

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)	Experiences of fatigued persons with multiple sclerosis with multimodal agility-based exercise training and the ReFEx study protocol. A qualitative extension of a feasibility study.
AUTHORS	Wolf, Florian; Folkerts, Ann-Kristin; Zimmer, Philipp; Nielsen, Jörn

VERSION 1 – REVIEW

REVIEWER	Flores, Victoria A. University of Illinois Chicago, Kinesiology and Nutrition
REVIEW RETURNED	17-Aug-2023

GENERAL COMMENTS	<p>This was an insightful and interesting qualitative research study. The topic is of great interest, and the design is novel in that MAT was implemented in an inpatient setting. The study's strengths lie in its innovative focus on a MAT framework, combined with traditional strength and endurance training, to provide insight into pwMS. The study captured the nuanced perceptions, both physical and psychological, as well as barriers and facilitators of MAT. The real-world inpatient rehabilitation environment in Germany, makes it applicable to clinical practice. High adherence rates, fatigue analysis, and inclusion of moderately-fatigued participants make this study highly relevant. Please see the following to make the delivery and clarity stronger for this much needed research topic.</p> <p>Abstract:</p> <p>Line 6 is confusing; is the overburdened, feeling pressured, and expectations for "traditional" strength training felt in the MAT group, or in the SET group? The MAT did not involve traditional strength and endurance training. Rewrite to be clear on what facilitators were observed per group.</p> <p>Abstract conclusions: A part of the listed objectives was to identify possible adaptations for a powered RCT - was there none found? Was the social comparison and negative self-evaluation too great of a barrier for a future high powered RCT? Insert a concluding sentence that addresses this part of the objectives to let readers know why the case to continue research in this area is strong/promising.</p> <p>Introduction:</p> <p>Page 4. Line 32: Cite which types of exercise interventions have been evaluated, along with their parameters (i.e., what type of exercise (aerobic, resistance, flexibility, balance, combined), intervention length, intervention delivery). It is important to note this as it needs to convince researchers of the importance of exercise in MS treatment.</p> <p>Line 35: Cite which studies explicitly tried to target fatigue, do not simply base this on Moss-Morris et al., 2021.</p>
-------------------------	--

Line 38: Please be more specific which studies in the cited systematic review contributed to the large effect size and are small in sample size.

Page 5. Line 3: Please cite the correct references pertaining to larger effect sizes found from balance training.

Line 42: replace 'eight' with '8', as later in the manuscript explained rates are numbers and not spelled out (i.e., 3x/week).

Please include hypotheses regarding objectives (i) and (ii), what was expected to be observed in fatigued pwMS? What were anticipated demands with ReFEx? This will better direct the readers to understand the thought process of the authors in investigating this topic and objectives.

Methods:

Page 7. Line 8: Please spell out 'max.' and where 'max.' appears throughout the manuscript. Lines 45-47: Please clarify in what way NRC patients that joined the SET group were changing (i.e., changing exercise? changing programs?).

Page 9. Line 20: Who was the third researcher, AFK? Please include initials. Also, please detail further the sources, methods, investigators (if others), and theories used to cross-validate the findings (triangulation). If this was not done, please add this as a limitation. Pages 9-10: It would have been arguably stronger evidence if the thematic blocks and coding utilized in the study involved external or independent coders for objectivity in the analysis. Please include this as a limitation.

Page 11. Lines 30-36: Please explain why those conducting the qualitative research did not return to participants with initial findings to verify if they accurately represented their experiences (member checking). (Supp 2 states that there were no repeat interviews). If this was not done, please add as a limitation. Also, please elaborate on why participatory action research (actively involving the participant in the research process, including constructing the interview guide and analysis processes) was not done. This qualitative method would assist in objectives (i) and (ii).

Results:

Table 4: Why have 18 of the 22 completed the surveys? Please clarify in the abstract and results. What does the - sign mean? Not available, not applicable, or was not obtained? Perhaps some clarification on what this symbolizes would be helpful in interpreting the response to surveys.

