PEER REVIEW HISTORY

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

(This paper received three reviews from its previous journal but only two reviewers agreed to published their review.)

ARTICLE DETAILS

TITLE (PROVISIONAL)	Effects of resistance training and/or beta-hydroxy-beta-methylbutyrate supplementation on muscle mass, muscle strength, and physical performance in older women with reduced muscle mass: protocol for a randomized, double-blind, placebo-controlled trial
AUTHORS	Osuka, Yosuke; Kojima, Narumi; Wakaba, Kyohsuke; Miyauchi, Daiji; Tanaka, Kiyoji; Kim, Hunkyung

VERSION 1 – REVIEW

REVIEWER	Tsan-Hon Liou Department of Physical Medicine and Rehabilitation, School of Medicine, College of Medicine, Taipei Medical University, Taiwan
REVIEW RETURNED	04-Sep-2018

GENERAL COMMENTS	Thank you for giving me this opportunity to review this manuscript. This is a study protocol of a randomized, double-blind, placebo-controlled trial which aims to examine the acute and residual effects of RT and/or HMB supplementation on muscle mass, muscle strength, and physical performance in older women with reduced muscle mass. This study was well designed with IRB approved and clinical trial registered. SPIRIT 2013 Checklist was also completely provided. I have only some concerns about this manuscript as following: Page 5, line 53 "A sample of 328 women fitting the eligibility criteria received invitation letters regarding the study intervention, of whom 156 participated in the baseline assessment." Could the authors explain how they choose these 156 participants from the pool of 328 women? The primary outcome is the longitudinal change in muscle mass. Secondary outcomes include the longitudinal changes in muscle strength, physical performance, muscle thickness, muscle quality, blood counts, blood biochemistry, calf circumference, skin viscoelasticity, habitual dietary intake, habitual physical activity levels, functional capacity, and health-related quality of life. There are so many outcomes measured in this study. Page 11, line 32 for sample size prediction, α-error to 0.01 (Bonferroni correction for multiple comparisons as post-hoc tests) was used. But the P-values is set < 0.05 will be considered to indicate significance.

	Why did the authors think Bonferronni correction for sample size prediction but not for the result analysis? Page 7, line 12 Placebo products do not include calcium or HMB, and the missing amounts are provided as carbohydrates. I am just curious about how much the added calcium and carbohydrate will impact on the primary and secondary outcomes? Page 7, line 22 “Wu et al. suggested a daily dose of 3,000 mg for older adults. However, as body weight is lower among Sarcopenic Japanese older women than among western older women, we set the active dose at 1,200 mg HMB daily.” Is there any reference that proved this amount adequate for building muscle with only 40% of recommendation level? Page 7, line 52 “Body composition measurements are conducted using the InBody720 device (Biospace Co., Ltd, Seoul, Korea), which has a validity for estimating appendicular lean mass in community-dwelling older populations in comparison of dual-energy X-ray absorptiometry systems.” I believe that BIA is quite improved in validity and also much more suitable for used in a community setting. But in this study, I wonder BIA could be sensitive enough to detect the muscle change in 3 months. Page 9, line 49 “Skin viscoelasticity is evaluated using the Cutometer® dual MPA 580 (Courage + Khazaka electronic GmbH, Cologne, Germany), which measures the elasticity of the upper layer of the skin using negative pressure to induce mechanical deformation” I am curious why the authors measure this outcome which is not mentioned in the introduction.

REVIEWER	Stuart Phillips McMaster University, Canada
REVIEW RETURNED	27-Sep-2018

GENERAL COMMENTS	The authors have outlined a protocol that they are running and that is registered: UMIN000028560 The trial protocol is well described and the protocol is outlined in sufficient detail. The authors are not, however, the first to test women as they note by citing Vukovich, so their bullet point on page 3 at top is not entirely correct Clarification is needed on the following items: The authors state on page 4 that, “Wu et al. reviewed the results of seven RCTs on the effects of HMB supplementation on muscle outcomes and concluded that HMB may contribute to preservation of muscle mass in older adults.” This is a very low quality meta-analysis that included trials from far too many studies of a very heterogeneous nature. The following commentary on the limitations of such meta-analyses bears important consideration https://www.ncbi.nlm.nih.gov/pubmed/28975260. Other authors have, appropriately, concluded that there are no grounds for a meta-analysis of HMB in aging persons https://www.ncbi.nlm.nih.gov/pubmed/24057808. The authors themselves admit that there are few studies on this by citing Vukovich and Stout, so how then is Wu’s meta-analysis valid? What form of HMB is being used, Ca or free acid?

