PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Short implants (≤6 mm) versus longer implants with sinus floor elevation in atrophic posterior maxilla: a systematic review and meta-analysis
AUTHORS	Yan, Qi; Wu, Xinyu; Su, Meiying; Hua, Fang; Shi, Bin

VERSION 1 – REVIEW

REVIEWER	Thomas Starch-Jensen, professor, chief surgeon, PhD Department of Oral and Maxillofacial Surgery Aalborg University Hospital Aalborg Denmark
REVIEW RETURNED	25-Feb-2019

GENERAL COMMENTS	Dear Adrian Aldcroft Editor in Chief, BMJ Open Ultrashort implants versus longer implants with sinus floor elevation in severely atrophic posterior maxilla: a systematic review and meta-analysis. The topic of the present systematic review and meta-analysis is very interesting and relevant. Moreover, only randomized controlled trials were included and PRISMA guidelines was followed, which makes the results of the present systematic review and meta-analysis trustworthy. However, newly published systematic reviews assessing identical topic with similar conclusions already exist, which limits the novelty of the current systematic review. Moreover, the present systematic review and meta-analysis includes maxillary sinus floor augmentation and osteotome-mediated sinus floor augmentation, without mentioning the residual alveolar vertical bone height of the included studies as well as a discussion of the influence of the residual alveolar vertical bone height on implant survival rate. See following publications (Del Fabbro M. Implant survival rates after osteotome-mediated maxillary sinus augmentation: a systematic review. Clin Implant Dent Relat Res. 2012 May;14 Suppl 1:e159-68 and Pjetursson BE. Maxillary sinus floor elevation using the (transalveolar) osteotome technique with or without grafting material. Part I: Implant survival and patients' perception. Clin Oral Implants Res. 2009 Jul;20(7):667-76. In addition, I have following comments to the manuscript: 1. According to PRISMA Checklist, PICOS question should be described on page 5. Where? 2. I would recommend using the term short implants (≤ 6 mm) and not ultrashort implant. 3. Key words: use MESH words in alphabetical order 4. The authors use the term severely atrophic posterior maxilla in the title, and includes studies involving osteotome-mediated sinus

	floor augmentation. Indications for osteotome-mediated sinus floor augmentation is not severely atrophic posterior maxilla. Moreover, The authors write in the objective “The present systematic review aimed to compare the effectiveness of ultrashort implants and longer implants with sinus floor elevation in atrophic posterior maxilla”. Confusing. 5. What is the hypothesis of the present systematic review? 6. In the objective of the systematic review, I would recommend to define ultrashort implants e.g. The present systematic review aimed to compare the effectiveness of ultrashort implants (≤ 6 mm) and longer implants. 7. Important outcome measures like survival of suprastructures, patient-related or professional-related outcome measures are not included. Why? 8. I would recommend using biological and technical complications and not adverse events. 9. The authors write in the introduction that “These conditions are exacerbated if patients have a history of wearing removable dentures”, is there enough evidence for this statement? 10. The authors write in the introduction that “The lateral window approach is used in up to 22.1% of all dental implantation procedures”. From my point of view, this is not true! 11. The authors write in eligible criteria “One year or longer follow-up period”, is it before or after loading. Please define. 12. Residual bone height was registered according to the specifically designed data extraction forms, but not reported or discussed in the manuscript. Why? 13. Ultrashort implants with a single crown as well as ultrashort implants with splinted prosthetic solutions were included in the present systematic review. I would recommend (if possible) describing eventually differences in implant survival or peri-implant marginal bone loss between these two different prosthetic solutions among the included studies.

REVIEWER	Derek Smith Vanderbilt University Medical Center
REVIEW RETURNED	10-Mar-2019

GENERAL COMMENTS	This is a well written, well-conducted review of ultra short implant vs. conventional +sinus lift. It appears to have been conducted according to the guidelines laid out in the Cochran review handbook. Additionally, all PRISMA reporting appears to be included and accurate. I have no major concerns with this study whatsoever. One minor comment, if the forest plots are to rendered in color then they should in fact be the same color.