Discussion:

In going back to the Abstract, what about implications for future RCTs? You do provide implications for future studies (page 24 of 41, Lines 10-11), but it would be beneficial to further elaborate on how these specific findings per group will directly inform the planning and execution of future RCTs.

Please discuss the psychological benefits; most mentioned were physical outcomes; however, the abstract included psychological benefits as well.

The discussion compares the findings with other studies of SET nature; please add more comparisons regarding MAT and fatigue management in pwMS to strengthen the understanding of how this study fits within the field (or challenges the current exercise paradigm for pwMS).

Page 23 of 41, lines 33 to 37. Please include a brief discussion of potential biases in participant responses since the interviews were conducted post-intervention.

	<p>The brief mention of seasonal factors (i.e., heat) is an interesting observation (page 24 of 41, lines 27-28); please elaborate on how they may affect exercise ability and adherence, and how this could be anticipated for future research.</p> <p>References: Of the 36 citations, only 5 are very recent (specifically 2022). There is no updated research in 2023. Please update references to include 2023 RCTs comparable with the objectives and multimodal training protocol (IF there are recent ones).</p>
--	--

REVIEWER	Yu, Laikang Beijing Sport University, Department of Sports Performance
REVIEW RETURNED	25-Aug-2023

GENERAL COMMENTS	<p>This manuscript aims to explore the experiences of fatigued persons with multiple sclerosis (pwMS) with a new multimodal agility-based exercise training (MAT) framework and to investigate the demands of the ReFEx (Rehabilitation, Fatigue, and Exercise) study protocol. The topic is very interesting, but from the process and results of the current study, it is not possible to fulfill the purpose.</p> <ol style="list-style-type: none"> 1. The Introduction is disorganized. Too many descriptions of the study process that should have been in the Methods section. 2. There is a very small research group, and I am concerned about the quality of the results and the possibility of drawing conclusions. Specifically, for face-to-face interviews, only 6 participants were included in the MAT group (f:m = 5:1, 2 young and 4 middle-aged) and SET group (f:m = 4:2, 1 young, 2 middle-aged, and 3 elderly), respectively. The gender and age of participants affect how they feel about the training, which makes their answers to the same questions inconsistent. In order to validate the effectiveness of a new training protocol, a large number of homogeneous participants are needed.
-------------------------	---

REVIEWER	Thorne, Sally University of British Columbia, Nursing
REVIEW RETURNED	07-Oct-2023

GENERAL COMMENTS	<p>My focus in this review was on the qualitative methodology. Although I am very familiar with the experience of fatigue in multiple sclerosis, I do not have expertise in physical therapy interventions in this regard.</p> <p>This manuscript describes a feasibility study using qualitative approaches within a larger clinical trial study. The rationale for conducting a qualitative investigation to complement that which can be assessed using more conventional measures is effectively articulated. In this context, the particular nature of the sample involved in this study is well described and justified. The recruitment strategy involved selection from the total population involved in the clinical trial, and allowed for selective invitation to reflect a balanced and appropriately diverse sample. None of those approached declined to be interviewed (which suggests enthusiasm for the study, and perhaps the authors' sense that these individuals would appreciate an opportunity to provide further input. Although such a strategy does raise the possibility that those invited were more likely to be more favorably inclined toward the intervention, the specific focus of the qualitative aspect</p>
-------------------------	---

