PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Acupuncture for fatigue in breast cancer survivors: a study protocol for a pragmatic, mixed-method, randomised controlled trial
AUTHORS	Alræk, Terje; Skjerve, Hilde; Sørensen, Anette; Lie, S; Presterud Ødegård, Hilde; Lu, Weidong; Mao, Jun; Deng, G; Lee, Myeong Soo; Birch, Stephen; Lamu, Admassu; Kim, Tae-hun; MacPherson, Hugh

VERSION 1 – REVIEW

REVIEWER	Davis, Mellar Geisinger Health System, Palliative Care
REVIEW RETURNED	13-Sep-2023

GENERAL COMMENTS	Critique • This is a proposed randomized study of acupuncture for fatigue in breast cancer survivors with the primary outcome measured by the Chalder Fatigue scale. This is a 14-week trial with acupuncture performed weekly • The exclusion criteria 1 “participants with any known pathology that can explain their fatigue,” is too vague and is left up to interpretation. • Within the study there is no mention of the timing of the blood studies. Are they in the morning? Are they in the afternoon? There are circadian changes in certain inflammatory cytokines. • The Chalder Fatigue Scale has been used to measure population-based fatigue. It has also been used to measure fatigue in multiple chronic illnesses including cancer.1 It has a 2-factor construct of physical and mental fatigue and has an adequate ROC sensitivity and reliability. A score of 29 discriminates the chronic fatigue syndrome from the average population.2 • It has been used to measure fatigue in cancer-related trials. • The Chalder Fatigue Scale changes are associated with major depressive disorder in lung cancer.3 This suggests that the mental fatigue factor may actually be evaluating depression rather than fatigue • This is clearly demonstrated in a randomized trial of a fatigue intervention in cancer patients where the changes in the Chalder Fatigue Scale at 12 weeks reflected depression measured by the CES-D and did not correlate with the VAS of fatigue.4 • The meaningful important change in the scale has not been established in the population to be studied. Therefore monitoring individual responses to treatment and the desired outcome in a responders analysis is not possible using this tool. • The investigators proposed in their power analysis to use a study of patients with irritable bowel had fatigue after Giardiasis using a distributional analysis rather than an anchor PRO analysis of

	change with a timeframe of 3 years rather than 14 weeks as the authors proposed in their study. This does not make sense at all.5 • Therefore, this study is fatally flawed for 3 reasons. The fatigue instrument measures depression rather than fatigue in the proposed cancer population. The instrument does not have an established meaningful, important change and thus may tell differences in groups statistically but cannot tell differences in individual patients for a responders analysis. The power analysis was done for a distinctly different population and timeframe which makes no sense 1. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. J Psychosom Res. 1993;37(2):147-153. 2. Cella M, Chalder T. Measuring fatigue in clinical and community settings. J Psychosom Res. 2010;69(1):17-22. 3. Maneeton B, Maneeton N, Reungyos J, Intaprasert S, Leelarphat S, Thongprasert S. Prevalence and relationship between major depressive disorder and lung cancer: a cross-sectional study. Onco Targets Ther. 2014;7:815-821. 4. Tamada S, Ebisu K, Yasuda S, et al. Kamikihito improves cancer-related fatigue by restoring balance between the sympathetic and parasympathetic nervous systems. Prostate Int. 2018;6(2):55-60. 5. Hanevik K, Wensaas KA, Rortveit G, Eide GE, Morch K, Langeland N. Irritable bowel syndrome and chronic fatigue 6 years after giardia infection: a controlled prospective cohort study. Clin Infect Dis. 2014;59(10):1394-1400.

REVIEWER	Grant, Suzanne Western Sydney University, NICM Health Research Institute
REVIEW RETURNED	17-Sep-2023

GENERAL COMMENTS	Abstract Line 22-27 These are two separate ideas, therefore two separate sentences. The abstract introduction could benefit from a stronger rationale – “used internationally” is vague. I think that the abstract is overall quite disconnected. Methods: opening sentence very clear. Study assessment timepoints are not clear (run on sentence). Perhaps separate the last sentence of the methods into two sentences. No mention of the inclusion criteria. The strengths seem disconnected from the introduction. The introduction highlights the need for a study that addresses comorbidities, and questionnaires that address different aspects of fatigue. Biomarkers are being measured – not mentioned in the methods section. Introduction Suggest replacing “patients” with people with cancer (or can use people affected by cancer if including carers/family). Last few sentences could be synthesised to make phrasing tighter.

