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)	A protocol for a feasibility and acceptability study using a brief ACT based intervention for people from Southwest Wales who live with persistent pain
AUTHORS	Saracutu, Madalina; Edwards, Darren; Davies, Helen; Rance, Jaynie

VERSION 1 – REVIEW

REVIEWER	Lance McCracken King's College London, United Kingdon
REVIEW RETURNED	27-Feb-2018

GENERAL COMMENTS	The submitted protocol describes research that is to be a part of the first author's PhD thesis. The treatment and population are interesting, appropriate, and deserving. The study itself seems to have a number of strengths. It is meant to be preliminary only, a study to lead to the design of another study. There is a lot to recommend this protocol and a few concerns, which the authors may be able to address. 1. I thought the abstract would be more informative if the projected sample size and primary outcomes were described. 2. One of the main weaknesses of the protocol is in ambiguity around its aims and objectives. It is described as a pilot and as a feasibility and acceptability study. Also there are points where the authors present it as if the main purpose is to test clinical outcomes and generalise the results. Pilot aims and feasibility aims are different to a degree – the authors should consult a definitive source for the meanings of these terms and then frame their study accordingly. The authors should be more precise about what they mean to do and achieve. 3. An example of the ambiguity caused in how the protocol is presented is in the data analysis section where the first method presented is a within subjects ANOVA looking at change. If this is indeed a meant to be a feasibility study the analyses ought to be focused around feasibility outcomes (recruitment rate, retention, data quality, treatment completion, patient acceptance) and feasibility related analyses. 4. In a similar vein, the "expected outcomes" section sounds more like the expected outcomes of some other future study. 5. Yet again, discussion of "transferability of the findings to other clinical settings" in the context of a feasibility study is not exactly the same as for a study focused on clinical outcomes. The findings of a feasibility study essentially need to inform the methods of a subsequent study.

REVIEWER	Julie Ashworth Institute for Primary Care and Health Sciences, Keele University, UK
REVIEW RETURNED	21-Mar-2018

GENERAL COMMENTS	Whilst acknowledging the importance of developing and testing brief psychologically based interventions for chronic pain which are accessible to patients, I have a number of concerns about this protocol which I think need to be addressed before it would be suitable for publication. First, the quantitative outcome measures need to be more closely aligned with the stated study objectives of testing the feasibility and acceptability of a new brief ACT intervention. The qualitative work seems, appropriately, to focus of the acceptability of the intervention. Adding quantitative measures of acceptability and adherence / engagement with the intervention would be an advantage. However the quantitative outcome measures are a mixture of process measures (of psychological flexibility) and treatment outcome measures (HAD and EQ5D). Whilst I accept that testing the outcome measures that may be used in a future full scale trial is appropriate with a view to refining questionnaires and processes, this sample is small and has no comparator group, and therefore seems unlikely to yield useful treatment outcome data. If the authors wish to test the feasibility of conducting a full scale trial of this intervention then it seems they should consider outcomes appropriate to that e.g. recruitment/ retention rates, treatment fidelity, follow-up rates, missing data on questionnaires etc. It would also be worthwhile consider what treatment outcomes would be included in the questionnaire for a future full scale trial in light of the IMMPACT recommendations for chronic pain trials (I note there is no outcome measure relating to pain or physical functioning). In addition, I have a number of comments regarding the introduction: • I think it would be more accurate to describe catastrophizing and fear avoidance as maladaptive coping behaviours rather than 'comorbidities'…….and as such people don't 'experience' catastrophizing – they 'do' it • The analogy of pain as an alarm system warning of potential damage is more appropriate to acute pain rather than chronic pain – worth clarifying this. • Chronic pain is often accompanied by distress – but it's not true to say that distress accompanies pain consistently • ACT is essentially a form of CBT - the distinction between ACT and traditional CBT could be better explained – there are more similarities than perhaps the authors suggest – in the context of chronic pain – both focus on improving quality of life and functioning in the presence of pain and good traditional CBT will also take into account values – the main distinction is perhaps best described in terms of cognitive restructuring in traditional CBT versus ACT which does not attempt to change thoughts or experiences and aims to increase psychological flexibility • The authors quote the American Psychology Association position in terms of ACT being strongly supported by empirical evidence but perhaps overstate the strength (quality) of evidence in relation to ACT for chronic pain. I would encouraged a more balanced view in terms of the evidence supporting ACT for chronic pain and how it compares to the evidence for other psychological based therapies in chronic pain taking into account recent meta-analyses.

	• The authors cite a previous pilot study of a brief ACT intervention in primary care (McCracken 2013) but do not say how they plan to improve on this, how the findings have influenced development of this intervention or why this intervention will be any more successful or acceptable • It would be helpful to know more about the findings of their preliminary qualitative study and how this influenced intervention development • The interdisciplinary collaboration between osteopathy and psychology would probably benefit from greater explanation – presumably physiotherapists could replace the osteopaths if necessary to improve feasibility in other centres which don't have osteopaths?

VERSION 1 – AUTHOR RESPONSE

Response to reviewer 1

We would like to thank Prof. Lance McCracken for his pertinent comments. They were valuable in clarifying and improving the aim and the design of the study.