	<p>means that should not constitute a design problem or compromise the usefulness of the findings generated from this body of interview material.</p> <p>The description of the data collection and analysis process reveals that a number of design steps were reconsidered throughout the process, and as the analysis evolved. This would be considered problematic in some circles, as it would appear that the authors were continuing to refine their understanding of the kinds of qualitatively derived data that would be important to the larger intentions of the wider clinical trials project. What this meant in practice is that the guiding question was refined over time, which I see as consistent with what is common in 'applied qualitative' studies. While the authors did not explicitly cite such a methodological approach as their intended direction, and it might be appropriate to consider that for another round should they take this feasibility study into a full program, the detailed description of what they did and decided along the way, and why each decision was taken, should be sufficient to allow readers to judge for themselves whether the findings are appropriately or inappropriately shaped by these particular analytic decisions. In my view, they do no harm to the integrity of the study (although readers with a strong grounding in formal (social science) qualitative approaches might disagree).</p> <p>Ultimately, the interviews were quite short (mean of 15 minutes) and seemed guided by a number of predetermined questions. For this reason, the emphasis within the description of analysis (p.11) seems somewhat oddly focused on the number of categories and subcategories that were discerned. What we see when we get to the findings, is that the researchers may well have been simply trying to ascertain what existed within the data set that might be most helpful in guiding future inquiry. When we look to the ultimate decision as to how the findings are displayed and reported, the structure is quite simple, and may well reflect the best possible structure for the intended purpose. Because of this, the extraneous detail about how many groupings were included at each stage along the way may add more confusion than necessary. And again, if the authors were to draw on an explicitly "applied" methodology, then the guidance would make that extra information unnecessary. Rather, focusing on how their thinking evolved through this process to the point that they came to an understanding of what not to report (because it was already well known within the field based on prior literature) and what new insights this study could add to the field. And that is what constitutes meaningful findings.</p> <p>Ultimately the findings presented seem entirely appropriate to the approach used to develop them, and are well aligned with the overall stated purpose within this feasibility phase. The authors are familiar with both the strengths and limitations of using this kind of small qualitative study within a larger clinical trials program. I do note, however, that they imply (p.27) in their discussion of limitations that they might have achieved saturation with a larger study. Since there is some possibility that they may extend this into their larger program of research over time, I would caution against such an implication. Saturation implies that they will have covered all relevant variations, which I suggest is an inappropriate assumption within a study of a complex human phenomenon such as MS fatigue. Rather, their claim would better resonate with</p>
--	---

	<p>clinicians/readers if they acknowledge that no qualitative investigation in such a context can reasonably claim “saturation,” and that the infinite range of possible variation in something like fatigue is what makes it so interesting to work with. There is considerable literature to draw from to support not taking up that social science theorizing concept in the world of clinical research where it is generally inaccurate and misleading.</p> <p>In conclusion, I believe that this feasibility study is worth reporting, and that the findings within it will be of interest, not only to the direction of the future research of this team of investigators, but also to others who are considering including a qualitative component to their clinical trials of evolving interventions for complex human experiences such as fatigue in chronic illness.</p>
--	--

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Dr. Victoria A. Flores, University of Illinois Chicago

Comments to the Author:

This was an insightful and interesting qualitative research study. The topic is of great interest, and the design is novel in that MAT was implemented in an inpatient setting. The study's strengths lie in its innovative focus on a MAT framework, combined with traditional strength and endurance training, to provide insight into pwMS. The study captured the nuanced perceptions, both physical and psychological, as well as barriers and facilitators of MAT. The real-world inpatient rehabilitation environment in Germany, makes it applicable to clinical practice. High adherence rates, fatigue analysis, and inclusion of moderately-fatigued participants make this study highly relevant. Please see the following to make the delivery and clarity stronger for this much needed research topic.

Response 1.1: Thank you for this positive feedback. We are glad to hear that this was a relevant study for you to review and that you acknowledge the innovative character and clinical applicability. We used some of your description for one of the strengths and limitations bullet points (p.5 of 104, I.3-6): The real-world inpatient rehabilitation environment in Germany makes it applicable to clinical practice. Please find our responses to your helpful comments below.

Abstract:

Line 6 is confusing; is the overburdened, feeling pressured, and expectations for "traditional" strength training felt in the MAT group, or in the SET group? The MAT did not involve traditional strength and endurance training. Rewrite to be clear on what facilitators were observed per group.

Response 1.2: Thank you for highlighting this. We agree that this statement was confusing, therefore we changed it as follows: (i) Facilitators regarding MAT were variety and playfulness, group setting, and challenging exercises. Barriers regarding MAT were feeling overburdened, feeling pressured in the group setting, and the wish to perform 'traditional' strength training (not part of MAT) (p. 4 of 103, I.3-11).