	Suggest adding a few sentences on effectiveness of current treatments and side effects etc – polypharmacy. Suggest synthesising the conclusions of the review rather than quoting. The research questions do not need to have the answers – this is obvious in the methods. Generally, the language of the whole paper needs tightening, overly verbose and reads more like a thesis than a manuscript/publication. Method In general, the Methods section is overly discursive and some points would be best kept in the Introduction section as rationale. No mention that the protocol has been reported according to the SPIRIT guidelines. The first paragraph is extraordinarily long and would benefit from separation into key concepts. Lines 31-42: This appear to relate to the pilot study not this study? Discussion/rationale on the choice of pragmatic study design does not belong in this section, perhaps introduction. Inclusion and exclusion criteria: including those on endocrine therapy or immunotherapy; metastatic disease? Rationale for using a VAS scale for eligibility – perhaps for screening but suggest that the BFI or Multidimensional Fatigue Symptom Inventory be considered. This will at least provide some comparability to other acupuncture and CRF studies about the characteristics of your study population. The Chalder Fatigue Scale is not commonly used to my knowledge. There is no maximum time limit given for time with cancer related fatigue? Or tme since cancer diagnosis? This may need to be considered in the sample size calculation – those with longer more protracted CRF may be less responsive? Intervention group: I think that the use of a whole acupuncture intervention is commendable. I would suggest that the authors highlight this in their intervention. Acupuncturists experience in treating people with cancer? Recording the initial TCM diagnosis? Qualitative studies: The rationale for exploring COVID-19 and CRF are unclear. This would also mean that the study participants will have to have been diagnosed CRF prior to the COVID-19? Or at least have been diagnosed and treated for the cancer prior to this? Outcomes measures: I note that this study is underway. However, I wonder about the participant burden of completing these measures, along with the potential for fatigue, responder disengagement particularly when the questionnaires overlap. Could the authors please provide a justification for the number of different questionnaires, and consider how they overlap. Eg the EORTC will also ask about sleep and fatigue etc

	Suggest that researchers also consider comorbidities along with BMI and body fat as effect modifiers. Cost effectiveness unclear if this data will be collected by access to the patient electronic medical record, self-report using a standard, validated questionnaire. Spell out NICE. The authors reported earlier in the paper that the FSS was selected for it's utility in estimating QALYs. However, this is not mentioned in this section and it would appear that the EQ-5D will be used instead. Statistical power Refer to the FQ – not mentioned earlier, assume this means the Chalders Fatigue Scale? Is a 10% decrease on this Scale considered clinically significant? User involvement Was a user involved in the initial pilot study or this expanded study design?

REVIEWER	Guo, Taipin Yunnan University of Chinese Medicine
REVIEW RETURNED	17-Nov-2023

GENERAL COMMENTS	Fatigue symptoms of breast cancer are common in clinic, and there is no particularly good chemical drug treatment at present. Acupuncture has been shown to have good clinical effect, but there is still a lack of rigorous and feasible design of randomized controlled trials, so this study has good clinical significance. There were no obvious flaws in the design of this study, and it would be more convincing if another sham acupuncture group could be added as a control group.

VERSION 1 – AUTHOR RESPONSE

Response to the Reviewer #1's Comments

Dr. Mellar Davis, Geisinger Health System

Comments to the Author:

Critique

• This is a proposed randomized study of acupuncture for fatigue in breast cancer survivors with the primary outcome measured by the Chalder Fatigue scale. This is a 14-week trial with acupuncture performed weekly

• The exclusion criteria 1 "participants with any known pathology that can explain their fatigue," is too vague and is left up to interpretation.