REVIEWER	Tzy-Jyun Yao Harvard TH Chan School of Public Health
REVIEW RETURNED	15-Apr-2019

GENERAL COMMENTS	Ultrashort implant versus longer implant with sinus floor elevation in severely atrophic posterior maxilla: a systemic review and meta-analysis • Abstract: Should report the analysis results of late adverse events in Results, not just mention in Conclusions. • Materials and Methods:

	o Should describe the background of the reviewers, whether they were area experts or methodologists. o The writing style of this section is somewhat confusing. It would be helpful to report what was conducted specific to this study, rather than statements with " we would have" (Sections 2.6, 2.8). More specific description of "outcome" should be described (Section 2.7) since this was how the author determined whether a meta-analysis would be conducted. o Section 2.7: Which specific method was used in the meta-analysis for relative risk and for mean difference? They should be described even though some were shown in Figures 4 & 5. o Section 2.7: The determination of whether to apply a fixed effect model or a randomeffect model was based on the number of studies in the analysis, rather than on the interpretation of these models. This seems arbitrary. A relevant point regarding sensitivity analysis is commented below for Results section 3.4.3. • Results: o Section 3.1: The date when the literature search was conducted should be reported. o Section 3.2: There should be some descriptions, such as setting, age, sex and country, for the participants in the included studies. o Section 3.3: For the comparison between these two treatments, the detection bias is unavoidable since the assessors could recognize sinus floor elevation. However, for unbalanced dropout, have the authors contacted authors of these studies for reasons of dropping out? Or, if information was not obtainable, conducted sensitivity analysis treating the dropouts as failures? o Section 3.4.1: should it be five instead of four studies reported 100% survival of ultrashort implants? o Section 3.4.2: For MBL one year or longer post-loading, the heterogeneity is very high. A sensitivity analysis using fixed effect model should therefore be performed. For MBL three years or longer post-loading, there were only three studies and thus according to Section 2.7, a fixed effect model would be applied. However, the heading of Figure 5 indicated random-effect models. Which one was applied? o Section 3.4.3. No meta-analysis was performed for adverse events, one of the three major outcomes. Even an implant-level analysis is better than no analysis at all. There were almost twice late adverse events in the ultrashort implant group than in the elevation group. This non-trivial finding should be further analyzed and discussed in Discussion. • Discussion: The authors discussed the influence of the implant length and surface structure in survival rate. Apparently, the surface structure was different in different studies, but was not reported even though the authors consider it an important factor. • Table 5 summary of findings: should include AE and it's risk assessment • Figure 1: Almost 80% of screened studies were excluded due to "Not regarding this study". More detailed exclusion reasons should be given. • Editorial point: Section 2.3, last line: "Corresponding authors were contacted for missing data or information." Instead of conducted.

VERSION 1 – AUTHOR RESPONSE

Comments from Reviewer #1:
1. According to PRISMA Checklist, PICOS question should be described on page 5. Where?

Response: Revision done in section 2.2. The PICOS question was in the eligible criteria (section 2.2) but not clearly described. We have re-written section 2.2 to make it clearer.

2. I would recommend using the term short implants (≤ 6 mm) and not ultrashort implant.

Response: Revision done.

3. Key words: use MESH words in alphabetical

order. Response: Thank you. Revision done.

4. The authors use the term severely atrophic posterior maxilla in the title, and includes studies involving osteotome-mediated sinus floor augmentation. Indications for osteotome-mediated sinus floor augmentation is not severely atrophic posterior maxilla. Moreover, the authors write in the objective "The present systematic review aimed to compare the effectiveness of ultrashort implants and longer implants with sinus floor elevation in atrophic posterior maxilla". Confusing.

Response: Thank you. We have deleted the expression 'severely'.

5. What is the hypothesis of the present systematic review?

Response: Thank you. We have added a hypothesis sentence to the last paragraph of 'Introduction'.

6. In the objective of the systematic review, I would recommend to define ultrashort implants e.g. The present systematic review aimed to compare the effectiveness of ultrashort implants (≤ 6 mm) and longer implants.

Response: We have used "short implant (≤6 mm)" to replace the pervious term "ultrashort implant". Thank you.

7. Important outcome measures like survival of suprastructures, patient-related or professional-related outcome measures are not included. Why?

