PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Ginkgo Biloba Extract for Prevention of Acute Mountain Sickness: A Systematic Review and Meta-analysis of Randomized Controlled Trials
AUTHORS	Tsai, Tou-Yuan; WANG, SHIH-HAO; Lee, Yi-Kung; Su, Yung-Cheng

VERSION 1 – REVIEW

REVIEWER	Ingrid Arevalo-rodriguez Clinical Biostatistics Unit, Ramón y Cajal Hospital (IRYCIS), Madrid, Spain
REVIEW RETURNED	12-Feb-2018

GENERAL COMMENTS	Thanks for the opportunity to provide comments about this important manuscript. I want to congratulate the authors for this initiative to assess a non-pharmacological intervention to prevent high altitude illness. I have some suggestions and one concern related to this research. Hope that my comments would be useful. Major concern: I have doubts about the inclusion of Gertsch 2002, due to this study does not seems like a randomized controlled trial at the end. For example, authors said: “Study administrators individually assigned the participants an experimental treatment on a random, first-come first-served basis after receiving signed consent.” Is this really a randomized method? Please consider whether the assessment of the risk of bias should be affected in this trial (authors have considered this study as at low RoB) or this study needs to be excluded (if there are no a truly randomized method, the study is not a truly randomized trial) In addition, authors of this trial said: “As previously noted, only 26 participants, the first of four planned groups, were enrolled before the trial was closed because of an unexpectedly high incidence of subjects meeting the predetermined safety criteria for severe AMS (11 subjects or 42%). There was additional impetus to close the study early because one of the research staff also suffered from severe AMS during the study. Thus, this study presents data on only 26 subjects when the intention was to enroll 100 subjects.” In this case, this study have an important limitation that need to be reflected in its quality assessment. At present, this issue is not considered by the authors of the review. Minor concerns: • Please define in methods how the included trials were considered as at “high Rob”. • Please consider if a funnel plot is useful when there are less than 10 studies to analyze (revise the Cochrane Handbook for more information about assessment of publication bias).

	• Authors said: “The results of several pre-planned subgroup analyses with all 7 datasets were similar”. however, these analyses were not detailed in methods. Please include this important information in the corresponding section.

REVIEWER	Kannan Arabian Gulf University, Bahrain
REVIEW RETURNED	21-Feb-2018

GENERAL COMMENTS	1. Was the protocol registered? The name of the registry and the concerned ID should be mentioned in the methodology. 2. What was the definition of 'population' that was included in the meta-analysis. Considering the variability that exists, this is important. 3. Assesment of publication bias by Funnel plot requires at least 10 studies whilst only 6 were included in the present review. 4. There are no details on the outcome measures that were considered in the methods section. 5. It looks like only one outcome measure was assessed - prevention of AMS; There are several clinically important outcomes that are missed by the investigators: incidence of severe AMS, headache and severe headache, oxygen saturation and adverse events. 6. Considering the multiplicity of statistical analyses and small number of events/patients, the authors should consider correcting the type 1 error for the number of analyses. 7. Compliance to PRISMA has not been followed and reported.

REVIEWER	Irene SL Zeng University of Auckland
REVIEW RETURNED	21-Mar-2018

GENERAL COMMENTS	The systematic review and meta-analysis study includes six eligible studies for investigating the pharmaceutical effectiveness of using GBE for prophylaxis of Mountain sickness. It was carefully written and included a large number of studies in the screening phase. I summarize its strength and potential improvements as follows: Strength: 1. The study was carefully analyzed and includes heterogeneity analysis with and without the studies of high risk of bias. 2. The study has thorough quality assessment of each study. 3. It clearly identified the potential confounding factors of efficacy. Potential improvements: 1. Primary endpoint: Please provide information about how the primary endpoint - acute mountain sickness (AMS) is defined in each study. 2. Meta-analysis: there are significant heterogeneities among the six studies. Since the number of studies is small, it is not reliable to analyze the subgroup based on the identified confounders. Authors can consider testing if there are significant associations between relative risk and dosage, days of treatment before ascent, gender and altitude using a method such as a simple meta-regression (only including one explanatory variable in the model). 3. In the result session, there is no interpretation of the relative risk. It will be informative to explain the relative risk. 4. Please consider adding PRISMA as an guideline. Other suggestions: 4. Why does the funnel plot only has 5 studies?