• Within the study there is no mention of the timing of the blood studies. Are they in the morning? Are they in the afternoon? There are circadian changes in certain inflammatory cytokines.

Response: The blood study/Biomarker is a nested study within the present RCT. Hence it has its own project description, including standard operating procedures (SOP) and a separate invitation

letter for the participants (n=80; 40 from the intervention group and 40 from the control group)). All documents have been approved by the Regional Ethics Committee. Due to word limitation in the present manuscript, this is not included. Maybe it can be included as a supplement if the BMJ Open editorial board appreciate that solution? When the results from the Biomarker study are ready for publication all information regarding blood samples and their SOP will be described in detail.

• The Chalder Fatigue Scale has been used to measure population-based fatigue. It has also been used to measure fatigue in multiple chronic illnesses including cancer.

1 It has a 2-factor construct of physical and mental fatigue and has an adequate ROC sensitivity and reliability. A score of 29 discriminates the chronic fatigue syndrome from the average population.

2 It has been used to measure fatigue in cancer-related trials.

The Chalder Fatigue Scale changes are associated with major depressive disorder in lung cancer.

3 This suggests that the mental fatigue factor may actually be evaluating depression rather than fatigue This is clearly demonstrated in a randomized trial of a fatigue intervention in cancer patients where the changes in the Chalder Fatigue Scale at 12 weeks reflected depression measured by the CES-D and did not correlate with the VAS of fatigue.

4 The meaningful important change in the scale has not been established in the population to be studied. Therefore monitoring individual responses to treatment and the desired outcome in a responders analysis is not possible using this tool. The investigators proposed in their power analysis to use a study of patients with irritable bowel had fatigue after Giardiasis using a distributional analysis rather than an anchor PRO analysis of change with a timeframe of 3 years rather than 14 weeks as the authors proposed in their study. This does not make sense at all

5 Therefore, this study is fatally flawed for 3 reasons. The fatigue instrument measures depression rather than fatigue in the proposed cancer population. The instrument does not have an established meaningful, important change and thus may tell differences in groups statistically but cannot tell differences in individual patients for a responders analysis. The power analysis was done for a distinctly different population and timeframe which makes no sense

1. Chalder T, Berelowitz G, Pawlikowska T, et al. Development of a fatigue scale. J Psychosom Res. 1993;37(2):147-153.

2. Cella M, Chalder T. Measuring fatigue in clinical and community settings. J Psychosom Res. 2010;69(1):17-22.

3. Maneeton B, Maneeton N, Reungyos J, Intaprasert S, Leelarphat S, Thongprasert S. Prevalence and relationship between major depressive disorder and lung cancer: a cross-sectional study. Onco Targets Ther. 2014;7:815-821.

4. Tamada S, Ebisu K, Yasuda S, et al. Kamikihito improves cancer-related fatigue by restoring balance between the sympathetic and parasympathetic nervous systems. Prostate Int. 2018;6(2):55-60.

5. Hanevik K, Wensaas KA, Rortveit G, Eide GE, Morch K, Langeland N. Irritable bowel syndrome and chronic fatigue 6 years after giardia infection: a controlled prospective cohort study. Clin Infect Dis. 2014;59(10):1394-1400.

•

Response to reviewer 1, especially concerning the reviewers' comment from 2-5:

We appreciate the thoughtful comments of reviewer 1, concerning our use of Chalder Fatigue Scale (CFS) in the study in question. However, as we have written in our manuscript this scale is used within The Norwegian Cancer Registry for registering fatigue in people with breast cancer. This is also reflected in recent publications (1). The scale has shown to be responsive to changes of fatigue levels and further, it is also recommended in multidimensional registration of fatigue in people with cancer (2). Also, the American Cancer Society, refer to Chalder as a useful tool for measuring scores for physical and mental fatigue in people with cancer related fatigue. (3). Further, in the present study, we have several secondary outcomes which influence fatigue and some of them register fatigue itself e.g., EORTC-QLQ 30. This allows us to compare the outcome