	Does KYOWA Co have an affiliation or is supplying the HMB under license from Metabolic Technologies Incorporated (MTI): http://www.mettechinc.com/ If so, this needs to be disclosed. The description of the sample size is incomplete. The authors state that "…required to detect clinically important differences…" What is the size of the minimally important clinical difference? Please provide a reference for this.

REVIEWER	Philip Chilibeck University of Saskatchewan, Canada
REVIEW RETURNED	05-Oct-2018

GENERAL COMMENTS	The manuscript describes a protocol for a randomized controlled trial of HMB supplementation and resistance training in older women with low muscle mass. My main comments pertain to the statistical analysis. I think you should be using a factorial analysis to better match the design of your study. Please see specific comments below. Abstract: A registration number is given; however, it is not clear where the study is registered. Please provide the name and/or website at the end of the abstract. Page 3, strengths and limitations section, fourth bullet point: It is stated here that resistance training is by using body weight, ankle weights, and elastic bands; however, in the abstract it is stated that machine-based exercises were also done. Page 5, line 46: please provide a reference for the sarcopenia diagnosis consensus issued by the Asian Working Group for Sarcopenia. On page 6, please provide details on the sets and repetitions used for the non-weight machine exercises. For the machine-based exercises, it is stated that 8–10 repetitions were performed, with the first four counts being performed with increasing strength and the last four with decreasing strength. Please clarify what is meant by "the first four counts being performed with increasing strength and the last four with decreasing strength". Outline the reproducibility (i.e. as a % coefficient of variation) for your main outcomes. Page 9, line 14: "Quadriceps femoris thickness is defined as the sum between the thickness of the rectus femoris and that of the vastus intermedius." Replace "between" with "of" Please provide a reference for the validity of the muscle ultrasound measurement of muscle quality. Page 11, adverse events: "If any adverse events associated with RT or supplements occur during the intervention period, the information is recorded and the program discontinued immediately." I would suggest that only for serious adverse events a participant may be withdrawn from the program, or the participant should have a choice as to whether or not they want to discontinue because of an adverse event. Also, if someone has an

	adverse event related to the supplement or placebo, can they still continue resistance training? Also, if someone has an injury because of the resistance training, can they still continue taking the supplement? I think they should still be allowed to in both cases, and they should be assessed at the conclusion of the trial and included in the group they were randomized to (i.e. as part of the "intent to treat" principle). Statistics section: You have used a factorial design; therefore, you should ensure you use a factorial statistical analysis. I don't think you have done this according to your description. The factorial ANOVA would include two group factors (i.e. HMB vs. placebo AND exercise vs. non-exercise) and a time factor (i.e. within-subjects or repeated measures factor). Using a factorial design will increase your statistical power (because two groups taking HMB are compared to the two groups taking placebo AND the two groups who are exercising are compared to the two groups who are not exercising). From this type of analysis, you will get supplement x time and exercise x time interactions from the ANOVA table that tell you whether the HMB or exercise were effective. You will also get an exercise x supplement x time interaction. If the supplement x time and exercise x time interactions are statistically significant, but the exercise x supplement x time interaction is not statistically significant, then you can say the two interventions were additive. If the exercise x supplement x time interaction is significant, then you can say the two interventions were synergistic (i.e. their combined effects were greater than the addition of the two effects). You state results will be presented as means and standard errors. You should be using standard deviations for your data. You are adjusting for baseline measures in your analysis. I don't think there is justification for doing this. It is stated you are doing a sub-group analysis with participants stratified by sarcopenia status. More detail needs to be provided here (i.e. how many groups are women being placed into for their sarcopenia status)? It is stated "Missing data will be applied multiple imputation." Please provide more details on how you are doing this. At the end of the manuscript, it is stated "One limitation is that this study will involve RT using body weight, elastic bands, and ankle weights, which does not provide objective information regarding muscle loading" As stated earlier, the participants also did machine-based exercises training; therefore, I don't think this statement is valid.