1. I thought the abstract would be more informative if the projected sample size and primary outcomes were described.
The sample size and the primary outcomes have been included in the abstract.

2. One of the main weaknesses of the protocol is in ambiguity around its aims and objectives. It is described as a pilot and as a feasibility and acceptability study. Also there are points where the authors present it as if the main purpose is to test clinical outcomes and generalise the results. Pilot aims and feasibility aims are different to a degree – the authors should consult a definitive source for the meanings of these terms and then frame their study accordingly. The authors should be more precise about what they mean to do and achieve.
3. An example of the ambiguity caused in how the protocol is presented is in the data analysis section where the first method presented is a within subjects ANOVA looking at change. If this is indeed a meant to be a feasibility study the analyses ought to be focused around feasibility outcomes (recruitment rate, retention, data quality, treatment completion, patient acceptance) and feasibility related analyses.

Many thanks for this. The objectives have been clarified and the main outcomes now focus on investigating the feasibility and acceptability of the intervention, as advised by the reviewers. Please read bellow:

'This study aims to investigate the feasibility and acceptability of a novel, ACT based psychosocial program (as described by the MRC framework) for people from Southwest Wales who live with persistent pain. The main outcomes will include the feasibility of the recruitment process and the measurement tools, the acceptability of the intervention for both the participants and the Osteopaths and the adherence to the program. '

The MRC guidance on developing and evaluating complex interventions includes a 'feasibility and piloting' stage: "The feasibility and piloting stage includes testing procedures for their acceptability, estimating the likely rates of recruitment and retention of subjects, and the calculation of appropriate sample sizes. Methodological research suggests that this vital preparatory work is often skimped. Evaluations are often undermined by problems of acceptability, compliance, delivery of the intervention, recruitment and retention, smaller-than-expected effect sizes, and so on, that could be anticipated by thorough piloting. ''
However, NIHR programmes have separate definitions for feasibility and pilot studies . They describe the focus of feasibility studies as the estimation of parameters to inform whether main studies are possible while pilot studies involve small-scale testing of the "smooth" delivery of research plans of larger studies.

These two slightly distinct definitions may cause some degree of confusion . We decided to use the MRC guidance and focus on determining the feasibility and acceptability of the intervention. The main outcomes are aligned with this aim.

We decided to use the following definition for acceptability:
'Acceptability is a multifaceted construct that reflects the extent to which people delivering and receiving a healthcare intervention consider it to be appropriate, based on anticipated or experienced cognitive and emotional responses to the intervention.'

Outcome measures
Primary and secondary outcomes
The primary outcomes include the acceptability of the proposed intervention and the feasibility of recruitment and measurement.
Feasibility
Recruitment process
• Number of people referred from Osteopaths and eligible for screening
• Number of people attending interview with researcher
• Recruitment and retention rates
Feasibility of measurement tools
• Time taken to fill in questionnaires
• Missing data from questionnaires
• Follow-up response rates (1 month & 3 months follow ups)
Acceptability
Prospective acceptability
• Barriers (screening interview)
• Burden (reasons for not taking part/ discontinuation or dropping out)
Intervention coherence and adherence
• Number of sessions attended
• Homework completion (workbook entries, Mindfulness diary)
• Time dedicated to homework practice
• ACT basic definitions quiz (six questions regarding ACT principles)
Experience (satisfaction, perceptions)- Qualitative interview (end of program)
Acceptability of the program to Osteopaths- Focus Group
The secondary outcomes (measured at baseline, upon completing of the program, at one month and three months follow- up) are meant to measure changes in outcomes such as depression and anxiety, acceptance of pain, mindfulness, fear avoidance and quality of life.

4. In a similar vein, the "expected outcomes" section sounds more like the expected outcomes of some other future study.

5. Yet again, discussion of “transferability of the findings to other clinical settings” in the context of a feasibility study is not exactly the same as for a study focused on clinical outcomes. The findings of a feasibility study essentially need to inform the methods of a subsequent study.

The ‘expected outcomes’ section has been revised:

In this feasibility and acceptability trial we shall determine the feasibility of recruiting patients living with persistent pain to a brief ACT based psychosocial intervention. This will include determining recruitment and retention rates, testing the procedures and exploring the acceptability of the program, as well as investigating potential changes in anxiety, depression, mindfulness, fear avoidance and quality of life. This study is in line with the MRC guidance for developing complex interventions (testing procedures, estimating recruitment, retention, determining sample size).

Although we cannot at this point test the effectiveness of ‘A Mindful Act’, we will learn whether it is feasible and acceptable to people living with long-term pain and psychological comorbidities from Southwest Wales. The data from the qualitative interviews will increase our understanding of the experience of taking part in the intervention. Furthermore, we will have a better knowledge of the barriers to taking part and reasons for discontinuation, and also an increased knowledge of the elements that are most valued by the participants. ‘

In response to Reviewer 2

Dear Dr Julie Ashworth,

Thank you for your comments. Please find bellow a response to your comments point by point.