from different questionnaires regarding cancer related fatigue. The Norwegian Cancer Registry use EORTC-QLQ 30 also, hence this gives us the opportunity to compare data from the participants in our study with the similar available data from the registry.
In the study by Hanevik et al. (2014) they show a clinically relevant difference of the FQ-scale. We adapted this difference as a relevant difference (on group level) also for the present study and based the sample-size calculation on this difference. The heterogeneity in the present population is larger (or smaller) than in Hanevik et al. Hence, the standard deviation may be larger (or smaller) then we assumed, which would lead to an unprecise sample-size calculation, irrespective of the follow-up time. We observe that MID for CFS have been discussed for other patient populations (4) and we agree that MIDs would be of interest. However, we argue that establishing MID for the present population is beyond the scope of the present study. Both issues will be discussed in the final publication in the relation to fatigue. Since, several instruments are used to measure fatigue of the secondary outcomes we feel that this is sufficiently covered.
One of them, as reviewer 1 refers to is depression (and anxiety). This can be a major problem for cancer survivors. Hence, we think we can address that since we have included HADS in our study. That gives us e.g., an opportunity via subgroup analysis to look at any correlation between fatigue and depression in our study.

1. Svendsen K, Nes LS, Meland A, Larsson IM, Gjelsvik YM, Børøsund E, Rygg CM, Myklebust TÅ, Reinertsen KV, Kiserud CE, Skjerven H, Antoni MH, Chalder T, Mjaaland I, Carlson LE, Eriksen HR, Ursin G. Coping After Breast Cancer: Protocol for a Randomized Controlled Trial of Stress Management eHealth Interventions. JMIR Res Protoc. 2023 May 24;12:e47195. doi: 10.2196/47195.
2. Stone PC, Minton O. Cancer-related fatigue. Eur J Cancer. 2008 May;44(8):1097-104. doi: 10.1016/j.ejca.2008.02.037.
3. Berger AM, Mitchell SA, Jacobsen PB, Pirl WF. Screening, evaluation, and management of cancer-related fatigue: Ready for implementation to practice? CA Cancer J Clin. 2015 May-Jun;65(3):190-211. doi: 10.3322/caac.21268.
4. Pouchot J, Kherani RB, Brant R, Lacaille D, Lehman AJ, Ensworth S, Kopec J, Esdaile JM, Liang MH. Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. J Clin Epidemiol. 2008 Jul;61(7):705-13. doi: 10.1016/j.jclinepi.2007.08.016

Reviewer: 2
Dr. Suzanne J Grant, Western Sydney University
Comments to the Author:
The authors should be commended for this comprehensive study protocol. My overall impression is that there is room for improvement. Please see the detailed comments:
Response to the Reviewer #2's Comments

Abstract
Line 22-27 These are two separate ideas, therefore two separate sentences.
Response: Changes have been made (page 1, line27)

The abstract introduction could benefit from a stronger rationale – "used internationally" is vague. I think that the abstract is overall quite disconnected.
Response: The rationale and the abstract have been rephrased (page 1, line 8 – 16)

Methods: opening sentence very clear. Study assessment timepoints are not clear (run on sentence).
Response: Changes have been made page 1, line 25-26.

Perhaps separate the last sentence of the methods into two sentences.
Response: Changes have been made accordingly, page 1, line 27.

No mention of the inclusion criteria.
Response: We think that our Abstract as it reads are in line with the description as given by BMJ Open.

The strengths seem disconnected from the introduction. The introduction highlights the need for a study that addresses comorbidities, and questionnaires that address different aspects of fatigue.

Biomarkers are being measured – not mentioned in the methods section.
Response: As described to reviewer 1:
The blood study/Biomarker is a nested study within the present RCT. Hence it has its own project description, including standard operating procedures (SOP) and a separate invitation letter for the participants (n=80). All documents have been approved by the Regional Ethics Committee. Due to word limitation in the present manuscript, this is not included. **Maybe it can be included as a supplement if the BMJ Open editorial board appreciate that solution?** When the results from the Biomarker study are ready for publication all information regarding blood samples and their SOP will be described in detail.