VERSION 1 – AUTHOR RESPONSE

Responses to comments made by Reviewer #1

Major comments

Thank you for giving me this opportunity to review this manuscript. This is a study protocol of a randomized, double-blind, placebo-controlled trial which aims to examine the acute and residual effects of RT and/or HMB supplementation on muscle mass, muscle strength, and physical performance in older women with reduced muscle mass. This study was well designed with IRB

approved and clinical trial registered. SPIRIT 2013 Checklist was also completely provided. I have only some concerns about this manuscript as following:

1. Page 5, line 53 “A sample of 328 women fitting the eligibility criteria received invitation letters regarding the study intervention, of whom 156 participated in the baseline assessment.” Could the authors explain how they choose these 156 participants from the pool of 328 women?
Response: We realize that it may seem we applied some selection criteria at that time, but this was not the case. Of the 328 eligible women invited to participate, only 156 agreed to undergo baseline assessment, whereas the remaining 172 declined to participate due to personal reasons (e.g., inconvenient schedule). We have clarified this point in the revised manuscript (page 5, lines 33; Figure 1).

2. The primary outcome is the longitudinal change in muscle mass. Secondary outcomes include the longitudinal changes in muscle strength, physical performance, muscle thickness, muscle quality, blood counts, blood biochemistry, calf circumference, skin viscoelasticity, habitual dietary intake, habitual physical activity levels, functional capacity, and health-related quality of life. There are so many outcomes measured in this study. Page 11, line 32 for sample size prediction, α-error to 0.01 (Bonferroni correction for multiple comparisons as post-hoc tests) was used. But the P-values is set < 0.05 will be considered to indicate significance. Why did the authors think Bonferronni correction for sample size prediction but not for the result analysis?
Response: At the suggestion of Reviewer #3, we have changed the main statistical analysis from mixed-model repeated-measures analysis of variance to two-way ANOVA. We have updated all descriptions accordingly (page 11, lines 29 to page 12, line 1).

3. Page 7, line 12 Placebo products do not include calcium or HMB, and the missing amounts are provided as carbohydrates. I am just curious about how much the added calcium and carbohydrate will impact on the primary and secondary outcomes?
Response: It is very difficult to discuss the impact of the added calcium and carbohydrate on the study outcomes, mainly because the replaced amounts were minimal, totaling approximately 120 g over the course of 12 weeks. Thus, we expect the effect of these differences to be limited.

4. Page 7, line 22 “Wu et al. suggested a daily dose of 3,000 mg for older adults. However, as body weight is lower among Sarcopenic Japanese older women than among western older women, we set the active dose at 1,200 mg HMB daily.” Is there any reference that proved this amount adequate for building muscle with only 40% of recommendation level?
Response: Thank you for bringing our attention to this important aspect. We could not find any explicit evidence that HMB at 1.2 g/day would represent the optimum dosage in older women with reduced muscle mass, mainly because very few studies have examined the effects of HMB in this population. Previous trials conducted in the US, which enrolled healthy older adults, used HMB doses of 0.03–0.04 g per kg of body weight (Table 1, please check the uploaded file). In our study, the participants’ weight was estimated at approximately 40 kg, based on body weight information from our previous study enrolling older women with reduced muscle mass (Kim et al., 2012, ref. #17 in the revised manuscript). In this case, we have confirmed that the daily HMB dosage used in our study (0.03 g/kg of body weight) is comparable to that used in previous studies (0.03–0.04 g/kg of body weight). We have added this information to the revised manuscript (page 7, lines 13-14). Additionally, 1.2 g of HMB would correspond to 12–24 g of leucine (Van Koevering et al., 1992, ref. #11). In a previous study enrolling sarcopenic older women, we found that daily supplementation with 6 g of leucine was associated with significantly increased muscle mass following exercise (Kim et al., 2012, ref. #17). Therefore, in the present study, we expect to observe significant change in muscle mass even for an HMB dosage of 1.2 g/day.