	5. It will be clearer to add incidence of disease (%) in Table 3; example is table 2 from Lionel Dumont L et al (2000, BMJ), Efficacy and harm of pharmacological prevention of acute mountain sickness: Quantitative systematic review paper. 6. Please consider removing the subgroup analysis forest plots; only keep the summarized RR in result. 7. It is not clear in the forest plot which direction is the placebo and which is the GBE. 8. Please consider adding Dumont L et al (2000, BMJ) paper as a reference, it has identified ascent rate as a significant factor for drug responsiveness and efficacy. Other minor suggestions: Line40 “Contributorship statement: can change to “author contributions”. Reference : Dumont, L., Mardirosoff, C., Tramèr, M.R. Efficacy and harm of pharmacological prevention of acute mountain sickness: Quantitative systematic review (2000) British Medical Journal

REVIEWER	Igor Locatelli University of Ljubljana, Faculty of Pharmacy, Ljubljana, Slovenia, EU
REVIEW RETURNED	26-Mar-2018

GENERAL COMMENTS	I have read the article about meta-analysis of ginkgo biloba extracts (GBE) for prevention of acute mountain sickness (AMS) and reviewed it with an emphasis on the statistical methods and analyses used. The article is well written, however I would recommend making tables and figures more concise and reduce the number of them. The major issue of this manuscript s the misinterpretation of the statistical analysis results when p values are above 5%. Is such cases the authors concluded that the null hypothesis was true, which is wrong thing to do. The effect sizes of GBE (in terms of relative risks) in some studies and the pooled relative risks are quite below 1, meaning that the effect might be clinically significant if higher number of subjects were included in the studies or more studies were conducted. In some studies none of the participants in GBE arm developed AMS while a half of the participants in control arm did develop AMS. Therefore, I cannot agree with your conclusions that currently available data suggest the GBE does not prevent AMS. This sentence should be at least written in a way Cochrane metaanalyses reports such conclusions, e.g. *There are not enough data to show the statistically significant effect of GBE for preventing AMA; further studies are warranted….* . Furthermore, as further studies are warranted, I would suggest including/discuss the results of nonrandomized controlled studies (if they were performed) at least in the discussion. Specific comments: 1. In the introduction I lack a paragraph about the information how AMS can be assessed/measured. There are some questionnaires mentioned in the table1 but a more detailed description of possible evaluation of AMS presence is needed in the introduction. Selection about AMS evaluation method should be included as inclusion/exclusion criteria, and in description of the outcome assessment. 2. Page 7, line 10. In what bases the final decisions about overall

	risks of bias were made? This should be explained here. 3. Page 7, line 40, the main outcome measure is the occurrence of AMS (defined as – see point 1). The RR is one of the options how to calculate the effect size. And you should state here base on what method RR, weights and pooled (overall) RR were calculated - was the Mantel-Haenszel method or the inverse variance method applied? Mantel-Haenszel method is generally preferable when there are few events (see: Cochrane Handbook for Systematic Reviews of Interventions, available online at http://handbook-5-1.cochrane.org/). What method was used to calculate between-trial variability in random effect model? Was the DerSimonian and Laird method? This is major. 4. In two studies none of the participants in GBE arm developed AMS; therefore, it is impossible to calculate RR unless you impute a small number (e.g. 0.5 or 0.3) instead of 0. This is important as the RR estimates can vary significantly. How was this treated in your analysis? 5. Page 7, line 46. Why did not you perform a subgroup analysis based on GBE dosing regimen (there is 2-fold difference in daily dose of GBE used between the included studies)? 6. Tables: Again as explained above you cannot conclude that some studies proved GBE does not prevent AMS – maybe they were underpowered to achieve statistical significance. So I strongly suggest not dividing the study to one that proved GBE for preventing AMS and other that did not. It is very unusual and statistically incorrect to do so. 7. Tables 1 and 2 can be put together. Results in the Table 3 can be a part of a forest plot. See some forest plot from Review Manager. 8. Figure 3 can be omitted and the polled (overall) RR (with number of studies and number of participants) can be presented in the text. You have done this, so I would just add the number of studies and number of participants and omit the Figure 3. 9. Figure 4 is irrelevant, funnel plot is meaningful in whole set of data. 10. Since you did not show and difference from subgroup analyses I would suggest to omit the figures 6, 7, and 8 and present the results in the text or create a separate table with results (pooled RR; number of included studies, number of patients in each arm) from sensitivity and subgroup analysis. 11. Figure 5. I would perform such analysis on the whole set of studies, the figure should be marked with line of equality (x axis = 1) not by 1.08. I do not see any benefits of such figure, though. Could be omitted.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer1:

1. Major concern: I have doubts about the inclusion of Gertsch 2002, due to this study does not seems like a randomized controlled trial at the end. For example, authors said:

“Study administrators individually assigned the participants an experimental treatment on a random, first-come first-served basis after receiving signed consent.” Is this really a randomized method?