Abstract conclusions: A part of the listed objectives was to identify possible adaptations for a powered RCT - was there none found? Was the social comparison and negative self-evaluation too great of a barrier for a future high powered RCT? Insert a concluding sentence that addresses this part of the objectives to let readers know why the case to continue research in this area is strong/promising.

Response 1.3: We did describe adaptations for a future study in the discussion section. However, it is correct that this was missing in the abstract, and we changed the conclusions in the abstract accordingly, while adhering to the word limit: MAT is appreciated by pwMS and includes facilitators less attainable with 'traditional' SET. Evaluation of MAT in a powered RCT is indicated, if rest breaks post-session, and screening for negative self-evaluation and social comparison are considered (p.4 of 103, l.25-35). Also, we further elaborated on how to tackle negative self-evaluation in a future study in the discussion: For a future study, an additional post-session self-rating will be helpful to match pwMS' individual performance levels and needs (e.g., "I felt not challenged enough/overburdened/just right"; in case of feeling overburdened, the therapist should seek further discussion with the participant and might adapt future sessions and re-evaluate the training fit) (p.25 of 103, l.10-21). We think this exemplifies how negative self-evaluation can be dealt with and should not be considered a barrier for a future study.

Introduction:

Page 4. Line 32: Cite which types of exercise interventions have been evaluated, along with their parameters (i.e., what type of exercise (aerobic, resistance, flexibility, balance, combined), intervention length, intervention delivery). It is important to note this as it needs to convince researchers of the importance of exercise in MS treatment.

Response 1.4: We added details on the types of exercise interventions, as requested: Consequently, many exercise interventions have been evaluated, including aerobic, resistance, flexibility, balance, general (i.e., no primary fitness target, such as yoga), and combined exercise, with some being conducted in an aquatic environment [9]. Most interventions have had a duration of 12 weeks or less, but some lasted for up to 26 weeks [9, 10] (p.6 of 103, l.11-18).

Line 35: Cite which studies explicitly tried to target fatigue, do not simply base this on Moss-Morris et al., 2021.

Response 1.5: As the cited meta-analysis included 13 exercise studies that were fatigue-targeted we think it is reasonable to just cite the meta-analysis. However, we did revise the sentence in the manuscript, now citing the meta-analysis and one study relevant for the specific context: A meta-analysis found 13 exercise studies that were explicitly targeted at fatigue[11], but only one small study was based in an inpatient rehabilitation setting[12] (p.6 of 103, l.20-26).

Line 38: Please be more specific which studies in the cited systematic review contributed to the large effect size and are small in sample size.

Response 1.6: We meant that studies are small in number in this category. We clarified this in the manuscript and cited the balance studies with a large effect: Generally, interventions broadly focused on 'balance' have shown large effects[13-15] but the number of existing studies is small[9, 11] (p.6 of 103, l.25-30).

Page 5. Line 3: Please cite the correct references pertaining to larger effect sizes found from balance training.

Response 1.7: The references now include the meta-analysis making this conclusion and the two fatigue-targeted studies from the balance category (p.6 of 103, l.52-55). We chose not to cite the Harrison et al., 2021 meta-analysis at this point even though it includes more studies in the balance category, because these are very heterogenous, and several studies included content which does not match the present intervention.

Line 42: replace 'eight' with '8', as later in the manuscript explained rates are numbers and not spelled out (i.e., 3x/week).

Response 1.8: Revised.

Please include hypotheses regarding objectives (i) and (ii), what was expected to be observed in fatigued pwMS? What were anticipated demands with ReFEx? This will better direct the readers to understand the thought process of the authors in investigating this topic and objectives.

Response 1.9: We agree that hypotheses regarding our aims might direct the reader in a certain way. On the other hand, in qualitative studies (especially in the context of feasibility) it can be a value to be open for both positive and negative and neutral aspects, which is the reason why we did not describe discrete hypotheses a priori. However, we did discuss three relevant topics from the literature for the qualitative analysis in the introduction, which guided the construction of the interview guide.

Methods:

Page 7. Line 8: Please spell out 'max.' and where 'max.' appears throughout the manuscript.

Response 1.10.: Changed accordingly.

Lines 45-47: Please clarify in what way NRC patients that joined the SET group were changing (i.e., changing exercise? changing programs?).