Reviewer >First, the quantitative outcome measures need to be more closely aligned with the stated study objectives of testing the feasibility and acceptability of a new brief ACT intervention. The qualitative work seems, appropriately, to focus of the acceptability of the intervention. Adding quantitative measures of acceptability and adherence / engagement with the intervention would be an advantage. However the quantitative outcome measures are a mixture of process measures (of psychological flexibility) and treatment outcome measures (HAD and EQ5D). Whilst I accept that testing the outcome measures that may be used in a future full-scale trial is appropriate with a view to refining questionnaires and processes, this sample is small and has no comparator group, and therefore seems unlikely to yield useful treatment outcome data. If the authors wish to test the feasibility of conducting a full scale trial of this intervention then it seems they should consider outcomes appropriate to that e.g. recruitment/ retention rates, treatment fidelity, follow-up rates, missing data on questionnaires etc.

Authors> Thank you. The objectives of the study have been clarified and the main outcomes now focus on investigating the feasibility and acceptability of the intervention. Please read bellow:

‘This study aims to investigate the feasibility and acceptability of a novel, ACT based psychosocial program (as described by the MRC framework) for people from Southwest Wales who live with persistent pain. The main outcomes will include the feasibility of the recruitment process and the measurement tools, the acceptability of the intervention for both the participants and the Osteopaths

and the adherence to the program. ‘

We followed your suggestion and added quantitative measures of acceptability and adherence / engagement with the intervention.

Outcome measures
Primary and secondary outcomes
The primary outcomes include the acceptability of the proposed intervention and the feasibility of recruitment and measurement.
Feasibility
Recruitment process
• Number of people referred from Osteopaths and eligible for screening
• Number of people attending interview with researcher
• Recruitment and retention rates
Feasibility of measurement tools
• Time taken to fill in questionnaires
• Missing data from questionnaires
• Follow-up response rates (1 month & 3 months follow ups)
Acceptability
Prospective acceptability
• Barriers (screening interview)
• Burden (reasons for not taking part/ discontinuation or dropping out)
Intervention coherence and adherence
• Number of sessions attended
• Homework completion (workbook entries, Mindfulness diary)
• Time dedicated to homework practice
• ACT basic definitions quiz (six questions regarding ACT principles)
Experience (satisfaction, perceptions)- Qualitative interview (end of program)
Acceptability of the program to Osteopaths- Focus Group
The secondary outcomes (measured at baseline, upon completing of the program, at one month and three months follow- up) are meant to measure changes in outcomes such as depression and anxiety, acceptance of pain, mindfulness, fear avoidance and quality of life.

Reviewer > It would also be worthwhile consider what treatment outcomes would be included in the questionnaire for a future full-scale trial in light of the IMMPACT recommendations for chronic pain trials (I note there is no outcome measure relating to pain or physical functioning).

Authors > Many thanks for this suggestion, we did consider this, however, Vowles et al. (2017) suggested that according to the ACT model, reductions in pain and distress are not necessary for improvement. ACT aims to improve functioning (e.g. engagement in valued activities, acceptance of pain, increased mindfulness) even when pain and distress persist. They also suggest that significant improvements in functioning may not require decreases in pain intensity.

Reviewer > In addition, I have a number of comments regarding the introduction:
• I think it would be more accurate to describe catastrophizing and fear avoidance as maladaptive coping behaviours rather than ‘comorbidities’. And as such people don’t ‘experience’ catastrophizing – they ‘do’ it

Authors >The following changes have been made based on your advice. ‘Similarly, people with persistent pain often catastrophize.’

‘Some of the most common comorbidities found in this population are depression, and anxiety, and also maladaptive behaviours such as fear avoidance and catastrophizing. ‘

Reviewer > The analogy of pain as an alarm system warning of potential damage is more appropriate to acute pain rather than chronic pain – worth clarifying this.

Authors > The following clarification has been made. ‘Pain is essentially an alarm system, warning us of a potential danger of injury, however, for people with long- term pain this is a weekly or daily occurrence that has a significant impact on their psychological wellbeing and quality of life.’

This statement emphasizes that pain serves as an alarm system that helps survival (when it is short-term, however, for those who experience persistent pain (long-term pain) it is a daily or weekly occurrence that affects their wellbeing and quality of life.

Reviewer > Chronic pain is often accompanied by distress – but it’s not true to say that distress accompanies pain consistently

Authors >We changed this to - ‘Persistent pain is often accompanied by distress. ‘

Reviewer >ACT is essentially a form of CBT - the distinction between ACT and traditional CBT could be better explained – there are more similarities than perhaps the authors suggest – in the context of chronic pain – both focus on improving quality of life and functioning in the presence of pain and good traditional CBT will also take into account values – the main distinction is perhaps best described in terms of cognitive restructuring in traditional CBT versus ACT which does not attempt to change thoughts or experiences and aims to increase psychological flexibility.

Authors >Many thanks for this, please read bellow for the slight change in accordance with your suggestion:

Although ACT and traditional CBT share many features, the distinctiveness between these two approaches consist in the emphasis of ACT on acceptance and engaging in behaviour in line with one’s values over cognitive restructuring and symptom reduction The ultimate goal of ACT is t to improve functioning by increasing psychological flexibility and the ability to act according to personal values, even in the presence of negative experiences. Acceptance and Commitment Therapy (ACT) helps people clarify what is truly important, then use that knowledge to guide, inspire, and motivate change in order for the individual to live a more full and meaningful life.’