Please consider whether the assessment of the risk of bias should be affected in this trial (authors have considered this study as at low RoB) or this study needs to be excluded (if there are no a truly randomized method, the study is not a truly randomized trial)

In addition, authors of this trial said:

“As previously noted, only 26 participants, the first of four planned groups, were enrolled before the trial was closed because of an unexpectedly high incidence of subjects meeting the predetermined safety criteria for severe AMS (11 subjects or 42%). There was additional impetus to close the study early because one of the research staff also suffered from severe AMS during the study. Thus, this study presents data on only 26 subjects when the intention was to enroll 100 subjects.”

In this case, this study has an important limitation that need to be reflected in its quality assessment. At present, this issue is not considered by the authors of the review.

Minor concerns:

2. Please define in methods how the included trials were considered as at “high Rob”.

3. Please consider if a funnel plot is useful when there are less than 10 studies to analyze (revise the Cochrane Handbook for more information about assessment of publication bias).

4. Authors said: “The results of several pre-planned subgroup analyses with all 7 datasets were similar”. however, these analyses were not detailed in methods. Please include this important information in the corresponding section.

Reply (Reviewer 1)

1. We appreciated the reviewer's important concern. After thorough discussion among co-authors, we decided to change Gerstch 2002 as high risk of bias. The study conducted by Gertsch et al.in 2002, used "first-come first-served basis" after receiving signed consent. Therefore, we judge it as "unclear random-sequence generation". In addition, we appraisal it as incomplete outcome data (attrition bias) because the study presented data on only 26 subjects when the intention was to enroll 100 subjects. We have revised the Table regarding risk of bias in included studies.
2. We appreciated the reviewer's important concern, and we clarified our criteria in method part. According to The Cochrane Collaboration's tool for assessing risk of bias in randomized trials, we defined studies as "high risk of bias" if one or more key domains is taken as high risk in the checklist.
3. We appreciated the reviewer's concern and have removed funnel plot in our content.
4. We appreciated the reviewer's concern and we have revised the method. We conducted subgroup analyses based on quality of studies, starting altitude below 2500 m, number of treatment days before ascending, and dosage of GBE.

Reviewer: 2

1. Was the protocol registered? The name of the registry and the concerned ID should be mentioned in the methodology.

2. What was the definition of 'population' that was included in the meta-analysis. Considering the variability that exists, this is important.

3. Assesment of publication bias by Funnel plot requires at least 10 studies whilst only 6 were included in the present review.

4. There are no details on the outcome measures that were considered in the methods section.

5. It looks like only one outcome measure was assessed - prevention of AMS; There are several clinically important outcomes that are missed by the investigators: incidence of severe AMS, headache and severe headache, oxygen saturation and adverse events.

6. Considering the multiplicity of statistical analyses and small number of events/patients, the authors should consider correcting the type 1 error for the number of analyses.

7. Compliance to PRISMA has not been followed and reported.

Reply (Reviewer 2)

1. We thank for the reviewer's suggestion. We did not register the protocol.
2. We appreciated the reviewer's concern and have clarified our method. We included non-acclimatized adult aged 18 to 60 years. We excluded studies which subjects were pregnant, included participants had symptoms consistent with AMS at baseline.
3. We appreciated the reviewer's suggestion and have removed funnel plot in our content.
4. We appreciated the reviewer's suggestion and have clarified our method about measurement of outcome. AMS defined as AMS-C score≧0.7 or an LLS score≧3 with headache. Primary outcome were the relative risks of AMS in participants receiving GBE for prophylaxis.
5. We appreciated the reviewer's comments. At first we designed (1) incidence of severe AMS, (2)headache and severe headache, (3)oxygen saturation, (4)high altitude pulmonary edema and (5)adverse events of GBE as our secondary outcome. However, there are no similar definitions among studies. Therefore, we decided to present primary outcome finally. We have included an appendix table to summarize the secondary outcomes measured in included studies.
6. We appreciated the reviewer's concern. Although most of our analyses show trends that GBE may prevent AMS, these relative risks are not statistically significant. As a result, we do not think type I error is an issue in our study.
7. We have included a PRISMA checklist in supplement (see online supplementary Checklist). We also mentioned PRISMA guide in method part.

