefly discusses the synthesis methods of triangulation and framework synthesis. Following on from this, the BCW approach is introduced as having the potential to filter the evidence from multiple sources in a way that can inform the choice of intervention components through the identification of barriers to change and the associated theoretical change processes :

In developing complex (especially behavioural) interventions, the synthesis of multiple types of information, including data from individual trials, surveys and interviews, and systematic reviews of qualitative and quantitative evidence, and other studies, is often helpful, yet there are few established methods of synthesis. By addressing the same question from more than one perspective or

technique, the findings from different methods can be compared and contrasted. This process is known as triangulation8. If the findings across the different methods are similar, or reinforce one another, then the findings can be considered more robust than those from each method alone. Framework synthesis utilises an a priori 'framework'9 and offers a highly structured approach to triangulation. In the context of intervention design, synthesis that can filter the evidence from multiple sources in a way that can inform the choice of intervention components is highly desirable yet, at present, there are no clearly defined methods for this.

Recent developments in the field of behaviour change encourage a systematic approach that has the potential to inform the choice of intervention components through the identification of barriers to change and associated theoretical change processes. Here, we report an overarching synthesis of three interconnected systematic reviews, undertaken by our team. These were syntheses of service-user views in interviews (review 1) and surveys (review 2), and evaluations (review 3) of a range of interventions and services. The overarching synthesis used the Theoretical Domains Framework (TDF)1 and the Behaviour Change Wheel (BCW)2. The reviews, summarised in Table 1 below, (and reported in full elsewhere) explored the self-care of minor ailments (MAs) in the UK3.

METHODS

It is not clear how "findings from the evaluations in review 3 were mapped onto the TDF and cross-tabulated, where possible, with the findings from the interview and survey studies". Maybe referring to one of the tables or using a new table to provide an example of how this mapping process might work - to be maximally useful and promote uptake of this innovative method, the paper needs to demonstrate HOW to perform this complex form of synthesis. I understand that this is challenging, but I'm not sure I could replicate this process from the current text.

Thank you.

The methods and results sections have been revised to reflect the methods focus of the paper and each includes the same three main sections (starts page 9):

- How interventions were mapped: identification of BCTs and TDF domains (step 1)
- Triangulation of findings using the TDF and COM-B system of the BCW (step 2)
- BCTs and intervention strategies (step 3)

In addition to mapping the presence or absence of determinants characterised using the TDF, Table 3 now shows the mapping of the determinants onto the COM-B model and highlights where the interventions (identified in review 3) target these.

A new column has been added to Table 4 entitled ‘existing interventions’ (to highlight if and how the determinant(s) are currently targeted by the interventions in review 3).

Supplementary tables (s1 to s3) have also been included to outline the mapping process not reported in the manuscript. Supplementary table 1 details the studies in the synthesis of evaluations (including information on behaviour, target, context, intervention provider and BCTs); Supplementary table 2 reports the BCTs identified in the interventions (with examples) and mapped onto the theoretical domains framework (TDF); and Supplementary Table 3 describes the choice of intervention functions.

RESULTS

I expected to see worked examples of how the various mapping processes worked as well as summary tables of the mapping (even if supplementary files). As it stands, there is something of a gap in the narrative of the paper - we go from a broad summary of the methods to summary level Results. There seems to be an explanatory /illustrative step missing here to complete the audit trail - how you we get from the data to the Results? I feel this is really important, so that others can replicate the methodology in future studies.

Thank you. We have now revised the results section and added supplementary tables outlining the mapping process (as discussed above).

It would be good to reserve judgement /discursive text for the Discussion (e.g. "This is surprising as the participants in the qualitative review requested information on symptom management ..."

We agree. All discursive text has been moved to the Discussion.

In general the Results section lacks structure and comes across as a list of observations (albeit gathered under two main headings). (As above) I cant see how the Results map onto the Methods described, or how the synthesis worked. Some of it is clearly a narrative synthesis of the findings, but this seems to be based on a more detailed mapping process which is not illustrated. I realise that this sounds rather negative, but I am actually very supportive of publication - I think this is an innovative and interesting study, but it needs to be written and structured much more clearly to make it a high quality paper. The standard to aim for (as with all science) is to ensure that the Methods that have been used /analyses applied are replicable by people reading the article. Example tables showing the mapping of data source /data onto elements of the TDF would help to illustrate how the analysis led to the findings. Supplementary data files could also be used to present more of (or the whole of) the mapping process? It may also help to provide Table 3 at the start of the section to provide the reader with a structure for processing the text from each domain.