Reviewer > The authors quote the American Psychology Association position in terms of ACT being strongly supported by empirical evidence but perhaps overstate the strength (quality) of evidence in relation to ACT for chronic pain. I would encouraged a more balanced view in terms of the evidence supporting ACT for chronic pain and how it compares to the evidence for other psychological based therapies in chronic pain taking into account recent meta-analyses.

Authors >Thank you. This study is centred on an ACT-based brief intervention for persistent pain. We believe that our account of the empirical evidence supporting the ACT model is balanced – but we have removed the word ‘strongly’ to deemphasize this a bit. We acknowledged the limitations of the

evidence base:

Whilst there is support for the use of ACT with persistent pain populations, more methodologically robust trials are needed to compare the effectiveness of ACT with other psychologically based approaches.

Reviewer > The authors cite a previous pilot study of a brief ACT intervention in primary care (McCracken 2013) but do not say how they plan to improve on this, how the findings have influenced development of this intervention or why this intervention will be any more successful or acceptable

Authors >Many thanks for this suggestion. Just to clarify, the pilot study by McCracken was merely used as an example of integrating ACT into different settings. Our intervention is neither informed by nor comparable to the mentioned pilot. We have made this more clear in the ms.

Reviewer > It would be helpful to know more about the findings of their preliminary qualitative study and how this influenced intervention development

Authors >Thanks. The results of the qualitative study will be presented in a different paper

Reviewer >• The interdisciplinary collaboration between osteopathy and psychology would probably benefit from greater explanation – presumably physiotherapists could replace the osteopaths if necessary to improve feasibility in other centres, which don't have osteopaths?

Authors >For a complete explanation of the collaboration between osteopathy and psychology please read our previous paper, a link to this has been provided to the notes:

https://www.journalofosteopathicmedicine.com/article/S1746-0689(16)30116-X/abstract
Although osteopathy itself is not a psychosocial intervention, it might be worth combining Osteopathic treatment with brief psychological packages. Integrating concepts and principles from third wave therapies like Acceptance and Commitment Therapy (ACT) could lead to an increase in the effectiveness osteopathic care, and moderate the impact of comorbidities. This type of pairing might have a strong synergistic effect, compared to standard care alone. In fact there are recommendations to combine different types of treatment (physical, psychological, rehabilitative) in order to match patients' needs.
For this project we are collaborating with the Osteopathy clinic within the Health and Wellbeing Academy at Swansea University.

There are other projects looking at pairing ACT with physiotherapy (please check Physiotherapy Informed by ACT (PACT) study:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4908884/)

VERSION 2 – REVIEW

REVIEWER	Lance McCracken King's College London, UK
REVIEW RETURNED	04-Jun-2018

GENERAL COMMENTS	The manuscript under review is a protocol for a feasibility and acceptability study. This is a revised submission following an earlier review. The authors have taken on board earlier comments and many of the issues raised before are now satisfactorily addressed. There are nonetheless a few outstanding issues. 1. I wondered if there is a scientific rationale attached to the setting in Southwest Wales. What is the significance or importance of this, from a research standpoint, or is there none? 2. Most of the focus on looking for changes in clinical outcomes is now removed, but the authors still say a couple times such things as “reveal differences.” The sample size is very small. It seems unlikely that they will be able to estimate much from their clinical measures, at least not reliably. 3. Related to the previous point, how did the authors determine their proposed sample size. It is very small, even for estimating recruitment and retention rate, and certainly for estimates for the clinical measures. The authors might want to consult a guide for the size of sample needed in feasibility studies or a statistician to consider how to derive stable estimates for the results they want. 4. The “objective” section is very lean, and some of the material in the “study design” section sounds like objective, or like their key feasibility questions. Along with their feasibility question, the authors might also want to describe the criteria or thresholds they will use to determine whether their results meet their standard for feasibility. 5. A one-group test of feasibility of a group treatment seems somewhat risky. As the people in the group itself can have an impact on the individual’s experience. Do the authors feel that running more than on group would be a better test of feasibility, otherwise this might need to be an acknowledged limitation. 6. Is Table one needed. Table 1 and the text overlap considerably. The description of first step and second step doesn’t really seem to lend itself very well to table form. Just an opinion, but something to think about. 7. Where under “outcome measures” and heading “secondary outcomes” appears, it seems to me that these are not really secondary outcomes. These are the clinical outcome measures. Secondary outcomes may be based in the data on the performance of these clinical outcomes as tools. To me the standardized clinical outcome measures are not really outcome measures for this feasibility study. There performances: time requires, completeness of data, perhaps participant view of relevance, are study outcomes. I don’t think I would to statistical tests on the clinical measures, and look at p-values. And, even calculating effect sizes seems a little unsatisfactory, again due to the small sample size. 8. This is just an opinion, but I did not see the place of the FABQ in a study of ACT and mindfulness based treatment. It is really a belief measures, and to me neither process or outcome. 9. What is the reason to include no clinical outcome measure that assessed general, physical, or social daily functioning? Short ones like the BPI inference scale could be useful.

REVIEWER	Julie Ashworth Keele University, UK
REVIEW RETURNED	04-Jun-2018

GENERAL COMMENTS	I commend the authors for taking on board much of the reviewer feedback and the manuscript is very much improved as a result.