Introduction
Suggest replacing "patients" with people with cancer (or can use people affected by cancer if including carers/family).
Response: The suggestion has been followed and changes have been made throughout the manuscript. We have now used "participants in the study" instead of "patients in the study"

Last few sentences could be synthesised to make phrasing tighter.
Suggest adding a few sentences on effectiveness of current treatments and side effects etc – polypharmacy.
Response: Due to word limitation, this is not doable. However, we will describe current treatments, polypharmacy, and potential side effects thereof, when we publish our article of the primary outcome.

Suggest synthesising the conclusions of the review rather than quoting.
Response: This has been done, page 2, line 62-65.

The research questions do not need to have the answers – this is obvious in the methods. Generally, the language of the whole paper needs tightening, overly verbose and reads more like a thesis than a manuscript/publication.
Response: The answers to the research questions have been removed (page 3, lines 87-89, 90-94 and 96-98). Our English, native speaking co-author has addressed the reviewer grammatical/language issues and changes have been made throughout the manuscrit.

Method
In general, the Methods section is overly discursive and some points would be best kept in the Introduction section as rationale.
Response: Text have been moved to the Introduction (page 3, line 79-92).

No mention that the protocol has been reported according to the SPIRIT guidelines.
Response: SPIRIT guidelines have been followed and is now added to the manuscript (page 4, line 132 – 133).

The first paragraph is extraordinarily long and would benefit from separation into key concepts.
Response: The long sentence has been divided into two (page 4, line 123).

Lines 31-42: This appear to relate to the pilot study not this study?
Response: The large RCT (n=250) is an ongoing study, it was registered in the NCT before it started, however there has been delays due to the Covid 19 pandemic in Norway.

Discussion/rationale on the choice of pragmatic study design does not belong in this section, perhaps introduction.
Response: The rationale for having a pragmatic study design has been moved to the Introduction.

Inclusion and exclusion criteria: including those on endocrine therapy or immunotherapy; metastatic disease? Rationale for using a VAS scale for eligibility – perhaps for screening but suggest that the BFI or Multidimensional Fatigue Symptom Inventory be considered. This will at least provide some comparability to other acupuncture and CRF studies about the characteristics of your study population.
Response: Since the AcuBreast is an ongoing study, we are not in the position to use the suggested methods for screening. All participants must have finished any form of curative cancer treatment, description of the cancer treatments they have received will be registered.

The Chalder Fatigue Scale is not commonly used to my knowledge.
Response: The Norwegian Cancer Registry uses Chalder (FQ) for registering fatigue in people with breast cancer. The scale has shown to be responsive to changes of fatigue levels and further, it is also recommended in multidimensional registration of fatigue in cancer.

There is no maximum time limit given for time with cancer related fatigue? Or time since cancer diagnosis? This may need to be considered in the sample size calculation – those with longer more protracted CRF may be less responsive?
Response: At baseline the time for cancer diagnosis, the cancer treatments received, and the length of cancer related fatigue will be registered. Hence, we can address the issues as mentioned by reviewer 2.

Intervention group: I think that the use of a whole acupuncture intervention is commendable. I would suggest that the authors highlight this in their intervention.
Response: This is described, page 6, line 202 – 209. Acupuncturists experience in treating people with cancer?

Recording the initial TCM diagnosis?
Response: Yes, and a sentence has been added with that information.

Qualitative studies: The rationale for exploring COVID-19 and CRF are unclear. This would also mean that the study participants will have to have been diagnosed CRF prior to the COVID-19? Or at least have been diagnosed and treated for the cancer prior to this?
Response: The focus of this qualitative study was to get knowledge about how the **lock-down** of the society, had any effect on their CRL. Hence, this has been clarified in the manuscript.

Outcomes measures: I note that this study is underway. However, I wonder about the participant burden of completing these measures, along with the potential for fatigue, responder disengagement particularly when the questionnaires overlap. Could the authors please provide a justification for the number of different questionnaires, and consider how they overlap. Eg the EORTC will also ask about sleep and fatigue etc.