Response: Revision done in 3.4.3. Initially, we categorized survival of supra-structures into adverse events and described them. We have used complications to replace 'adverse events'. We have categorized survival of supra-structures into technical complications and conducted meta-analysis for each category of complication.
We did not report patient satisfaction because only three studies 1-3 reported patient satisfaction, and the methods used to evaluate patient satisfaction were different among studies. Therefore, we have added patient satisfaction outcomes in section 3.4.4 without conducting a meta-analysis. Thank you.

8. I would recommend using biological and technical complications and not adverse

events. Response: Revision done. Thank you.

9. The authors write in the introduction that "These conditions are exacerbated if patients have a history of wearing removable dentures", is there enough evidence for this statement?

Response: Thank you. This statement was mentioned briefly in a previous study -- “a common problem that hinders restoration of edentulous sites with implant-supported prostheses is bone resorption, which may be more pronounced if the patient has been wearing a removable prosthesis”. 4

10. The authors write in the introduction that “The lateral window approach is used in up to 22.1% of all dental implantation procedures”. From my point of view, this is not true!

Response: Thank you. The data was mentioned in a previous cross-sectional study. 5 Considering that these data may not be applicable to all, we have used “commonly” to replace the specific figure.

11. The authors write in eligible criteria “One year or longer follow-up period”, is it before or after loading. Please define.

Response: The eligible criteria should be “a follow-up length of one year or longer post-loading”. Revision Done in section 2.2. Thank you.

12. Residual bone height was registered according to the specifically designed data extraction forms, but not reported or discussed in the manuscript. Why?

Response: We have added the inclusion criteria for residual bone height of each RCT to table 1 “Characteristics of the included studies”. Thank you.

13. Ultrashort implants with a single crown as well as ultrashort implants with splinted prosthetic solutions were included in the present systematic review. I would recommend (if possible) describing eventually differences in implant survival or peri-implant marginal bone loss between these two different prosthetic solutions among the included studies.

Response: Thank you for your suggestion. Two studies included only single crowns while the other five studies included single crown and splinted prosthetic. Subgroup analysis was not applicable here.

Comments from Reviewer #2:
1. This is a well written, well-conducted review of ultrashort implant vs. conventional +sinus lift. It appears to have been conducted according to the guidelines laid out in the Cochran review handbook. Additionally, all PRISMA reporting appears to be included and accurate. I have no major concerns with this study whatsoever.

Response: Thank you for your comments.

2. One minor comment, if the forest plots are to rendered in color then they should in fact be the same color.

Response: Thank you. Please note that the forest plots were generated with RevMan 5.3, which use different colours to demonstrate different outcome measures automatically. Blue means dichotomous data calculated using Mantel-Haenszel method, while Green means continuous data calculated using inverse variance method.

Comments from Reviewer #3:
1. Abstract: Should report the analysis results of late adverse events in Results, not just mention in Conclusions.

Response: Thank you for your comments. We have replaced “adverse events” with “complications” according to the suggestion of reviewer #1 and we have added outcome for complications in the abstract.

2. Should describe the background of the reviewers, whether they were area experts or methodologists.
Response: Revision done. Thank you.

3. The writing style of this section is somewhat confusing. It would be helpful to report what was conducted specific to this study, rather than statements with “we would have” (Sections 2.6, 2.8). More specific description of “outcome” should be described (Section 2.7) since this was how the author determined whether a meta-analysis would be conducted.

Response: In our protocol (PROSPERO CRD42018103531), we stated that we would conduct funnel plot or the Egger’s test 6 if ten or more studies were included. Although less than ten studies were included in this systematic review, we should still report it as required by PRISMA. The statement involving “we would have” is commonly used for this kind of situation in Cochrane reviews.7
Please note that in section 2.7 we did mention our standards for carrying out meta-analyses -- “Meta-analyses were undertaken only when at least two studies that made similar comparisons reported the same outcomes.” Thank you.

4. Section 2.7: Which specific method was used in the meta-analysis for relative risk and for mean difference? They should be described even though some were shown in Figures 4 & 5.

Response: Revision done in “2.7 synthesis of results”. Thank you.

5. Section 2.7: The determination of whether to apply a fixed effect model or a random- effect model was based on the number of studies in the analysis, rather than on the interpretation of these models. This seems arbitrary. A relevant point regarding sensitivity analysis is commented below for Results section 3.4.3.