Thank you. We have now revised the methods and results sections and added supplementary tables outlining the mapping process (as noted above). Where reported in the text, we have placed tables at the start of the relevant section to provide improved structure for the reader.

NB: Is the word "barriers" the right word to head the column on Table 3 - maybe it is more about "influences" or determinants as the influences can be both positive or negative (e.g. both barriers and enablers are referred to elsewhere in the article)?

We agree. We have adopted the term determinants throughout.

P11 line 51-53: Grammar error? "... data from the views studies, highlighted the importance of ..."

Thank you. We have now corrected this error.

Similarly, Table 4 could come at the start of the 'Step 2' section to ground the reader in what is being discussed (in what is otherwise a difficult read for the non-specialist reader). Again "barriers" might be better phrased as "influences" or "determinants". As above, a figure outlining what the BCW is seems necessary here for de-coding the table /the findings

Thank you for your helpful comments. As noted above, tables reported in the text, have been placed at the start of the relevant section to provide improved structure for the reader. The term 'determinants' has been adopted throughout. We have also added a figure of the BCW to the introduction to support comprehension of this approach.

DISCUSSION

The sub-heading structure seems appropriate (e.g. Summary of the principal methodological findings). However, the text included under the first heading does not always seem to fit the heading - this could be shortened considerably - what are the key points you are trying to make here?

Some very long sentences here that need breaking down /simplifying. E.g. "The 'limited roles' sub-theme (categorised within the Environmental context and resources domain) might also have been coded within the Social influences (professional role and identity) TDF domain, but after extensive deliberation and some uncertainty about the definition ('A coherent set of behaviours and displayed personal qualities of an individual in a social or work setting') the broader domain of Environmental context and resources was

chosen (including any circumstance of a person's situation or environment that discourages or encourages the development of skills and abilities, independence, social competence, and adaptive behaviour)." I have read this 3 times and still don't get what it is trying to say.

Thank you. This section has been revised and the text reduced considerably.

In the original draft most of conceptual issues related to the synthesis in the primary reviews (reviews 1 and 2). The revised version of the manuscript is focused on the overarching synthesis where there were fewer issues.

For more information, please see response to reviewer 1, #19.

Strengths and limitations:

Again, the text is quite long. Instead of saying "In the absence of evaluation data, themes and sub-themes were identified as important based on concordance of findings across both the views and surveys, where at least a quarter of survey participants identified the theme as important. The threshold of 25%, used to determine the relevance of themes in the survey

data, was arbitrary.", why not just say "The threshold of 25%, used to determine the relevance of themes in the survey data, was arbitrary."?

We agree and have reduced text where possible to make it more accessible.

Other limitations should include the limited number and diversity of studies in each review and the lack of any evidence from trials targeting the specific influences /determinants identified. In addition, given that the analysis here is complex (and primarily a narrative synthesis) are there any limitations in relation to what level of confidence we can place in the findings - how do we know if the data is "saturated" or if the inferences being drawn are based on sufficiently powered statistical analyses? - how do we know how robust the findings of this type of data synthesis might be? For example, in quantitative review studies you could do a statistical power calculation, or run a publication bias test (for reviews), in qualitative studies you could use a quality assessment framework (e.g. Lincoln & Guba), but here there are (as yet) no clear criteria to allow researchers to judge the relative strength or weaknesses /limitations of the findings. Further work may be needed to define criteria delimiting the requirements for trustworthiness /confidence is outputs when using this type of synthesis?