5. Page 7, line 52 “Body composition measurements are conducted using the InBody720 device (Biospace Co., Ltd, Seoul, Korea), which has a validity for estimating appendicular lean mass in community-dwelling older populations in comparison of dual-energy X-ray absorptiometry systems.” I believe that BIA is quite improved in validity and also much more suitable for used in a community setting. But in this study, I wonder BIA could be sensitive enough to detect the muscle change in 3 months.
Response: We agree that DEXA may be better than BIA for detecting the change in muscle mass over this short period. However, in a previous study enrolling sarcopenic older women and employing BIA measurements, we noted a significant increase in muscle mass (3.1%) following 3 months of exercise + amino acid supplementation (Kim et al., 2012, ref. #17 in the revised manuscript). Therefore, we expect that BIA measurements will be useful in the present study.

6. Page 9, line 49 “Skin viscoelasticity is evaluated using the Cutometer® dual MPA 580 (Courage + Khazaka electronic GmbH, Cologne, Germany), which measures the elasticity of the upper layer of the skin using negative pressure to induce mechanical deformation” I am curious why the authors measure this outcome which is not mentioned in the introduction.
Response: We understand your concern. HMB increases protein synthesis through activation of the mammalian target of rapamycin and decreases protein catabolism through down-regulation of the ubiquitin proteasome pathway. Thus, HMB is expected to have an important role in promoting skin regeneration in older adults. In fact, HMB provided as a nutritional supplement was shown to be associated with enhanced wound repair after surgery (Williams et al., Ann Surg, 2002). However, since our primary focus is muscle mass, we omitted from the Introduction any mention of the effects of HMB on skin quality and are collecting such data as secondary or exploratory outcomes.

Responses to comments made by Reviewer #2

Major comments
The authors have outlined a protocol that they are running and that is registered: UMIN000028560. The trial protocol is well described and the protocol is outlined in sufficient detail.

7. The authors are not, however, the first to test women as they note by citing Vukovich, so their bullet point on page 3 at top is not entirely correct.
Response: Thank you for your bringing up this important point. Indeed, this study is not the first trial to examine the effects of HMB and exercise in older women. However, all previous studies focused on healthy older women and were not designed as four-arm RCTs to examine the effects of RT and/or HMB supplementation. Therefore, we would like to emphasize that this study is the first four-arm RCT to examine the combined effect of RT and HMB supplementation in older people with reduced muscle mass. We have revised the bullet point accordingly (page 3, lines 2-3).

Clarification is needed on the following items:
8. The authors state on page 4 that, “Wu et al. reviewed the results of seven RCTs on the effects of HMB supplementation on muscle outcomes and concluded that HMB may contribute to preservation of muscle mass in older adults.” This is a very low quality meta-analysis that included trials from far too many studies of a very heterogeneous nature. The following commentary on the limitations of such meta-analyses bears important consideration https://www.ncbi.nlm.nih.gov/pubmed/28975260. Other authors have, appropriately, concluded that there are no grounds for a meta-analysis of HMB in aging persons https://www.ncbi.nlm.nih.gov/pubmed/24057808. The authors themselves admit that there are few studies on this by citing Vukovich and Stout, so how then is Wu’s meta-analysis valid?
Response: We agree that the findings of the meta-analysis conducted by Wu et al. may be equivocal and thus have removed the discussion of such findings from this context. Instead, we cited the review

article you indicated, and which we found very useful. We have also revised that part of the Introduction to clarify the context of our research (page 4, lines 21-27).

9. What form of HMB is being used, Ca or free acid?
Response: We use the Ca form of HMB. We have now explicitly stated this fact in the Intervention section (page 7, line 4).

10. Does KYOWA Co have an affiliation or is supplying the HMB under license from Metabolic Technologies Incorporated (MTI): http://www.mettechinc.com/ If so, this needs to be disclosed.
Response: We use a domestic HMB supplement (i.e., produced in Japan) and are unaware of any affiliation Kyowa Co might have with MTI.