Response 1.11.: Thank you for this question. We elaborated on this aspect more extensively in the methods section: Cycle ergometer sessions were running all day in the clinic. Thus, the SET participants had flexible schedules and trained together with around five other patients from the NRC, who did not participate in the study. Due to the flexible scheduling, participants in the cycling sessions were changing from day-to-day, and therefore, the cycling sessions for the SET-participants did not occur in a closed group (p.9 of 103, l.32-43).

Page 9. Line 20: Who was the third researcher, AFK? Please include initials.

Response 1.12: The third researcher was Ümran Sema Seven, mentioned in the acknowledgments. We included this clarification in the text: A guide containing 24 questions and five thematic blocks was then discussed with an independent third researcher (ÜSS, see acknowledgments) (p.11 of 103, l.28-33).

Also, please detail further the sources, methods, investigators (if others), and theories used to cross-validate the findings (triangulation). If this was not done, please add this as a limitation.

Response 1.13: Thank you for emphasizing aspects of triangulation. The application of investigator triangulation is now clearly stated in the manuscript: The further process of analysis was critically accompanied by discussions among the investigators, about how to group the findings, whereupon a final coding system was agreed on (i.e., investigator triangulation) (p.12-13 of 103, l.57-4). Lack of further cross-validation is now described in the limitations section: lack of cross-validation beyond investigator triangulation (e.g., data triangulation) (p.28 of 103, l.18-21).

Pages 9-10: It would have been arguably stronger evidence if the thematic blocks and coding utilized in the study involved external or independent coders for objectivity in the analysis. Please include this as a limitation.

Response 1.14: The lack of independent coders is now included in the strength and limitations bullet points: Due to the feasibility stage, the sample size was small, no participant validation was performed, and no independent coders were used (p.5 of 103, l.15-21). It is also mentioned in the limitations section: ...lack of cross-validation beyond investigator triangulation (e.g., data triangulation), and that coders were not independent (p.28 of 103, l.18-21).

However, we want to note that, as described above and in the text, thematic blocks and the interview guide were discussed with an otherwise uninvolved researcher (ÜSS).

Page 11. Lines 30-36: Please explain why those conducting the qualitative research did not return to participants with initial findings to verify if they accurately represented their experiences (member checking). (Supp 2 states that there were no repeat interviews). If this was not done, please add as a limitation.

Response 1.15: We added the lack of participant validation in the limitations section: The small sample size for this qualitative feasibility study, the lack of validating the results by the study participants,... (p.28 of 103, l.15-18).

Also, please elaborate on why participatory action research (actively involving the participant in the research process, including constructing the interview guide and analysis processes) was not done. This qualitative method would assist in objectives (i) and (ii).

Response 1.16: We completely agree, that this approach would have been valuable in the present case and added a sentence in the 'Patient and public involvement' section: This was due to limited staff and time resources (p.13 of 103, l.49-53).

Results:

Table 4: Why have 18 of the 22 completed the surveys? Please clarify in the abstract and results. What does the - sign mean? Not available, not applicable, or was not obtained? Perhaps some clarification on what this symbolizes would be helpful in interpreting the response to surveys.

Response 1.17: Thank you for your question. We added information on the participants not completing the survey in table 4: Of the remaining four participants two had dropped out, while another two did not complete the survey (p.23 of 103, l.41-47). We also changed the symbol to n.a. (not applicable), as several questions were only applicable to one of the groups.

Discussion:

In going back to the Abstract, what about implications for future RCTs? You do provide implications for future studies (page 24 of 41, Lines 10-11), but it would be beneficial to further elaborate on how these specific findings per group will directly inform the planning and execution of future RCTs.