Response: The participant's burden was addressed in the study information included in the signed informed consent. Here we described that filling in the questionnaire was a time-consuming effort. Hence, we advise them to take breaks in between. A total potential timeframe was 45-50 minutes. This description was in line with the suggestion from Regional Ethics Committee.

Suggest that researchers also consider comorbidities along with BMI and body fat as effect modifiers.
Response: This is not doable since the present study is an ongoing study.

Cost effectiveness unclear if this data will be collected by access to the patient electronic medical record, self-report using a standard, validated questionnaire. Spell out NICE.
Response: Thank you for the comments. For cost-effectiveness analyses, we consider costs and consequences (outcomes) in two comparison groups: the intervention (acupuncture treatment) and the control (usual treatment without acupuncture). EQ-5D is the most widely used instrument for such economic evaluation, and recommended by the health authority in several countries, including NICE (The National Institute for Health and Care Excellence). The Norwegian Medicine Agency also recommends EQ-5D for a single technology assessment. It is a ***generic*** preference-based measure of health that is dominantly used for QALY calculation that enables comparison across diseases or alternative interventions. Thus, self-reported EQ-5D has been used as an intervention outcome in cost-effectiveness analysis. This has now been revised in the manuscript. Patient electronic medical records will not be used as health outcome measure for cost-effectiveness.

The authors reported earlier in the paper that the FSS was selected for it's utility in estimating QALYs. However, this is not mentioned in this section and it would appear that the EQ-5D will be used instead.
Response: Thank you for spotting this. We have now revised, and only EQ-5D will be used as a health outcome for QALY calculation. In line with this, the sentence: *"As a sensitivity analysis, we also converted FSS scores to health state utility values to estimate QALY's for use in cost-effectiveness analyses"* has been removed from page 12, line 348.
FSS is a symptom-specific non-preference instrument, and hence can't be used for QALY calculation.

Statistical power Refer to the FQ – not mentioned earlier, assume this means the Chalders Fatigue Scale?
Response: FQ is a modified version of the Chalder FS, 3 questions are removed from the original Chalder FS. For power calculation we have used the modified version, hence we have rephrased the sentence on page 9, line 272-273.
Is a 10% decrease on this Scale considered clinically significant?
Response: Yes

User involvement Was a user involved in the initial pilot study or this expanded study design? **Response:** User involvement was present in the pilot study and as described in the manuscript. Based on that we made necessary changes for the present ongoing study. Furthermore, we have users' representatives in the ongoing study. Described under the heading "Patient and Public Involvement" page 10-11, line 307 – 321.

Reviewer: 3
Dr. Taipin Guo, Yunnan University of Chinese Medicine
Comments to the Author:
Fatigue symptoms of breast cancer are common in clinic, and there is no particularly good chemical drug treatment at present. Acupuncture has been shown to have good clinical effect, but there is still a

lack of rigorous and feasible design of randomized controlled trials, so this study has good clinical significance. There were no obvious flaws in the design of this study, and it would be more convincing if another sham acupuncture group could be added as a control group.

Response to the Reviewer #3's Comments

Response: Several of the authors, in the present study, have been involved in discussions around the problems withsham controlled trials, hence we decided to go for a pragmatic RCT. We have published several papers about the problems of using sham acupuncture in controlled trials:

1. Lee B, Kwon CY, Lee HW, Nielsen A, Wieland LS, Kim TH, Birch S, Alraek T, Lee MS. Needling Point Location Used in Sham Acupuncture for Chronic Nonspecific Low Back Pain: A Systematic Review and Network Meta-Analysis. JAMA Netw Open. 2023 Sep 5;6(9):e2332452. doi: 10.1001/jamanetworkopen.2023.32452.
2. Birch S, Lee MS, Kim TH, Alraek T. On defining acupuncture and its techniques: A commentary on the problem of sham. Integr Med Res. 2022 Jun;11(2):100834. doi: 10.1016/j.imr.2022.100834.

If the reviewer wants us to include more information about why we did not use a sham controlled design, that is doable in review, but journal space does not permit the extensive documentation we can provide placed in the article.