	Accepting that this project is part of a PhD, there remain a few areas which would benefit from further improvement. Abstract Although the title and outcomes have been amended to clarify the primary objective of the study the, the methods section of the abstract fails to mention how feasibility of recruitment will be assessed and instead offers greater detail on clinical outcomes – the latter could be summarised as measures of potential effectiveness and this would perhaps allow for something to be added about recruitment. Strengths & weaknesses The authors haven't really explained the advantage of using multiple methods of data collection - triangulation of findings? The authors should also consider adding 'lack of generalisability' given that osteopaths are not routinely involved in chronic pain management across the NHS and consider commenting on whether the intervention could be adapted to other groups of healthcare professionals in future if needed to improve the potential for implementation in the NHS. In addition, although the authors have added a paragraph to the discussion which provides some context regarding their decision to use osteopaths to deliver the intervention, they still haven't explained 'why osteopaths?' or why a collaboration between psychologists and osteopaths is desirable in either the introduction or strengths / weakness sections. It may be obvious to the authors, but perhaps not to many readers. The authors should consider including some of the paragraph about expanding the repertoire of osteopaths from the discussion in the introductory paragraphs to set the scene. Introduction Non-negligible seems an odd way of describing an estimated between one-third and one-half of the UK population When comparing ACT & CBT 'distinctiveness between ' needs rewording - distinction? P 7, 2nd paragraph - depression & anxiety should be in brackets after psychological comorbidities and catastrophizing & fear avoidance in brackets after 'maladaptive coping strategies Inconsistent spelling of catastrophizing (not catatrophising) Methods Study design The authors should consider rewording the first paragraph regarding study design 'acceptability of the intervention to (not for) patients and osteopaths' Pain and function should be included as outcome measures in any trial of an intervention in patients with chronic pain - Vowles et al. include pain and function as outcome measures - the fact that function can improve without changes in pain or distress is true but doesn't remove the need to measure important outcomes (see IMMPACT recommendations for trials in chronic pain) even if they are not targeted directly by the intervention -If it is possible to do

	so at this stage, I recommend that the authors consider including a suitable measure of pain and function e.g. the Brief Pain Inventory - or even the abbreviated form (PEG) - if questionnaire burden is a concern. Alternatively a comment should be made that it would be included in any future full trial looking at effectiveness. Expected outcomes Although now acknowledging that this study will not test effectiveness, the paragraph at the end of p 17 starts by talking about the full trial – the authors should consider starting the paragraph by saying secondary outcome is to look at potential effectiveness. How will acceptability and feasibility be defined? The authors should consider the criteria for recommending progression to a full scale trial (albeit that this will be a separate study) given that this is repeatedly mentioned or shift the focus towards the outcome of this study being to refine the intervention, which could then be tested in a future trial.

VERSION 2 – AUTHOR RESPONSE

Response to reviewer 1

Thank you for acknowledging our previous changes. We believe that we have now clarified the objectives of our study. We have followed the recommendations of the 'Feasibility and pilot studies: a guide for NIHR Research Design Service advisors' (Williams, 2016).

1. I wondered if there is a scientific rationale attached to the setting in Southwest Wales. What is the significance or importance of this, from a research standpoint, or is there none?

First of all, the setting was chosen for convenience, as the PhD has been funded by Swansea University. In addition, it was agreed to collaborate with the Health and Wellbeing Academy (an organisation within the College of Human and Health Sciences) whose aim is to complement the services provided by the NHS and allow people from Southwest Wales to make informed and positive lifestyle choices to improve their health and wellbeing.

In Wales as in the rest of the UK, it is estimated that around 13% of the population are likely affected by persistent pain (Breivik et al, 2006) .

Furthermore, preliminary results from the first large scale observational study of opioid prescription in Wales suggests that between 2005-2015 there has been a 323.4% increase in strong opioid prescription for people with non- cancer pain (Trends in Opioid Prescribing and Associated resource utilisation in Wales, 2018).

2. Most of the focus on looking for changes in clinical outcomes is now removed, but the authors still say a couple times such things as “reveal differences.” The sample size is very small. It seems unlikely that they will be able to estimate much from their clinical measures, at least not reliably.

Please read the information below, which has been extracted from Feasibility and pilot studies: a guide for NIHR Research Design Service advisors’ (Williams, 2016).

‘Should feasibility and pilot studies include an assessment of the effectiveness of the intervention? The simple answer is “no”; whilst pilot studies and some feasibility studies may measure participant outcomes, they are by definition underpowered to make useful assessments of effectiveness.

The NIHR description of feasibility studies is unequivocal in stating that they “do not evaluate the outcome of interest”. The description of pilot studies is more nuanced, stating that they will resemble the main study in many respects, including an assessment of the primary outcome. In some cases this will be the first phase of the substantive study and data from the pilot phase may contribute to the final analysis; this can be referred to as an internal pilot. Or at the end of the pilot study the data may be analyzed and set aside, a so-called external pilot.

Since the focus of this guidance is purely on external (stand-alone) pilot studies, then the advice relating to feasibility and pilot studies is the same; that any outcome measurements should not be used to assess the effectiveness of the intervention. It should also be noted that most journals do not expect to see an assessment of the effectiveness of interventions in articles reporting on feasibility or stand-alone pilot studies. ‘

This is a feasibility study, and its aim is not to assess effectiveness.