Response 1.18: Thank you for highlighting implications for future RCTs. First, as described in Response 1.3, we further elaborated on how to tackle negative self-evaluation in a future study in the discussion: For a future study, an additional post-session self-rating will be helpful to match pwMS' individual performance levels and needs (e.g., "I felt not challenged enough/overburdened/just right"; in case of feeling overburdened, the therapist should seek further discussion with the participant and might adapt future sessions and re-evaluate the training fit) (p.25 of 103, l.10-21). Furthermore, we decided to summarize implications for future studies, mentioned throughout the discussion, in the conclusion section: Future group-based exercise studies should include participants with similar levels of motor performance and include additional self-ratings of exercise demands, post-session. The present results, supplemented by our quantitative results[22], further show that future studies conducted in an inpatient rehabilitation setting can involve fatigued pwMS in high-frequency exercise schedules if this includes adequate rest breaks. Furthermore, new forms of exercise interventions

should be well-described in the study information sheet to minimize participants favouring traditional exercise approaches, and it will be beneficial to include patient representatives in similar projects like this to construct the interview guide and aid in the analysis process. Lastly, qualitative methods should be used alongside quantitative measures in the study of fatigue in the future, with one field of application being the investigation of the important factors of the inpatient rehabilitation environment contributing to fatigue reduction in pwMS (p.29 of 103, l.32-60). We hope this gives the reader an adequate overview on implications for future studies.

Please discuss the psychological benefits; most mentioned were physical outcomes; however, the abstract included psychological benefits as well.

Response 1.19: Thank you for stressing the psychological benefits, which are now discussed in this paragraph in the manuscript: However, as part of the benefits category, some participants also described a sense of accomplishment after finishing a session, and effects related to self-efficacy to continue exercising at home. As we did not perform interviews pre-intervention, we do not have a detailed understanding on how participants' issues with self-esteem, or self-efficacy regarding exercise changed during the intervention. Russell et al.[31] reported improvements in these domains after a 10-week social cognitive behaviour change physical activity intervention, indicating that incorporating workshops on principles of social cognitive theory in the program might be beneficial. Nevertheless, the current psychological benefits already support that the exercise content felt relevant to pwMS, which will assist compliance in the future (p.25 of 103, l.20-43). Furthermore, psychological benefits are now also mentioned in the brief summary at the start of the discussion: A physical benefit highlighted by the participants was improved balance, while psychologically, experiencing a sense of accomplishment was emphasized (p.24 of 103, l.22-31).

The discussion compares the findings with other studies of SET nature; please add more comparisons regarding MAT and fatigue management in pwMS to strengthen the understanding of how this study fits within the field (or challenges the current exercise paradigm for pwMS).

Response 1.20: As mentioned in the discussion section, to our knowledge, this is the first qualitative study assessing barriers and facilitators regarding MAT in pwMS, and regarding the demands of high frequency exercise during inpatient rehabilitation. Therefore, adding comparisons regarding MAT is difficult. However, we expanded some more on the aspect of 'vacation from daily grind' with quantitative results from two RCTs conducted in similar inpatient settings: Two recent RCTs conducted in similar inpatient settings have not yet elaborated on this aspect, but it is noteworthy that in both studies, the control groups tended to also show reductions in fatigue experience, i.e., potentially displaying 'vacation from daily grind'[42, 43]. Moreover, it has to be noted that quantifying change in fatigue with established questionnaires, including the ones used in the two recent RCTs and the present feasibility study, proves to be difficult[22, 44, 45], which is why qualitative investigations can be seen as a valuable methodology in these instances (p.27-28 of 103, l.52-9). Regarding your last comment (see below), these are also relatively recent (2021 & 2023).

Page 23 of 41, lines 33 to 37. Please include a brief discussion of potential biases in participant responses since the interviews were conducted post-intervention.

Response 1.21: Thank you for mentioning biases in participant responses, which we briefly discussed in the following section: It is possible that participants' responses at the time of the interview were influenced by other processes of their stay in the NRC (e.g., overall satisfaction with their stay, perceived overall success, satisfaction with the recommendation regarding their future work situation), and by the fact that participants knew they were part of an intervention study (p.28 of 103, l.47-57).

The brief mention of seasonal factors (i.e., heat) is an interesting observation (page 24 of 41, lines 27-28); please elaborate on how they may affect exercise ability and adherence, and how this could be anticipated for future research.