Thank you. We agree and have revised the strengths and limitations section to include the limited number and diversity of studies in each review, the limited evidence from evaluations targeting the specific determinants identified in the interviews (review 1) and surveys (review 2), and possible biases in the primary reviews, which were conducted by ourselves. It is also noted that without longitudinal modelling studies or intervention designs, it is difficult to establish which determinants are most important and whether the links between theoretical assessments and behaviour change techniques are valid. However, we highlight that consistency across multiple sources provides more confidence in the findings than each method alone. As you suggest, we note that criteria to assess the quality of outputs when using this type of synthesis would be beneficial, but to our knowledge, these are not currently available in the literature.

CONCLUSIONS

These are nicely put. However, I wasn't sure about the last sentence. Possibly "The theoretical scaffold provides a means to COLLATE /SYNTHESISE the evidence" - but, I don't know what you mean by "giving longevity to the research" and cant see how the evidence presented "facilitates the generalisation of the findings to similar contexts." Maybe it is just a matter of phrasing, but please consider clarifying or editing this sentence.

Thank you. This section has been revised in accordance with your feedback. We were trying to make the point that theory-based models of behaviour can be applied to other similar behavioural domains and tested, so evidence can be accumulated rather than starting from scratch each time.

The following is stated in the revised version: The theoretical scaffold provides a means to accumulate the evidence and could potentially be used to understand behaviour in similar contexts.

1. O'Cathain A, Murphy E, Nicholl J. Three techniques for integrating data in mixed methods studies. BMJ. 2010;341:c4587.

Thank you for this reference. We have included it in the first paragraph of the introduction when triangulation methods are briefly discussed.

VERSION 2 – REVIEW

REVIEWER	Sarah Denford University of Exeter, England
REVIEW RETURNED	23-Jan-2019

GENERAL COMMENTS	The authors have addressed all comments very clearly. It is now much clearer how the review was conducted. Overall, I think this is a really interesting piece of work.

REVIEWER	Colin Greaves University of Birmingham
REVIEW RETURNED	04-Feb-2019

GENERAL COMMENTS	Abstract (and Intro): "Synthesis that can filter the evidence from multiple sources to inform choice of intervention components is highly desirable yet, at present, there are no clearly defined methods" There is quite a large literature on meta-synthesis of both qualitative and quantitative data (and mixtures of the two), including methods for umbrella reviews (reviews of reviews). For example, meta-regression has been used extensively to try to identify intervention components that are associated with

	effectiveness. Maybe you could just make this statement a bit more specific to say that there are currently no clearly defined methods which use A THEORETICAL BEHAVIOUR CHANGE FRAMEWORK to inform this type of synthesis? (the bit in caps is the unique aspect of this study) Scientific Summary: To make the 2nd bullet less ambiguous (as per previous comment), I would suggest to say "allowing specific recommendations to be made for intervention design and future research in this area." Methods: The revised content and structure seems more appropriate. However, can you check through for consistency /clarity of terminology - when you say "the evaluations" what is that referring to (intervention content, the findings of the prior review? Perhaps specifically it is findings relating to the effectiveness of specific intervention components?). Do you mean the same thing each time you say "evaluations". Not clear what you mean by "To develop the content of the BCTs ..." in step 3 (para 2) - whatever this is, it is also referred to as a "decision-making process" in the following sentence. The flow of logic /methodological steps in this paragraph isn't clear Results: The following text is discursive /could be rephrased: "Which additional strategies, then, could be used to target the identified barriers to the self-care of MAs?" could this be rephrased? Is Table 4 a result of the analysis, or a set of recommendations for future interventions? Can you clarify where this comes from /how this relates to the analysis conducted? I would suggest that "The BCW SUGGESTS that change for psychological capability can be enhanced through interventions that target the Knowledge/Skills domain" (not "shows that"). As with the above comment, it is not clear where this whole paragraph (or the following text through to stakeholder involvement) comes from - is this coming from the analysis? - it seems to be simply a description of how the BCW /taxonomy defines specific techniques. Not sure this belongs here? If that is what it is, perhaps put this text in a Box or table as a 'look up' resource for readers? Discussion This sentence needs a verb? "However, the content of the strategies, whilst based on the salient determinants in the overarching synthesis, required some creativity" Suggest to add "for people with minor ailments" to the sentence "Mapping the salient TDF domains onto the COM-B system of behaviour change showed that all aspects were relevant targets for promoting self-care behaviour" typo? "Both of the views reviews (interviews and surveys) were already synthesised ..."