VERSION 2 – REVIEW

REVIEWER	Grant, Suzanne Western Sydney University, NICM Health Research Institute
REVIEW RETURNED	11-Apr-2024

GENERAL COMMENTS	Minor suggestions: Abstract Suggest this is evident in previous sentence: "This study will address these issues and it is the first study with acupuncture as intervention in breast cancer survivors with fatigue in a Norwegian population". Change to "This study of acupuncture as an invention in ….is the first to be conducted in a Norwegian population". In abstract – methods please briefly mention the biomarkers sub-study. Page 3 Delete " and worth paying for it?" this is implicit in cost effectiveness methodology. These sentences are a little unclear: "The pilot study was run to examine the feasibility of our approach, which is like the current fully powered multi-center, pragmatic randomized controlled study. Hence, the design and methodological approach of the study in question are similar to the pilot study. By doing this we have been able to evaluate the feasibility of recruitment, randomization, collaboration, assessment procedures, and implementation of acupuncture." Suggest rephrasing for simplification to "The pilot study examined the feasibility of our the design and methodological approach, recruitment, randomization, collaboration, assessment procedures, and implementation of acupuncture. Based on the findings of the pilot study, we have developed a fully powered multi-center, pragmatic randomized controlled study outlined in this study

	protocol. The present study was modified to address challenges encountered in the pilot study including..” Page 4 Please include that this is a mixed method study in the opening sentence on method, and in the abstract. Page 5. Consider mentioning that the study will be reported according to the CONSORT extension for pragmatic trials. Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., ... & Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. Bmj, 337. Page 10. Is it next of kin? Or carers? Page 11 Last sentence of first paragraph does not make sense. Page 12 You have not outlined data analysis for the qualitative component of the study P14 Some minor grammatical changes needed in this paragraph: “The late, Hugh MacPherson, was a colleague and close friend to most of the authors of the present protocol. The first author (TA) invited Hugh into be a member of the research group “AcuBreast”. By this Hugh played an important role in the development of the submitted protocol submitted to the Pink ribbon in 2019, and was successfully funded through this., which received funding. Hence, we would like to pay a tribute to his work therefore he is listed as the last author. We would also like to acknowledge, tThe two previous project coordinators of AcuBreast, have had two project coordinators, Merete Lindén Dahle and Mona Solberg. The present study will not have been possible without their help.”

REVIEWER	Guo, Taipin Yunnan University of Chinese Medicine
REVIEW RETURNED	27-Jan-2024

GENERAL COMMENTS	Thanks for the authors' feedback, the content of the study on this disease has good clinical significance. However, the acupuncture intervention part of the article seems to be different from the traditional design, and does not specify specific acupuncture points and acupuncture operations, etc., but describes acupuncture treatment in accordance with the real world. I hope the author can explain the necessary reasons here. Finally, it is suggested to record the operation plan of each point selection for each patient.

VERSION 2 – AUTHOR RESPONSE

Minor suggestions from reviewer 2: Dr. Suzanne Grant, Western Sydney University

Comments to the Author:
Thank for addressing my previous comments comprehensively. The manuscript is greatly improved.

Response: Dear reviewer, thank You very much for your constructive and useful comments. I do hope we have met your suggestions for to reach a better manuscript.

I agree that the inclusion of the biomarker sub-study as a supplement would be very useful. However, this will require some attention to grammar.

Response: With regard to your comment on the grammar, our manuscript has now been edited by a professional copyediting service and they have also been editing the supplementary protocol of the biomarker sub-study.

Response to the specific comments attached from reviewer 2

Abstract Suggest this is evident in previous sentence: "This study will address these issues and it is the first study with acupuncture as intervention in breast cancer survivors with fatigue in a Norwegian population". Change to "This study of acupuncture as an invention in ….is the first to be conducted in a Norwegian population".

Response: Suggested changes have been made in the abstract, page 2, line 8-9

In abstract – methods please briefly mention the biomarkers sub-study.

Response: The biomarker sub-study is now mentioned on page 2-3, line 28-31.