Reviewer: 3

1. Primary endpoint: Please provide information about how the primary endpoint - acute mountain sickness (AMS) is defined in each study.

2. Meta-analysis: there are significant heterogeneities among the six studies. Since the number of studies is small, it is not reliable to analyze the subgroup based on the identified confounders. Authors

can consider testing if there are significant associations between relative risk and dosage, days of treatment before ascent, gender and altitude using a method such as a simple meta-regression (only including one explanatory variable in the model).

3. In the result session, there is no interpretation of the relative risk. It will be informative to explain the relative risk.

4. Please consider adding PRISMA as a guideline.

Other suggestions:

5. Why does the funnel plot only has 5 studies?

6. It will be clearer to add incidence of disease (%) in Table 3; example is table 2 from Lionel Dumont L et al (2000, BMJ), Efficacy and harm of pharmacological prevention of acute mountain sickness: Quantitative systematic review paper.

7. Please consider removing the subgroup analysis forest plots; only keep the summarized RR in result.

8. It is not clear in the forest plot which direction is the placebo and which is the GBE.

9. Please consider adding Dumont L et al (2000, BMJ) paper as a reference, it has identified ascent rate as a significant factor for drug responsiveness and efficacy.

Other minor suggestions:

10. Line40 “Contributorship statement: can change to “author contributions”.

Reply (Reviewer 3)

1. We thank to the reviewer’s advice and we have clarified the definition of AMS in each study (table 1).
2. We appreciate the reviewer’s suggestion. However, the use of meta-regression is not a recommended option when the number of studies is small. Nevertheless, we have conducted meta-regression to evaluate the associations between relative risk and dosage, days of

treatment before ascent and altitude. Since there are no female-only participants in our enrolled studies, we did not perform the regression based on gender. When performed univariate meta-regression regarding dosage (binary variable based on 200mg daily), days of treatment before ascent (binary variable based on 3 days) and altitude (continuous variable), the p-Values are 0.18, 0.77, and 0.96. None of them are statistically significant.

3. We thank to the reviewer's advice, and we clarified interpretation of relative risk in result part. In the primary meta-analysis of all 7 study groups, GBE showed trend of AMS prophylaxis, but it is not statistically significant (RR =0.68; 95% CI: 0.45 to 1.04; p-value=0.08) (Figure 2). The I^2 statistic was 58.7% (p-value=0.02), indicating substantial heterogeneity.
4. We have included a PRISMA checklist in supplement (see online supplementary Checklist). We also mentioned PRISMA guide in method part.
5. We appreciate the reviewer's question. There are two studies (Roncin 1996, and Moraga 2007) reported no event in GBE group. Because of inadequate study numbers, we take another reviewers' suggestion and decided not to include funnel plot as publication bias presentation.
6. We appreciate the reviewer's suggestion and have modified Table 3.
7. We thank to the reviewer's suggestion and have removed subgroup analysis forest plots for simplicity.
8. We modified forest plot figures as the reviewer's recommendation.
9. We have added the reference as the reviewer's suggestion. Although different ascend rate is an interesting factor which may alter the incidence of AMS, we did not have enough information from the enrolled studies.
10. We appreciated the reviewer's suggestion and have revised the title.

Reviewer: 4

1. The article is well written, however I would recommend making tables and figures more concise and reduce the number of them. The major issue of this manuscript is the misinterpretation of the statistical analysis results when p values are above 5%. Is such cases the authors concluded that the null hypothesis was true, which is wrong thing to do. The effect sizes of GBE (in terms of relative risks) in some studies and the pooled relative risks are quite below 1, meaning that the effect might be clinically significant if higher number of subjects

were included in the studies or more studies were conducted. In some studies none of the participants in GBE arm developed AMS while a half of the participants in control arm did develop AMS. Therefore, I cannot agree with your conclusions that currently available data suggest the GBE does not prevent AMS. This sentence should be at least written in a way Cochrane meta-analyses reports such conclusions, e.g. There are not enough data to show the statistically significant effect of GBE for preventing AMA; further studies are warranted…. . Furthermore, as further studies are warranted, I would suggest including/discuss the results of nonrandomized controlled studies (if they were performed) at least in the discussion.