Response 1.22: Thank you for highlighting this finding and presenting ideas to further elaborate on this aspect in the discussion section. We expanded on this and included an additional (and recent) reference in the manuscript: Heat might not only heighten fatigue levels, but also negatively impact balance control, as reported by pwMS[39]. Therefore, future studies should monitor in advance, whether the training locations are susceptible to heat, whether there are options for cooler environments, or time of training during the day (e.g., morning hours) and how exercise will be adapted in case of heightened fatigue and lowered balance control due to heat (p.26 of 103, l.40-52).

References: Of the 36 citations, only 5 are very recent (specifically 2022). There is no updated research in 2023. Please update references to include 2023 RCTs comparable with the objectives and multimodal training protocol (IF there are recent ones).

Response 1.23: See Response 1.20. The manuscript now includes five 2023 references.

Reviewer: 2

Prof. Laikang Yu, Beijing Sport University

Comments to the Author:

This manuscript aims to explore the experiences of fatigued persons with multiple sclerosis (pwMS) with a new multimodal agility-based exercise training (MAT) framework and to investigate the demands of the ReFEx (Rehabilitation, Fatigue, and Exercise) study protocol. The topic is very interesting, but from the process and results of the current study, it is not possible to fulfill the purpose.

1. The Introduction is disorganized. Too many descriptions of the study process that should have been in the Methods section.
2. There is a very small research group, and I am concerned about the quality of the results and the possibility of drawing conclusions. Specifically, for face-to-face interviews, only 6 participants were included in the MAT group (f:m = 5:1, 2 young and 4 middle-aged) and SET group (f:m = 4:2, 1 young, 2 middle-aged, and 3 elderly), respectively. The gender and age of participants affect how they feel about the training, which makes their answers to the same questions inconsistent. In order to validate the effectiveness of a new training protocol, a large number of homogeneous participants are needed.

Response 2.1: Thank you for reviewing our manuscript. We agree that the introduction includes several parts relating to the design of the study project. However, we think it is important to give the reader some background on where this interview study fits in the overall study project. We decided to include this information in the introduction and focused on the qualitative methods, which were relevant for this manuscript, in the methods section.

As described in the introduction, this was a feasibility study and not designed to test the effectiveness of the MAT framework. We are fully aware, that our sample was not sufficient to test for effectiveness, and the small sample size is also mentioned as a limitation in the manuscript (as it is mentioned in the quantitative results paper (Wolf et al., 2023, BMC Neurology)). In qualitative research in the context of feasibility studies it is usually favorable to achieve a diverse sample to acquire perspectives from participants with different demographical backgrounds, as, for example, the training program might fit for young but not for older people, which would be a desirable result in a qualitative feasibility study.

Reviewer: 3

Dr. Sally Thorne, University of British Columbia

Comments to the Author:

My focus in this review was on the qualitative methodology. Although I am very familiar with the experience of fatigue in multiple sclerosis, I do not have expertise in physical therapy interventions in this regard.

Response 3.1: Thank you for providing this detailed review of our qualitative methodology!

This manuscript describes a feasibility study using qualitative approaches within a larger clinical trial study. The rationale for conducting a qualitative investigation to complement that which can be assessed using more conventional measures is effectively articulated. In this context, the particular nature of the sample involved in this study is well described and justified. The recruitment strategy involved selection from the total population involved in the clinical trial, and allowed for selective invitation to reflect a balanced and appropriately diverse sample. None of those approached declined to be interviewed (which suggests enthusiasm for the study, and perhaps the authors' sense that these individuals would appreciate an opportunity to provide further input. Although such a strategy does raise the possibility that those invited were more likely to be more favorably inclined toward the intervention, the specific focus of the qualitative aspect means that should not constitute a design problem or compromise the usefulness of the findings generated from this body of interview material.