Page 3 Delete " and worth paying for it?" this is implicit in cost effectiveness methodology.

Response: Deleted

These sentences are a little unclear: "The pilot study was run to examine the feasibility of our approach, which is like the current fully powered multi-center, pragmatic randomized controlled study. Hence, the design and methodological approach of the study in question are similar to the pilot study. By doing this we have been able to evaluate the feasibility of recruitment, randomization, collaboration, assessment procedures, and implementation of acupuncture." Suggest rephrasing for simplification to "The pilot study examined the feasibility of our the design and methodological approach, recruitment, randomization, collaboration, assessment procedures, and implementation of acupuncture. Based on the findings of the pilot study, we have developed a fully powered multi-center, pragmatic randomized controlled study outlined in this study protocol. The present study was modified to address challenges encountered in the pilot study including.."

Response: The chapter "Experiences from a pilot study" has now been rephrased in accordance with your suggestions, page 5-6, line 114-128

Page 4 Please include that this is a mixed method study in the opening sentence on method, and in the abstract.

Response: The two suggested inclusion of the “mixed method” have been placed at page 2, line 16 and page 6, line 131.

Page 5. Consider mentioning that the study will be reported according to the CONSORT extension for pragmatic trials. Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., ... & Moher, D. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. Bmj, 337.

Response: On page 8, line 205 we have introduced STRICTA guidelines, which is an official extension to the CONSORT Statement. Does the reviewer still want us to also include the article by Zwarenstein M, et al.?

Page 10. Is it next of kin? Or carers?

Response: It shall read: next of kin, as it written now on page 13, line 336

Page 11 Last sentence of first paragraph does not make sense.

Response: we totally agree, and the sentence has now been deleted page 13, line 346-347

Page 12 You have not outlined data analysis for the qualitative component of the study

Response: The requested information has now been added on page 15, line 405-410

P14 Some minor grammatical changes needed in this paragraph: “The late, Hugh MacPherson, was a colleague and close friend to most of the authors of the present protocol. The first author (TA) invited Hugh into be a member of the research group “AcuBreast”. By this Hugh played an important role in the development of the submitted protocol submitted to the Pink ribbon in 2019, and was successfully funded through this., which received funding. Hence, we would like to pay a tribute to his work therefore he is listed as the last author. We would also like to acknowledge, The two previous project coordinators of AcuBreast, have had two project coordinators, Merete Lindén Dahle and Mona Solberg. The present study will not have been possible without their help.”

Response: This chapter has now been rephrased and is named: “Special Acknowledgment”. Beside Your suggestions we have made other changes as well.

Reviewer: 3 Dr. Taipin Guo, Yunnan University of Chinese Medicine

Comments to the Author:
Thanks for the authors' feedback, the content of the study on this disease has good clinical significance.

Dear reviewer, thank you very much for you constructive and useful comments.

However, the acupuncture intervention part of the article seems to be different from the traditional design, and does not specify specific acupuncture points and acupuncture operations, etc., but describes acupuncture treatment in accordance with the real world. I hope the author can explain the necessary reasons here.

Response: Thank you very much for these important comments. I do hope we addressed them with the following clarification. As we have written our study is based on TCM (page 8 line 203), we have now added the information that we will report our acupuncture treatment according to the STRICTA guidelines. “These guidelines, which has become an official extension to the CONSORT Statement, were developed to improve the completeness and transparency of reporting in acupuncture research.

Hence, they will allow readers to understand, interpret, and evaluate the results of a clinical trial by providing clear, precise, and comprehensive details about the acupuncture intervention", see page 8, line 205-209. Real world is a description that can be used when a clinical acupuncturist uses relevant knowledge from e.g. education, experiences, and research. Hence, the participating acupuncturists provide their care as in normal practice.

Finally, it is suggested to record the operation plan of each point selection for each patient.

Response: For each patient we will keep an electronic TCM diagnostic sheet. This will be filled in by the treating acupuncturist, hence the selection of acupuncture points will be described and any reasons for changing points will be recorded during the 12 treatments. This is described on page 8-9, line 209-211