2. In the introduction I lack a paragraph about the information how AMS can be assessed/measured. There are some questionnaires mentioned in the table1 but a more detailed description of possible evaluation of AMS presence is needed in the introduction. Selection about AMS evaluation method should be included as inclusion/exclusion criteria, and in description of the outcome assessment.
3. Page 7, line 10. In what bases the final decisions about overall risks of bias were made? This should be explained here.
4. Page 7, line 40, the main outcome measure is the occurrence of AMS (defined as – see point 1). The RR is one of the options how to calculate the effect size. And you should state here base on what method RR, weights and pooled (overall) RR were calculated - was the Mantel-Haenszel method or the inverse variance method applied? Mantel-Haenszel method is generally preferable when there are few events (see: Cochrane Handbook for Systematic Reviews of Interventions, available online at http://handbook-5-1.cochrane.org/). What method was used to calculate between-trial variability in random effect model? Was the DerSimonian and Laird method? This is major.
5. In two studies none of the participants in GBE arm developed AMS; therefore, it is impossible to calculate RR unless you impute a small number (e.g. 0.5 or 0.3) instead of 0. This is important as the RR estimates can vary significantly. How was this treated in your analysis?
6. Page 7, line 46. Why did not you perform a subgroup analysis based on GBE dosing regimen (there is 2-fold difference in daily dose of GBE used between the included studies)?
7. Tables: Again as explained above you cannot conclude that some studies proved GBE does

not prevent AMS – maybe they were underpowered to achieve statistical significance. So I strongly suggest not dividing the study to one that proved GBE for preventing AMS and other that did not. It is very unusual and statistically incorrect to do so.

8. Tables 1 and 2 can be put together. Results in the Table 3 can be a part of a forest plot. See some forest plot from Review Manager.
9. Figure 3 can be omitted and the pooled (overall) RR (with number of studies and number of participants) can be presented in the text. You have done this, so I would just add the number of studies and number of participants and omit the Figure 3.
10. Figure 4 is irrelevant, funnel plot is meaningful in whole set of data.
11. Since you did not show and difference from subgroup analyses I would suggest to omit the figures 6, 7, and 8 and present the results in the text or create a separate table with results (pooled RR; number of included studies, number of patients in each arm) from sensitivity and subgroup analysis.
12. Figure 5. I would perform such analysis on the whole set of studies, the figure should be marked with line of equality (x axis = 1) not by 1.08. I do not see any benefits of such figure, though. Could be omitted.

Reply (Reviewer 4):

1. We appreciate the reviewer's suggestion and have removed some figures. We appreciate the reviewer's comments about the results interpretation. We do believe type II error might exist in this situation. However, if there is a large study coming out in the future with no preventive effects of GBE on AMS, the pooled result will be moved more toward the null. Furthermore, there is significant heterogeneity if studies with high risk of bias were included. After excluding three high-risk-bias studies, the heterogeneity is decreased and I^2 statistic became 40.2% (p-value=0.17), and the pooled RR is 0.84 (95% CI 0.59 to 1.21; p-value=0.36). We have revised the conclusion based on the reviewer' suggestion. There are not enough data to show the statistically significant effect of GBE for preventing AMS, and further studies are warranted. Of note, there is no non-randomized controlled study associated GBE for preventing AMS so far.

2. We appreciate the reviewer's comments and have modified our introduction and method. Lake Louise Symptom (LLS) Questionnaires and Environmental Symptom Questionnaire III acute mountain sickness–cerebral (AMS-C) section are two tools to diagnose and evaluate severity of AMS. Studies were included in the systematic review if they were (1) randomized controlled trials (RCTs) of healthy non-acclimatized adult between age 18 and 60 years; (2) compared GBE with placebo; (3) conducted in humans; and (4) studies diagnosing AMS with the Lake Louise Criteria or AMS-C. We excluded studies which subjects were pregnant, had symptoms consistent with AMS at baseline. Studies were also excluded if they were irrelevant to the study's aim, were animal studies, lacked a placebo group, or were published as review articles, case reports, editorials, or letters. AMS defined as AMS-C score≧0.7 or an LLS score≧3 with headache. Primary outcome were the relative risks of AMS in participants receiving GBE for prophylaxis.

3. We appreciate the reviewer's suggestion and have revised this part. The Cochrane Collaboration's tool was used to assess the risk of selection, performance, detection, attrition, and reporting biases in the included randomized trials. We defined studies as "high risk of bias" if one or more key domains is taken as high risk in the checklist.