3. Related to the previous point, how did the authors determine their proposed sample size. It is very small, even for estimating recruitment and retention rate, and certainly for estimates for the clinical measures. The authors might want to consult a guide for the size of sample needed in feasibility studies or a statistician to consider how to derive stable estimates for the results they want.

In regards to the small sample size, there are other feasibility studies with a similar number of participants. Please see Fischer-White et al. (2014) who investigated the feasibility of an 8-week restorative yoga intervention and employed a number of 12 people diagnosed with fibromyalgia.

‘The size of a feasibility or pilot study will be influenced by a range of factors including the type of information that will be collected, the purpose for which it is needed, the desired accuracy of any parameter estimates, and pragmatic considerations such as time and resource requirements. When there is no need to estimate parameters with particular accuracy, the size of feasibility and pilot studies is often a pragmatic balance between information needs and resources.’ (Williams, 2016).

Perhaps if we provide some more information, this will lead to a better understanding of the pragmatic considerations that Williams refers to.

It is not possible to recruit a large number of participants in the time frame of a PhD, also considering the whole process (conducting preliminary research to inform the intervention, designing the intervention, obtaining ethical approval, recruiting participants for the intervention).

In addition, recruitment has been carried out in partnership with the osteopaths working within the Health and Wellbeing Academy, who are treating a limited number of people living with persistent pain. They volunteered to refer people to the intervention program without any form of remuneration.

Last but not least, the development and piloting of the intervention have been carried out with very limited resources. There were no additional funds or sponsorships.

4. The “objective” section is very lean, and some of the material in the “study design” section sounds like objective, or like their key feasibility questions. Along with their feasibility question, the authors might also want to describe the criteria or thresholds they will use to determine whether their results meet their standard for feasibility.

Feasibility and acceptability will be assessed by analysing the qualitative data from the one-to-one interviews and focus group. The main focus is on determining whether participants and those delivering the intervention consider it appropriate and acceptable.

Additional data will also be collected (time taken to fill in questionnaires, missing data, follow-up responses, number of people referred and eligible, number of people attending initial interview). However, this will only intend to complement the qualitative data. Recruitment and retention rate have been deleted.

This is in line with Sekhon et al (2017) :

‘Successful implementation depends on the acceptability of the intervention to both intervention deliverers (e.g. patients, researchers or healthcare professionals) and recipients (e.g. patients or healthcare professionals). From the patient’s perspective, the content, context and quality of care received may all have implications for acceptability. If an intervention is considered acceptable, patients are more likely to adhere to treatment recommendations and to benefit from improved clinical outcomes. From the perspective of healthcare professionals, if the delivery of a particular intervention to patients is considered to have low acceptability, the intervention may not be delivered as intended (by intervention designers), which may have an impact on the overall effectiveness of the intervention.’

5. A one-group test of feasibility of a group treatment seems somewhat risky. As the people in the group itself can have an impact on the individual’s experience. Do the authors feel that running more than on group would be a better test of feasibility; otherwise this might need to be an acknowledged limitation.

This is indeed a limitation and it has been now acknowledged in our limitations. Future trials will ideally involve running a number of different groups.

6. Is Table one needed. Table 1 and the text overlap considerably. The description of first step and second step doesn’t really seem to lend itself very well to table form. Just an opinion, but something to think about.

There is some overlap, however it is important to clarify the stages of the recruitment and the steps taken by the osteopaths and the main investigator.

7. Where under “outcome measures” and heading “secondary outcomes” appears, it seems to me that these are not really secondary outcomes. These are the clinical outcome measures. Secondary outcomes may be based in the data on the performance of these clinical outcomes as tools. To me the standardised clinical outcome measures are not really outcome measures for this feasibility study.

There performances: time requires, completeness of data, perhaps participant view of relevance, are study outcomes. I don't think I would to statistical tests on the clinical measures, and look at p-values. And, even calculating effect sizes seems a little unsatisfactory, again due to the small sample size.

As expressed before, the focus will be on the feasibility and acceptability. Other studies with similar sample sizes employed similar additional quantitative measures; see Fischer-White et al (2015) , Price et al. (2007).

8. This is just an opinion, but I did not see the place of the FABQ in a study of ACT and mindfulness based treatment. It is really a belief measures, and to me neither process or outcome.

Fear avoidance plays an important role in persistent pain. Recent findings from Larsson et al. (2016) suggest that fear avoidance beliefs play a more important role in predicting future physical activity levels than pain characteristics. In addition, they found that the level of physical activity was significantly lower among those with persistent pain and was significantly associated with kinesiophobia.

9. What is the reason to include no clinical outcome measure that assessed general, physical, or social daily functioning? Short ones like the BPI inference scale could be useful.

EQ5D already includes dimensions such as pain (or discomfort), self –care, usual activities (functioning) or mobility. We will consider including BPI in future studies.

In addition, given certain experiences reported by people living with pain ('brain fog', the need to move around, discomfort when sitting for long periods of time) we decided not to overburden them.

Response to reviewer 2

Thank you for your comments. Please find below a response point by point.