Response: As stated in our review protocol, we determined whether to choose fixed- or random-effect model based on the number of included studies. This method was described in a previous methodological study 8 and has been routinely used in Cochrane reviews.7 9 10

6. Section 3.1: The date when the literature search was conducted should be reported.
Response: We mentioned the date in section 2.3. Because search date usually followed database in methods in Cochrane reviews.7 Thank you.

7. Section 3.2: There should be some descriptions, such as setting, age, sex and country, for the participants in the included studies.

Response: Revision done. We have added information into table 1 “Characteristics of the included studies”, including gender, age and residual bone height of patients in each RCT. Thank you

8. Section 3.3: For the comparison between these two treatments, the detection bias is unavoidable since the assessors could recognize sinus floor elevation. However, for unbalanced dropout, have the

authors contacted authors of these studies for reasons of dropping out? Or, if information was not obtainable, conducted sensitivity analysis treating the dropouts as failures?

Response: 1. Thank you. Unbalanced drop-out could be resulted from reasons and / or proportion of dropping out. According to the Cochrane handbook, when no detail (either reason or number) for drop-out is provided, authors should be contacted for missing information. In this systematic review, only one study11 did not provide specific reasons for dropping out. We contacted the corresponding author, but have not received any reply. In the rest of the studies, the reasons, number and percentage for dropping out were provided and presented in table 2 “Details on the risk of bias for each included study”. When the reasons and proportion of drop-out were balanced between two groups and the proportion was<10%, we considered the risk of bias low; when reasons were not provided and proportion was unbalanced or proportion≥10%, we considered the risk of bias high. 3. Treating drop-outs as failures may not be appropriate. In clinical practice, patients who did not show up in follow-up appointments usually have no apparent symptom.

9. Section 3.4.1: should it be five instead of four studies reported 100% survival of ultrashort implants?

Response: Revision done. Thank you.

10. Section 3.4.2: For MBL one year or longer post-loading, the heterogeneity is very high. A sensitivity analysis using fixed effect model should therefore be performed. For MBL three years or longer post-loading, there were only three studies and thus according to Section 2.7, a fixed effect model would be applied. However, the heading of Figure 5 indicated random-effect models. Which one was applied?

Response: For MBL, more than three studies were included in the meta-analysis and therefore we chose to use random-effects model. This is a method routinely used in reviews conducted by the Cochrane Oral Health Group7. We have added a sensitivity analysis by using fixed-effect model, which indicated that our result was robust. Thank you for your suggestion.

11. Section 3.4.3. No meta-analysis was performed for adverse events, one of the three major outcomes. Even an implant-level analysis is better than no analysis at all. There were almost twice late adverse events in the ultrashort implant group than in the elevation group. This non-trivial finding should be further analyzed and discussed in Discussion.

Response: We have replaced “adverse events” with “complications” and categorized them. For each specific category of complication, we have conducted meta-analysis if two or more studies reporting the same complication categorization were detected (figures 6-8). Thank you.

12. Discussion: The authors discussed the influence of the implant length and surface structure in survival rate. Apparently, the surface structure was different in different studies, but was not reported even though the authors consider it an important factor.

Response: Surface structure was an important factor in improving implant survival in the past. However, as technique advancing, implant surface available in market and used in studies could meet the need of most clinical practice. Different surface modification among implants was not a determining factor. So, we put surface structure in risks of bias assessment. Thank you for your suggestion.

13. Table 5 summary of findings: should include AE and it’s risk assessment

Response: Revision done. Thank you.

14. Figure 1: Almost 80% of screened studies were excluded due to “Not regarding this study”. More detailed exclusion reasons should be given.

Response: Revision done. Thank you.

15. Editorial point: Section 2.3, last line: “Corresponding authors were contacted for missing data or information.” Instead of conducted.
Response: Revision done. Thank you for your suggestion.