4. We appreciated the reviewer's question. Random effect models with DerSimonian and Laird method were selected for these analyses. Under the random effects model, we assume these studies are drawn from a range of populations in which the effect size varies. We prefer our analyses using random effect model since the estimates would be more conservative.

5. We appreciated the reviewer's question. For the zero cells dealing we add 0.5 to all cells of the 2 × 2 table for the study. This is the general acceptable principle and we would not like to manipulate the data just for significant outcomes. Interestingly, these two studies pose high risk of bias in the study design. After excluding three high-risk-bias studies, the I^2 statistic became 40.2% (p-value=0.17) and the result did not change (RR =0.84; 95% CI 0.59 to 1.21; p-value=0.36).
6. We appreciated the reviewer's suggestion and have conducted the subgroup analysis. Dosage of GBE was not prophylactic for AMS when given "less than 200mg per day" (RR =0.16; 95% CI 0.01 to 2.57; p-value=0.19) or "more than 200mg per day" (RR =0.84; 95% CI 0.59 to 1.21; p-value=0.36).
7. We appreciated the reviewer's suggestion and have modified the tables.

8. There is too much information if we combine Table1 and Table2. We have combined Table 3 and forest plot as the reviewer's suggestion.

9. We thank to the reviewer's suggestion have removed subgroup analysis forest plots.

10. We appreciated the reviewers' concern and have removed funnel plot in our content.

11. We appreciate the reviewer's suggestion and have removed subgroup analysis forest plots.

12. We appreciated the reviewer's suggestion. In the figure of sensitivity analysis, the X-axes as 95% confidence interval. When we did sensitivity analysis to studies in low risk of bias, the result was RR =0.79, and 95% CI 0.58 to 1.08; p-value=0.144. As a result, X-axes are marked on 0.58 and 1.08, respectively. We will take the reviewer's suggestion and perform the sensitivity analysis based on all studies. We also will move the figure as an appendix and leave the description in the manuscript for simplicity.

VERSION 2 – REVIEW

REVIEWER	Irene SL Zeng University of Auckland, New Zealand
REVIEW RETURNED	15-May-2018

GENERAL COMMENTS	The revised version has improved greatly. There remains several places need authors' clarification, listed as follows: 1. In the abstract, it is not clear why there are 7 studies from 6 publications. The information is revealed later in the results section. Please give more clear instruction in this sentence of the abstract: "Six published articles with a total of 451 participants met all eligibility criteria" 2. In the subgroup analysis result, please provide numbers of participants and studies. 3. Method: In the method section, it states: "Pooled relative risk (RR) with corresponding 95% confidence intervals (CIs) for each outcome of interest were calculated." This sentence is not clear to me. Pooled RR is derived from all studies and there is only one outcome. I think the authors indicate pooled RRs are derived for all studies and different subgroups of interest. In the two forest plots, figure one demonstrates the risk difference (RD) and figure two demonstrates the risk ratio (RR). In figure two, please provide pooled outcome as well. Since the risk difference (RD) is also used, it will be better to mention the reason of using RD in the method section. 4. Result: "The number of AMS events and its incidence are summarized in

	Figure 2." The presenting figure 2 is not complied with this description in the result. 5. Please include full label for RD (95% CI) in the legend of figure 1. In the legend, please indicate the direction of the pooled outcome. The pooled Risk Difference is from the active treatment verse placebo?

REVIEWER	Kannan Sridharan Arabian Gulf University, Bahrain.
REVIEW RETURNED	15-May-2018

GENERAL COMMENTS	Previous meta-analyses in this subject has concluded the same as that of this manuscript. Hence, there is no novelty and no new conclusion has been drawn at the end of this meta-analysis.