The description of the data collection and analysis process reveals that a number of design steps were reconsidered throughout the process, and as the analysis evolved. This would be considered problematic in some circles, as it would appear that the authors were continuing to refine their understanding of the kinds of qualitatively derived data that would be important to the larger intentions of the wider clinical trials project. What this meant in practice is that the guiding question was refined over time, which I see as consistent with what is common in 'applied qualitative' studies. While the authors did not explicitly cite such a methodological approach as their intended direction, and it might be appropriate to consider that for another round should they take this feasibility study into a full program, the detailed description of what they did and decided along the way, and why each decision was taken, should be sufficient to allow readers to judge for themselves whether the findings are appropriately or inappropriately shaped by these particular analytic decisions. In my view, they do no harm to the integrity of the study (although readers with a strong grounding in formal (social science) qualitative approaches might disagree).

Response 3.2: Thank you for pointing out the 'applied' nature of our qualitative approach, which will be considered for potential future work.

Ultimately, the interviews were quite short (mean of 15 minutes) and seemed guided by a number of predetermined questions. For this reason, the emphasis within the description of analysis (p.11) seems somewhat oddly focused on the number of categories and subcategories that were discerned. What we see when we get to the findings, is that the researchers may well have been simply trying to ascertain what existed within the data set that might be most helpful in guiding future inquiry. When we look to the ultimate decision as to how the findings are displayed and reported, the structure is quite simple, and may well reflect the best possible structure for the intended purpose. Because of this, the extraneous detail about how many groupings were included at each stage along the way may add more confusion than necessary. And again, if the authors were to draw on an explicitly "applied" methodology, then the guidance would make that extra information unnecessary. Rather, focusing on how their thinking evolved through this process to the point that they came to an understanding of what not to report (because it was already well known within the field based on prior literature) and what new insights this study could add to the field. And that is what constitutes meaningful findings.

Response 3.3: Thank you for supporting our category structure for our intended purpose, which you described in your comment. We agree that the analysis section included some details that might have added unnecessary confusion. Therefore, we tried to reduce the listing of numbers of categories in this section (p.12 of 103, l.40-43).

Ultimately the findings presented seem entirely appropriate to the approach used to develop them, and are well aligned with the overall stated purpose within this feasibility phase. The authors are familiar with both the strengths and limitations of using this kind of small qualitative study within a larger clinical trials program. I do note, however, that they imply (p.27) in their discussion of limitations that they might have achieved saturation with a larger study. Since there is some possibility that they may extend this into their larger program of research over time, I would caution against such an implication. Saturation implies that they will have covered all relevant variations, which I suggest is an inappropriate assumption within a study of a complex human phenomenon such as MS fatigue. Rather, their claim would better resonate with clinicians/readers if they acknowledge that no qualitative investigation in such a context can reasonably claim “saturation,” and that the infinite range of possible variation in something like fatigue is what makes it so interesting to work with. There is considerable literature to draw from to support not taking up that social science theorizing concept in the world of clinical research where it is generally inaccurate and misleading.

Response 3.4: We took up your advice and revised this part of the limitations section, as follows: Overall, we only acquired a small glance at the diversity of fatigue experiences, while it is unclear whether saturation can be achieved at all for this complex phenomenon (p.28 of 103, l.30-36).

In conclusion, I believe that this feasibility study is worth reporting, and that the findings within it will be of interest, not only to the direction of the future research of this team of investigators, but also to others who are considering including a qualitative component to their clinical trials of evolving interventions for complex human experiences such as fatigue in chronic illness.

Response 3.5: We thank you for your support and thoughtful review. We are similarly convinced that this form of mixed-methods research is very valuable in the context of fatigue and exercise.

VERSION 2 – REVIEW

REVIEWER	Flores, Victoria A. University of Illinois Chicago, Kinesiology and Nutrition
REVIEW RETURNED	17-Jan-2024

GENERAL COMMENTS	All concerns have been addressed. The paper is much stronger and clearer. I have no additional comments and look forward to the next step of this line of research.
-------------------------	---

REVIEWER	Thorne, Sally University of British Columbia, Nursing
REVIEW RETURNED	27-Dec-2023

GENERAL COMMENTS	Thank you for your careful and thoughtful attention to all of the original reviewer comments (mine included). It is clear that you have understood the rationale for the original suggestions and have made realistic and appropriate responses to them. I look forward to seeing this manuscript in published form and feel confident it will find an enthusiastic reading audience.
-------------------------	---