References

1. Gastaldi G, Felice P, Pistilli R, et al. Short implants as an alternative to crestal sinus lift: a 3-year multicentre randomised controlled trial. European Journal of Oral Implantology 2017;10(4):391-400.
2. Bechara S, Kubilius R, Veronesi G, et al. Short (6-mm) dental implants versus sinus floor elevation and placement of longer (>/=10-mm) dental implants: a randomized controlled trial with a 3-year follow-up. Clinical Oral Implants Research 2017;28(9):1097-107.
3. Guljé FL, Raghoebar GM, Vissink A, et al. Single crowns in the resorbed posterior maxilla supported by either 6-mm implants or by 11-mm implants combined with sinus floor elevation surgery: a 1-year randomised controlled trial. European Journal of Oral Implantology 2014;7(3):247-55.
4. Malchiodi L, Caricasulo R, Cucchi A, et al. Evaluation of Ultrashort and Longer Implants with Microrough Surfaces: Results of a 24- to 36-Month Prospective Study. International Journal of Oral & Maxillofacial Implants 2017;32(1):171-79.
5. Cha HS, Kim JW, Hwang JH, et al. Frequency of bone graft in implant surgery. Maxillofacial plastic and reconstructive surgery 2016;38(1):19.
6. Egger M, Davey SG, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. Bmj British Medical Journal 1997;316(7129):469-71.
7. Esposito M, Grusovin MG, Felice P, et al. Interventions for replacing missing teeth: horizontal and vertical bone augmentation techniques for dental implant treatment. The Cochrane database of systematic reviews 2009(4):Cd003607.
8. Borenstein M, Hedges LV, Higgins JP, et al. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods 2010;1(2):97-111.
9. Esposito M, Ardebili Y, Worthington HV. Interventions for replacing missing teeth: different types of dental implants. The Cochrane database of systematic reviews 2014(7):Cd003815.
10. Esposito M, Felice P, Worthington HV. Interventions for replacing missing teeth: augmentation procedures of the maxillary sinus. The Cochrane database of systematic reviews 2014(5):Cd008397.
11. Pohl V, Thoma DS, Sporniak-Tutak K, et al. Short dental implants (6 mm) versus long dental implants (11-15 mm) in combination with sinus floor elevation procedures: 3-year results from a multicentre, randomized, controlled clinical trial. Journal of Clinical Periodontology 2017;44(4):438-45.

VERSION 2 – REVIEW

REVIEWER	Tzy-Jyun Yao Harvard T.H. Chan School of Public Health, Boston, USA
REVIEW RETURNED	31-Jul-2019

GENERAL COMMENTS	Short implant (≤6 MM) versus longer implant with sinus floor elevation in atrophic posterior maxilla: a systemic review and meta-analysis • It was apparent that patient satisfaction was not one of the pre-determined outcomes as an eligible criterion. Only 3 out of 7 studies reported some collected information on patient satisfaction, no systemic analysis could be performed. Therefore, patient satisfaction should not be treated with equal weight as the other outcomes, but rather just a descriptive report from information in the 3 studies. Any description about patient satisfaction should be given less weight, and should not be part of conclusion. All relevant sections should be modified to reflect the secondary nature of this outcome. • Materials and Methods: o Section 2.9: Categories of complication were only listed in Table 4 and never described in the text. Other than post-surgery reaction, which I assume was (how?) soon after surgery, the timing when these complications occurred was not well defined. Therefore, it is not clear how subgroup analysis by categories of complications could "control for the possibility that function time might influence implant survival". • Results: o Section 3.4.3: For all types of complications, there were only three or less studies had reported information, and thus according to Section 2.7, a fixed effect model would be applied. However, fixed effect model was only applied for post-surgery reaction, all other applied random-effect models, contradicting to analysis plan. o Section 3.4.3: "…while in 20% patient (1/10) in sinus floor elevation group", either 20% or 1/10 is wrong. Why the denominators of complications ("Total" column in Figures 68) were less than the sample size of the study? For example, the sample size for SI group in Bolle 2018 should be 20, but 10 was listed under "Total", why half of the participants were excluded in the analysis? o Section 3.4.3: It was mentioned that there were many more late adverse events in the short implant group than in the elevation group. Which complications were late? Why now short implant group had fewer complications than the elevation group in all categories? o Section 3.4.4: Bechara et al studied patient satisfaction in multiple aspects, but only the result of cost was mentioned. What about the other aspects? It looked selective to only report partial findings, all in favor of the short implant group. In addition, this result in cost was emphasized several times throughout the article. It doesn't seem appropriate in a systemic review report to highlight finding based on one aspect of on single publication. • Discussion: The rationale of subgroup analysis by categories of complications was never given, neither was why these particular categories were chosen. It is hard to judge whether the subgroup analysis by categories of complications actually reduce or increase bias.