Abstract- The abstract is merely a brief summary of the study; therefore feasibility of recruitment will not be explained in detail. Please refer to pages 15- 16 for a definition of feasibility and acceptability and how they will be measured.

Please read the information below, which has been extracted from Feasibility and pilot studies: a guide for NIHR Research Design Service advisors' (Williams, 2016).

'Should feasibility and pilot studies include an assessment of the effectiveness of the intervention? The simple answer is "no"; whilst pilot studies and some feasibility studies may measure participant outcomes, they are by definition underpowered to make useful assessments of effectiveness.

The NIHR description of feasibility studies is unequivocal in stating that they "do not evaluate the outcome of interest". The description of pilot studies is more nuanced, stating that they will resemble the main study in many respects, including an assessment of the primary outcome. In some cases this will be the first phase of the substantive study and data from the pilot phase may contribute to

the final analysis; this can be referred to as an internal pilot. Or at the end of the pilot study the data may be analyzed and set aside, a so-called external pilot."

Since the focus of this guidance is purely on external (stand-alone) pilot studies, then the advice relating to feasibility and pilot studies is the same; that any outcome measurements should not be used to assess the effectiveness of the intervention. It should also be noted that most journals do not expect to see an assessment of the effectiveness of interventions in articles reporting on feasibility or stand-alone pilot studies. '

This is a feasibility study, and its aim is not to assess effectiveness.

Strengths & Weaknesses

Having followed the recommendations of the 'Feasibility and pilot studies: a guide for NIHR Research Design Service advisors' (Williams, 2016) we provided a complete explanation of why we included the measures. Please refer to the guidance.

'Lack of generalizability' has been included before, however the other reviewer advised us to remove it. It is implied that due to the nature of this study (feasibility and acceptability) the findings will not be generalized.

In addition, this intervention has not been designed to be adapted into the NHS, but to be the start point for developing an alternative route using complementary therapies and brief psychosocial interventions.

In regards to the question 'why Osteopaths'? please refer to our paper entitled 'The effects of osteopathic treatment on psychosocial factors in people with persistent pain: A systematic review' (Saracutu et al., 2017). This paper can provide a good explanation in regards to the rationale.

Introduction

• 'Non negligible' has been replaced

• 'Distinctiveness' has been replaced with 'distinction'

• Depression and anxiety and also catastrophyzing and fear avoidance have been moved

• the spelling of catastrophyzing is now consistent

Methods

'to ' replaced by 'for'

We included EQ5D, which is a measure of health status and includes dimensions such as: self-care, mobility, daily activities, pain/ discomfort, depression/ anxiety. We will include a measure of pain in a future trial. This has been mentioned in the discussion.

In addition, given certain experiences reported by people living with pain ('brain fog', the need to move around, discomfort when sitting for long periods of time) we decided not to overburden them.

Expected outcomes

Please refer to the previous response in regards to effectiveness (Abstract). This study is not looking at effectiveness.

In regards to feasibility and acceptability, please read pages 15-17 of the manuscript, where definitions of feasibility and acceptability are provided.

Feasibility and acceptability will be assessed mainly by analysing the qualitative data from the one-to-one interviews and focus group. The main focus is on determining whether participants and those delivering the intervention consider it appropriate and acceptable.

Other data will also be considered (time taken to fill in questionnaires, missing data, follow-up responses, number of people referred and eligible, number of people attending initial interview). However, this will only intend to complement the qualitative data. Recruitment and retention rate have been deleted.

This is in line with Sekhon et al (2017) :

'Successful implementation depends on the acceptability of the intervention to both intervention deliverers (e.g. patients, researchers or healthcare professionals) and recipients (e.g. patients or healthcare professionals). From the patient's perspective, the content, context and quality of care received may all have implications for acceptability. If an intervention is considered acceptable, patients are more likely to adhere to treatment recommendations and to benefit from improved clinical outcomes. From the perspective of healthcare professionals, if the delivery of a particular intervention to patients is considered to have low acceptability, the intervention may not be delivered as intended (by intervention designers), which may have an impact on the overall effectiveness of the intervention.'

VERSION 3 – REVIEW

REVIEWER	Lance McCracken King's College London, UK
REVIEW RETURNED	26-Jul-2018

GENERAL COMMENTS	Second line of introduction should read “there is a large percentage....” I have no further comments.

REVIEWER	Julie Ashworth Research Instituite for Primary care and Health Sciences, Keele University
REVIEW RETURNED	30-Jul-2018