REVIEWER	Igor Locatelli Faculty of Pharmacy, University of Ljubljana
REVIEW RETURNED	23-May-2018

GENERAL COMMENTS	1. We appreciate the reviewer's suggestion and have removed some figures. We appreciate the reviewer's comments about the results interpretation. We do believe type II error might exist in this situation. However, if there is a large study coming out in the future with no preventive effects of GBE on AMS, the pooled result will be moved more toward the null. Furthermore, there is significant heterogeneity if studies with high risk of bias were included. After excluding three high-risk-bias studies, the heterogeneity is decreased and I2 statistic became 40.2% (p-value=0.17), and the pooled RR is 0.84 (95% CI 0.59 to 1.21; p-value=0.36). We have revised the conclusion based on the reviewer' suggestion. There are not enough data to show the statistically significant effect of GBE for preventing AMS, and further studies are warranted. Of note, there is no non-randomized controlled study associated GBE for preventing AMS so far. Response: Done 2. We appreciate the reviewer's comments and have modified our introduction and method. Lake Louise Symptom (LLS) Questionnaires and Environmental Symptom Questionnaire III acute mountain sickness–cerebral (AMS-C) section are two tools to diagnose and evaluate severity of AMS. Studies were included in the systematic review if they were (1) randomized controlled trials (RCTs) of healthy non-acclimatized adult between age 18 and 60 years; (2) compared GBE with placebo; (3) conducted in humans; and (4) studies diagnosing AMS with the Lake Louise Criteria or AMS-C. We excluded studies which subjects were pregnant, had symptoms consistent with AMS at baseline. Studies were also excluded if they were irrelevant to the study's aim, were animal studies, lacked a placebo group, or were published as review articles, case reports, editorials, or letters. AMS defined as AMS-C score≧0.7 or an LLS score≧3 with headache. Primary outcome were the relative risks of AMS in participants receiving GBE for prophylaxis. Response: Done 3. We appreciate the reviewer's suggestion and have revised this part. The Cochrane Collaboration's tool was used to assess the risk

	of selection, performance, detection, attrition, and reporting biases in the included randomized trials. We defined studies as “high risk of bias” if one or more key domains is taken as high risk in the checklist. Response: Done 4. We appreciated the reviewer’s question. Random effect models with DerSimonian and Laird method were selected for these analyses. Under the random effects model, we assume these studies are drawn from a range of populations in which the effect size varies. We prefer our analyses using random effect model since the estimates would be more conservative. Response: Done 5. We appreciated the reviewer’s question. For the zero cells dealing we add 0.5 to all cells of the 2 × 2 table for the study. This is the general acceptable principle and we would not like to manipulate the data just for significant outcomes. Interestingly, these two studies pose high risk of bias in the study design. After excluding three high-risk-bias studies, the I2 statistic became 40.2% (p-value=0.17) and the result did not change (RR =0.84; 95% CI 0.59 to 1.21; p-value=0.36). Response: I agree, however. You should add a description of “For the zero cells dealing we add 0.5 to all cells of the 2 × 2 table for the study” in the methods (last paragraph) I 6. We appreciated the reviewer’s suggestion and have conducted the subgroup analysis. Dosage of GBE was not prophylactic for AMS when given “less than 200mg per day” (RR =0.16; 95% CI 0.01 to 2.57; p-value=0.19) or “more than 200mg per day” (RR =0.84; 95% CI 0.59 to 1.21; p-value=0.36). Response: Done 7. We appreciated the reviewer’s suggestion and have modified the tables. Response: Done 8. There is too much information if we combine Table1 and Table2. We have combined Table 3 and forest plot as the reviewer’s suggestion. Response: New figure 2 has the info regarding event rates, but there should also be some information regarding each study RR , and pooled RR. Add the information about weights (as had been shown in previous version) 9. We thank to the reviewer’s suggestion have removed subgroup analysis forest plots. Response: Done 10. We appreciated the reviewers’ concern and have removed funnel plot in our content. Response: Done 11. We appreciate the reviewer’s suggestion and have removed subgroup analysis forest plots. Response: Done 12. We appreciated the reviewer’s suggestion. In the figure of sensitivity analysis, the X-axes as 95% confidence interval. When we did sensitivity analysis to studies in low risk of bias, the result was RR =0.79, and 95% CI 0.58 to 1.08; p-value=0.144. As a result,

	X-axes are marked on 0.58 and 1.08, respectively. We will take the reviewer's suggestion and perform the sensitivity analysis based on all studies. We also will move the figure as an appendix and leave the description in the manuscript for simplicity. Response: Your pooled results of 7 included studies is 0.68 (95% CI: 0.45 – 1.04) so I assume that these sre the limits you want to show on supplementary figure 2. Currently you have 0.46, 0.69, and 1.04). In the text you mention "only changed the pooled estimate from 0.79 to 0.74" Shouldn't be from 0,69 to 0.74?