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 2

Reviewer Name: Sarah Denford

Institution and Country: University of Exeter, England

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below The authors have addressed all comments very clearly. It is now much clearer how the review was conducted. Overall, I think this is a really interesting piece of work.

Thank you.

Reviewer: 3

Reviewer Name: Colin Greaves

Institution and Country: University of Birmingham

Please state any competing interests or state ‘None declared’: None declared

Please leave your comments for the authors below Abstract (and Intro): "Synthesis that can filter the evidence from multiple sources to inform choice of intervention components is highly desirable yet, at present, there are no clearly defined methods" There is quite a large literature on meta-synthesis of both qualitative and quantitative data (and mixtures of the two), including methods for umbrella reviews (reviews of reviews). For example, meta-regression has been used extensively to try to identify intervention components that are associated with effectiveness. Maybe you could just make this statement a bit more specific to say that there are currently no clearly defined methods which use A THEORETICAL BEHAVIOUR CHANGE FRAMEWORK to inform this type of synthesis? (the bit in caps is the unique aspect of this study)

Thank you. We agree and have revised the sentence as suggested to make the statement more specific to the method of using a behaviour change framework to synthesise multiple sources of evidence to design interventions. The following is now stated:

Synthesis that can filter the evidence from multiple sources to inform the choice of intervention components is highly desirable yet, at present, there are few examples of systematic reviews, that explicitly define the methods to inform this type of synthesis using a theoretical behaviour change framework

Scientific Summary: To make the 2nd bullet less ambiguous (as per previous comment), I would suggest to say "allowing specific recommendations to be made for intervention design and future research in this area."

Thank you. We have revised the scientific summary as suggested. The following is now stated:

Framing of the determinants, in terms of the TDF and COM-B, supported the identification, using the BCW approach, of potential interventions that target the likely determinants of self-care behaviour, allowing specific recommendations to be made for intervention design and future research in this area.

Methods:

The revised content and structure seems more appropriate. However, can you check through for consistency /clarity of terminology - when you say "the evaluations" what is that referring to

(intervention content, the findings of the prior review? Perhaps specifically it is findings relating to the effectiveness of specific intervention components?). Do you mean the same thing each time you say "evaluations".

Thank you. This is really helpful. We have been through the manuscript to check for consistency in our use of the word “evaluations”.

We can now clarify that the term evaluations is used to refer to review 3 which assessed the effectiveness of behavioural interventions and services that support self-care for MAs.

Since the evaluations of services were not coded for BCTs (as the active ingredients concerned elements other than the content of the treatment, including provider type, delivery format and setting) the phrase evaluations of interventions is used for discussing the content of the interventions.

To support this, in the Methods section, we have also made a clear distinction between evaluations of services and evaluations of interventions.

Not clear what you mean by "To develop the content of the BCTs ..." in step 3 (para 2) - whatever this is, it is also referred to as a "decision-making process" in the following sentence. The flow of logic /methodological steps in this paragraph isn't clear

Thank you. We agree that this section was unclear. The intervention functions, and BCTs were selected to address the key determinants of self-care behaviour for MAs identified in the interviews, surveys and evaluations of interventions. The process was supported using the APEASE criteria in collaboration with stakeholders.

This section has been altered to make this clear:

BCTs were selected, therefore, to target the most salient determinants of self-care for MAs identified in interviews, surveys and evaluations of interventions. This process was supported using the APEASE criteria (Acceptability, Practicability, Effectiveness/cost-effectiveness, Affordability, Safety/side-effects, and Equity)20 in consultation with stakeholders (see below), and the results were put into a matrix.

Results:

The following text is discursive /could be rephrased: "Which additional strategies, then, could be used to target the identified barriers to the self-care of MAs?" could this be rephrased?

Thank you. The discursive elements of the sentence have been removed and now reads:

Additional intervention strategies that could be used to target the identified barriers to the self-care of MAs were, therefore, suggested.

Is Table 4 a result of the analysis, or a set of recommendations for future interventions? Can you clarify where this comes from/how this relates to the analysis conducted?