11. The description of the sample size is incomplete. The authors state that "…required to detect clinically important differences…" What is the size of the minimally important clinical difference? Please provide a reference for this.
Response: We have changed the main statistical analysis from mixed-model repeated-measures analysis of variance to two-way ANOVA, as suggested by Reviewer #3, and have updated the description for sample size accordingly (page 11, lines 29 to page 12, lines 1).

Responses to comments made by Reviewer #3

Major comments
The manuscript describes a protocol for a randomized controlled trial of HMB supplementation and resistance training in older women with low muscle mass. My main comments pertain to the statistical analysis. I think you should be using a factorial analysis to better match the design of your study. Please see specific comments below.

12. Abstract: A registration number is given; however, it is not clear where the study is registered. Please provide the name and/or website at the end of the abstract.
Response: The trial was registered with the University Hospital Medical Information Network (UMIN) Clinical Trials Registry, as mentioned in the Methods (page 5, lines 5-7). Due to restrictions regarding the word count, we could not include the full name of the registry, the trial registration number, and the full URL information within the Abstract. Per your instructions, we have added the URL at the end of the Abstract (page 2, line 29).

13. Page 3, strengths and limitations section, fourth bullet point: It is stated here that resistance training is by using body weight, ankle weights, and elastic bands; however, in the abstract it is stated that machine-based exercises were also done.
Response: Thank you for bringing our attention to this oversight. Indeed, RT was conducted also in the form of machine-based exercises. We have clarified the statement in the Summary (page 3, lines 9-12).

14. At the end of the manuscript, it is stated "One limitation is that this study will involve RT using body weight, elastic bands, and ankle weights, which does not provide objective information regarding muscle loading" As stated earlier, the participants also did machine-based exercises training; therefore, I don't think this statement is valid.
Response: We have revised this part (page 13, lines 10-12).

15. Page 5, line 46: please provide a reference for the sarcopenia diagnosis consensus issued by the Asian Working Group for Sarcopenia.
Response: We have provided the reference in question (page 5, line 28; ref. #16).

16. On page 6, please provide details on the sets and repetitions used for the non-weight machine exercises.
Response: We have included a detailed description of the prescription for machine and non-machine-based exercises (page 6, lines 28-29).

17. For the machine-based exercises, it is stated that 8–10 repetitions were performed, with the first four counts being performed with increasing strength and the last four with decreasing strength. Please clarify what is meant by “the first four counts being performed with increasing strength and the last four with decreasing strength”.
Response: We have clarified the description in question (page 6, lines 29-31).

18. Outline the reproducibility (i.e. as a % coefficient of variation) for your main outcomes.
Response: We have added information about the reproducibility of measurements conducted using the InBody720 device (page 7, lines 35-36).

19. Page 9, line 14: “Quadriceps femoris thickness is defined as the sum between the thickness of the rectus femoris and that of the vastus intermedius.” Replace “between” with “of”
Response: We have made the suggested revision (page 9, line 7).

20. Please provide a reference for the validity of the muscle ultrasound measurement of muscle quality.
Response: We have cited a study supporting the validity of muscle ultrasound measurements to assess muscle quality (page 9, lines 10-12; ref. #32).

21. Page 11, adverse events: “If any adverse events associated with RT or supplements occur during the intervention period, the information is recorded and the program discontinued immediately.” I would suggest that only for serious adverse events a participant may be withdrawn from the program, or the participant should have a choice as to whether or not they want to discontinue because of an adverse event. Also, if someone has an adverse event related to the supplement or placebo, can they still continue resistance training? Also, if someone has an injury because of the resistance training, can they still continue taking the supplement? I think they should still be allowed to in both cases, and they should be assessed at the conclusion of the trial and included in the group they were randomized to (i.e. as part of the “intent to treat” principle).
Response: Thank you for these excellent suggestions, which we plan to apply should adverse events occur. We have revised the description of such procedures accordingly (page 10, lines 35 to page 11, lines 2).