VERSION 2 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 3

1. In the abstract, it is not clear why there are 7 studies from 6 publications. The information is revealed later in the results section. Please give more clear instruction in this sentence of the abstract:

"Six published articles with a total of 451 participants met all eligibility criteria"

2. In the subgroup analysis result, please provide numbers of participants and studies.

3. Method:

In the method section, it states:

"Pooled relative risk (RR) with corresponding 95% confidence intervals (CIs) for each outcome of interest were calculated."

This sentence is not clear to me. Pooled RR is derived from all studies and there is only one outcome. I think the authors indicate pooled RRs are derived for all studies and different subgroups of interest.

In the two forest plots, figure one demonstrates the risk difference (RD) and figure two demonstrates the risk ratio (RR). In figure two, please provide pooled outcome as well. Since the risk difference (RD) is also used, it will be better to mention the reason of using RD in the method section.

4. Result:

"The number of AMS events and its incidence are summarized in Figure 2."

The presenting figure 2 is not complied with this description in the result.

5. Please include full label for RD (95% CI) in the legend of figure 1.

In the legend, please indicate the direction of the pooled outcome. The pooled Risk Difference is from the active treatment verse placebo?

Reply:

1. We appreciate the reviewer's advice and we did clear instruction in the abstract.
2. We appreciate the reviewer's concern. Regarding to the numbers of participants, there are 357 participants enrolled in High quality subgroup; 94 in low quality subgroup; 208 participants in the starting altitude was below 2500m subgroup; 114 participants in "3–5 days prior to ascent" subgroup; 337 participants in "0–2 days prior to ascent" subgroup; 94 participants in "dosage less than 200mg per day" subgroup; 357 participants in "dosage more than 200mg per day" subgroup. Information regarding number of participants and enrolled studies in each subgroup are summarized in supplementary table 2.

Supplementary table 2. Number of participants and enrolled studies in each subgroup

Subgroup	Number of participants
Low risk of bias studies	357
Gertsch, 2004	
Chow, 2005	
Leadbetter, 2009	
Starting altitude below 2500m	208
Roncin, 1996	
Gertsch, 2002	
Chow, 2005	
Moraga, 2007	
Leadbetter, 2009	
3–5 days prior to ascent	114
Chow, 2005	
Leadbetter, 2009	
0–2 days prior to ascent	337
Roncin, 1996	
Gertsch, 2002	
Gertsch, 2004	
Moraga, 2007	
Dosage less than 200mg per day	94
Roncin, 1996	
Gertsch, 2002	
Moraga, 2007	

Dosage more than 200mg per day		357
	Gertsch, 2004	
	Chow, 2005	
	Leadbetter, 2009	

3. We appreciate the reviewer's advice and we modified our sentence. We also mention the reason of using RD in the method section.
4. We are sorry for the confusion and have revised our figure legends.
5. We have revised the figure 3 as the pooled results of risk difference and have added direction of treatment effects.

Reviewer: 2

Previous meta-analyses in this subject has concluded the same as that of this manuscript. Hence, there is no novelty and no new conclusion has been drawn at the end of this meta-analysis.

Reply:

We appreciate the reviewer's comments.

Reviewer: 4

1. You should add a description of "For the zero cells dealing we add 0.5 to all cells of the 2 × 2 table for the study" in the methods (last paragraph)
2. New figure 2 has the info regarding event rates, but there should also be some information regarding each study RR, and pooled RR. Add the information about weights (as had been shown in previous version)
3. Your pooled results of 7 included studies is 0.68 (95% CI: 0.45 – 1.04) so I assume that these are the limits you want to show on supplementary figure 2. Currently you have 0.46, 0.69, and 1.04). In the text you mention "only changed the pooled estimate from 0.79 to 0.74" Shouldn't be from 0,69 to 0.74?

Reply:

1. We appreciated for the reviewer's advice, and we added the clear instruction of zero cells dealing in the last paragraph of the method.
2. We have revised the figure 2 as the reviewer's suggestion.
3. We appreciated to the reviewer's question. Supplementary figure 1 revealed sensitivity analysis, and marked lines indicated RR and confidence interval of meta-analysis in primary outcome (RR =0.68; 95% CI: 0.45 to 1.04). We have revised the numbers. We are sorry for the confusion and we believe 0.79 is a typo.

VERSION 3 – REVIEW

REVIEWER	Irene SL Zeng University of Auckland, New Zealand
REVIEW RETURNED	24-Jun-2018

GENERAL COMMENTS	Authors have addressed all points raised in my previous review. I recommend accepting it for the publication.