We can confirm that Table 4 is both a result of the analysis and a set of recommendations for future interventions. The methods for this section are described underneath the sub-heading “BCTs and intervention strategies (step 3)” and have been revised to improve clarity:

The next step was to identify the strategies that are likely to be effective in promoting self-care for MAs. Nine intervention functions were specified in the BCW and mapped on to the COM-B domains: education, persuasion, incentivisation, coercion, training, restriction, environmental restructuring, modelling and enablement (see table 2.3 in the Behaviour Change Wheel guide20). Using the guidance from the BCW, those intervention functions that were most likely to address the key determinants of self-care behaviour (identified in the previous step) were selected. BCTs that were linked to the intervention functions were identified using the BCTv111 and the results from an expert consensus exercise that mapped BCTs onto intervention functions (see table 3.3 in the Behaviour Change Wheel guide20). BCTs were selected, therefore, to target the most salient determinants of self-care for MA identified in interviews, surveys and evaluations of interventions. This process was supported using the APEASE criteria (Acceptability, Practicability, Effectiveness/cost-effectiveness, Affordability, Safety/side-effects, and Equity)20 in consultation with stakeholders (see below), and the results were put into a matrix.

I would suggest that "The BCW SUGGESTS that change for psychological capability can be enhanced through interventions that target the Knowledge/Skills domain" (not "shows that"). As with

the above comment, it is not clear where this whole paragraph (or the following text through to stakeholder involvement) comes from - is this coming from the analysis? - it seems to be simply a description of how the BCW /taxonomy defines specific techniques. Not sure this belongs here? If that is what it is, perhaps put this text in a Box or table as a 'look up' resource for readers?

The text presents illustrative examples derived from the analyses (step 3) for each COM-B domain presented in Table 4. To improve clarity, this has now been explicitly clarified in the text and the examples have been shortened and moved so that they immediately follow the introduction to table 4.

The text now reads:

Table 4 shows the salient TDF domains, mapped onto the COM-B domains, systematically selected intervention functions, strategies and BCTs to deliver the relevant intervention functions. Illustrative strategies, derived from the analyses, are briefly discussed for each COM-B domain below.

Capability

Greater emphasis on knowledge of self-care resources in addition to symptom management may enhance people's capability through improved access to the support they need to self-care effectively. Furthermore, an emphasis on skills training (such as self-monitoring of the behaviour, behavioural rehearsal and demonstration of the behaviour) may improve physical capability particularly (over and above didactic knowledge acquisition or theory), e.g., through using a daily symptom diary, when sick, to improve skills in the recognition and treatment of MAs and the identification of danger signs and symptoms.

Motivation

In addition to prescribing interventions examined in review 3 (delayed and none), other strategies may be usefully implemented and tested. For example, automatic motivation may be targeted through enabling service-users to identify anxiety as a trigger to visit GP/A&E services and to initiate coping strategies to overcome such urges. In terms of reflective motivation, targeting beliefs about illness severity and susceptibility, especially among parents of children, may be beneficial, e.g., using persuasion to strengthen beliefs that the opinions of pharmacists and nurses are trustworthy.

Opportunity

For social opportunity, a suggested strategy was the enablement of self-care through the provision of reassurance (e.g., from a pharmacist) that self-care is appropriate; and for physical opportunity, restructuring of the environment was indicated, e.g., by training more nurses and pharmacists with full prescribing rights.

Discussion

This sentence needs a verb? "However, the content of the strategies, whilst based on the salient determinants in the overarching synthesis, required some creativity"

Thank you. A verb has been added:

However, defining the content of the strategies, whilst based on the salient determinants in the overarching synthesis, required some creativity.

Suggest to add "for people with minor ailments" to the sentence "Mapping the salient TDF domains onto the COM-B system of behaviour change showed that all aspects were relevant targets for promoting self-care behaviour"

Thank you. We agree and have revised as suggested.

typo? "Both of the views reviews (interviews and surveys) were already synthesised ..."

Thank you. This sentence has been revised to make it more accessible to the reader:

Both of the interview and survey reviews were already synthesised using the TDF.