GENERAL COMMENTS	The authors have done a good job of taking on board previous reviews and the manuscript has further improved. I think it's nearly there now with some minor corrections as per my comments below Strengths & Weaknesses Capital letter to start the last 2 bullet points of strengths & weaknesses Methods, study design missing ‘and ‘ at end of 2nd sentence Final sentence In addition, quantitative data will be collected at baseline, upon completion of the program and at one-month and three months follow-up to reveal any differences in psychological flexibility, depression, anxiety, fear avoidance and general health status (see Fig.1). Has fig 1 now been removed? – if so need to take this bit out and I recommend saying that quantitative data will be collected (details)……..and "will include" rather than "to reveal differences". This is consistent with my further comments below Participants 3rd sentence I suggest replacing "sample size" with "group size", since that is what the authors are really referring to, otherwise it conflicts with the potential limitation acknowledged around sample size/ desirability of using more than one group Recruitment, 2nd paragraph The following sentence needs a little rewording to make sense “Once the Osteopaths will identify patients who are eligible (by applying inclusion and exclusion criteria), the MI will ask the patients a couple of questions to find out more about their general health and also to reiterate the nature of the intervention and answer questions.” Secondary outcomes The authors have included a sentence which goes some way to acknowledging that the questionnaire measure will only provide preliminary data on outcome measures, but I would prefer something a little more convincing e.g E.g. Data collected to assess the suitability of treatment and process outcome measures will also provide some preliminary data on outcomes……. Planned data analysis quantitative evaluation Consistent with the above, I suggest restricting this to saying data on treatment and process outcome measures at baseline and follow-up with be analysed (give details) ……. and changes described using descriptive statistics (give details)……..

	Discussion The authors have now much more clearly articulated the limitations of their study in response to reviewer feedback. Unfortunately this has led to the potential advantages / strengths of the intervention & the study being buried in the middle of the discussion and, in the spirit of always starting with the positive, I recommend moving things around a bit. E.g. starting the discussion with all the strengths relating to the intervention and feasibility study, before moving on to limitations - which should be grouped together. The added sentence "In addition, future trials may benefit from including measures of pain and functioning" would sit better in with the other limitations leaving the rest of the final paragraph as a nice conclusion where it is.

VERSION 3 – AUTHOR RESPONSE

Reviewer: 1

Thank you for your recommendations.

Second line of introduction should read “there is a large percentage....”
This has been now modified.

Reviewer: 2

Thank you for acknowledging our previous changes. Please find bellow a response to the additional corrections that were suggested.

Strengths & Weaknesses
Capital letter to start the last 2 bullet points of strengths & weaknesses
Methods, study design
missing ‘and ‘ at end of 2nd sentence

This has been modified according to your suggestion.

Final sentence
Has fig 1 now been removed? – if so need to take this bit out and I recommend saying that quantitative data will be collected (details)…….and "will include" rather than "to reveal differences". This is consistent with my further comments below

Figure 1. hasn’t been removed, but is presented separately. BMJ Open requires figures to be submitted in a separate file. The paragraph now reads:

Quantitative data will be collected at baseline, upon completion of the program and at one-month and three months follow-up and will include the following outcomes: psychological flexibility, depression, anxiety, fear avoidance and general health status (see Fig.1).Participants
3rd sentence I suggest replacing "sample size" with "group size", since that is what the authors are really referring to, otherwise it conflicts with the potential limitation acknowledged around sample size/ desirability of using more than one group
‘Sample size‘ has been replaced with ‘group size’

Recruitment, 2nd paragraph
The following sentence needs a little rewording to make sense
“Once the Osteopaths will identify patients who are eligible (by applying inclusion and exclusion criteria), the MI will ask the patients a couple of questions to find out more about their general health and also to reiterate the nature of the intervention and answer questions.”

This paragraph has been replaced with:

‘The Osteopaths will identify patients who are eligible (by applying inclusion and exclusion criteria). Following that, the MI will meet the patients in person in order to find out more about their general health and also to reiterate the nature of the intervention and answer questions. ‘

Secondary outcomes
The authors have included a sentence which goes some way to acknowledging that the questionnaire measure will only provide preliminary data on outcome measures, but I would prefer something a little more convincing e.g
E.g. Data collected to assess the suitability of treatment and process outcome measures will also provide some preliminary data on outcomes…….

We consider that the sentence you are referring to provides a clear and strong rationale for including the outcome measures mentioned.

The secondary outcomes (measured at baseline, upon completing of the program, at one month and three months follow- up) are only meant to provide some preliminary data on outcomes such as depression and anxiety, acceptance of pain, mindfulness, fear avoidance and quality of life.

Planned data analysis quantitative evaluation
Consistent with the above, I suggest restricting this to saying data on treatment and process outcome measures at baseline and follow-up with be analysed (give details) ……. and changes described using descriptive statistics (give details)……..

This has been modified according to your recommendation. There is now more emphasis on descriptive statistics.

‘Preliminary data will be collected on outcomes such as depression and anxiety, acceptance of pain, mindfulness, fear avoidance and quality of life. Descriptive statistics (means, standard deviations) will provide some insight into population characteristics and also an indication of potential changes in mean scores between the four time points (pre-, post intervention and two follow-ups). Within-subjects effects will also be reported (F, degrees of freedom, effect sizes, CI, p values) and presented in a table. ‘

Discussion
The authors have now much more clearly articulated the limitations of their study in response to reviewer feedback. Unfortunately this has led to the potential advantages / strengths of the intervention & the study being buried in the middle of the discussion and, in the spirit of always starting with the positive, I recommend moving things around a bit.
E.g. starting the discussion with all the strengths relating to the intervention and feasibility study, before moving on to limitations - which should be grouped together. The added sentence "In addition, future trials may benefit from including measures of pain and functioning" would sit better in with the other limitations leaving the rest of the final paragraph as a nice conclusion where it is.

The strengths are now presented at the beginning of the discussion and followed by the limitations.