VERSION 2 – AUTHOR RESPONSE

Response to Reviewer#3

1 It was apparent that patient satisfaction was not one of the pre-determined outcomes as an eligible criterion. Only 3 out of 7 studies reported some collected information on patient satisfaction, no systemic analysis could be performed. Therefore, patient satisfaction should not be treated with equal weight as the other outcomes, but rather just a descriptive report from information in the 3 studies. Any description about patient satisfaction should be given less weight, and should not be part of conclusion. All relevant sections should be modified to reflect the secondary nature of this outcome.

Response: thank you. We have revised the manuscript to give less weight to patient satisfaction, and deleted relevant content from the conclusion.

2 Section 2.9: Categories of complication were only listed in Table 4 and never described in the text. Other than post-surgery reaction, which I assume was (how?) soon after surgery, the timing when these complications occurred was not well defined. Therefore, it is not clear how subgroup analysis by categories of complications could “control for the possibility that function time might influence implant survival”.

Response: We have added a sentence in section 2.9 describing categories of complications. In addition, in section 2.9, the first sentence was not clear (“Subgroup analysis by length of follow-up and categories of complications was performed to control for the possibility that function time might influence implant survival”). Subgroup analyses by length of follow-up was used to control for the possibility that function time might influence implant survival. While subgroup analyses by complication categories were used to evaluate whether the incidence of different categories of complications were different between the intervention and control groups. We have reorganized the language to make our meaning clearer.

3 Section 3.4.3: For all types of complications, there were only three or less studies had reported information, and thus according to Section 2.7, a fixed effect model would be applied. However, fixed effect model was only applied for post-surgery reaction, all other applied random-effect models, contradicting to analysis plan.

Response: We have revised the effect models and made a comparison table (table 4) to describe the results of complications.

4 Section 3.4.3: “...while in 20% patient (1/10) in sinus floor elevation group”, either 20% or 1/10 is wrong. Why the denominators of complications (“Total” column in Figures 6- 8) were less than the sample size of the study? For example, the sample size for SI group in Bolle 2018 should be 20, but 10 was listed under “Total”, why half of the participants were excluded in the analysis?

Response: Revision done. We input a wrong denominator of complication categories of the study of Bolle’s. The total number in each group should have been 19 instead of 10. We have revised this mistake. The results of “post-surgery reaction” and “sinus perforation or infection” remained significant. In addition, we have checked all the other data relating to complication (section 3.4.3 and table 4) to avoid such careless mistakes.

5 Section 3.4.3: It was mentioned that there were many more late adverse events in the short implant group than in the elevation group. Which complications were late? Why now short implant group had fewer complications than the elevation group in all categories?

Response: Initially, “late complication” was defined as complications that occurred after delivery of definitive crowns, including peri-implant mucositis and peri-implantitis, screw loosening, crowns loosening, chipping and decementation. In fact, short implant groups had more “late complications” described above (15 in total) than sinus floor elevation group (7 in total). Results of meta-analyses suggested that the difference was not significant. While “early complication” was defined as complications occurring from surgery to the delivery of definitive crowns, including post-surgery reaction, sinus perforation or infection and implant mobile. Short implant groups had less “early complications” described above than sinus floor elevation group.

6 Section 3.4.4: Bechara et al studied patient satisfaction in multiple aspects, but only the result of cost was mentioned. What about the other aspects? It looked selective to only report partial findings, all in favor of the short implant group. In addition, this result in cost was emphasized several times throughout the article. It doesn’t seem appropriate in a systemic review report to highlight finding based on one aspect of on single publication.

Response: Revision done. 1) We have added a sentence in section 3.4.4 to describe all results about patient satisfaction. In Bechara’s study, the authors evaluated patient satisfaction in five aspects while only the results of patient satisfaction towards cost was significantly different between intervention and control group. 3.4.4. 2) We have limited the use of “cost” only in introduction and discussion.

7 Discussion: The rationale of subgroup analysis by categories of complications was never given, neither was why these particular categories were chosen. It is hard to judge whether the subgroup analysis by categories of complications actually reduce or increase bias

Response: One sentence added in section 2.9. Thank you.