22. Statistics section: You have used a factorial design; therefore, you should ensure you use a factorial statistical analysis. I don’t think you have done this according to your description. The factorial ANOVA would include two group factors (i.e. HMB vs. placebo AND exercise vs. non-exercise) and a time factor (i.e. within-subjects or repeated measures factor). Using a factorial design will increase your statistical power (because two groups taking HMB are compared to the two groups taking placebo AND the two groups who are exercising are compared to the two groups who are not exercising). From this type of analysis, you will get supplement x time and exercise x time interactions from the ANOVA table that tell you whether the HMB or exercise were effective. You will also get an exercise x supplement x time interaction. If the supplement x time and exercise x time interactions are statistically significant, but the exercise x supplement x time interaction is not statistically significant, then you can say the two interventions were additive. If the exercise x supplement x time interaction is significant, then you can say the two interventions were synergistic (i.e. their combined effects were greater than the addition of the two effects).

Response: Once again, thank you for this very important observation and excellent suggestion, according to which we have changed the main statistical analysis from mixed-model repeated-measures analysis of variance to two-way ANOVA in order to increase statistical power. Our hypothesis is that, compared to RT alone, HMB supplementation alone, and placebo, combined RT and HMB supplementation would provide higher benefit in muscle mass, muscle strength, and physical performance. To test this hypothesis, we focus on the interaction between two factors (RT vs education and HMB vs placebo) to evaluate the longitudinal changes in outcomes (page 11, lines 29 to page 12, lines 1). We have revised the description of sample size accordingly (page 11, lines 14-22).

23. You state results will be presented as means and standard errors. You should be using standard deviations for your data.
Response: We have revised this point per your recommendation (page 11, lines 29).

24. You are adjusting for baseline measures in your analysis. I don't think there is justification for doing this.
Response: We agree and have deleted this par.

25. It is stated you are doing a sub-group analysis with participants stratified by sarcopenia status. More detail needs to be provided here (i.e. how many groups are women being placed into for their sarcopenia status)?
Response: This is an important point. Unfortunately, we cannot provide information about the expected size of the sub-groups defined according to sarcopenia status because baseline data (including sarcopenia status) will only become unblinded after the end of the second intervention.

26. It is stated "Missing data will be applied multiple imputation." Please provide more details on how you are doing this.
Response: We have added details of the multiple imputation procedure (page 12, lines 4-6).

VERSION 2 – REVIEW

REVIEWER	Tsan-Hon Liou Department of Physical Medicine and Rehabilitation, School of Medicine, College of Medicine, Taipei Medical University, Taiwan
REVIEW RETURNED	30-Dec-2018

GENERAL COMMENTS	The authors have already addressed all the questions I raised. I think this manuscript is ready to be accepted.

REVIEWER	Stuart Phillips McMaster University
REVIEW RETURNED	20-Dec-2018

GENERAL COMMENTS	Nice job on revision.

REVIEWER	Philip Chilibeck University of Saskatchewan, Canada
REVIEW RETURNED	31-Dec-2018

GENERAL COMMENTS	Page 6, line 49: "Each movement was performed slowly, in 8 counts (increasing strength in the first 4 counts and decreasing strength in the last 4 counts)." I think the authors might have meant "lifting for first 4 counts and lowering for the last 4 counts"

	or “concentric for the first 4 counts and eccentric for the last 4 counts”. Please clarify. Page 12, line 7: add “standard” before “deviations” Page 12, line 21: Change “will be not invited” to “will not be invited”

VERSION 2 – AUTHOR RESPONSE

Responses to comments made by Reviewer #3
Minor comments
1. Page 6, line 49: “Each movement was performed slowly, in 8 counts (increasing strength in the first 4 counts and decreasing strength in the last 4 counts).” I think the authors might have meant “lifting for first 4 counts and lowering for the last 4 counts” or “concentric for the first 4 counts and eccentric for the last 4 counts”. Please clarify.
Response: Thank you for the kind suggestions. We have clarified the description (page 6, lines 29 to 31).

2. Page 12, line 7: add “standard” before “deviations”.
Response: We have deleted this sentence since such sentence is already included at the beginning of the paragraph.

3. Page 12, line 21: Change “will be not invited” to “will not be invited”.
Response: We have revised this part (page 12, lines 11 to 12).