PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	How effective is undergraduate palliative care teaching for medical students? A systematic literature review
AUTHORS	Boland, Jason; Brown, Megan; Duenas, Angelique; Finn, G; Gibbins, Jane

VERSION 1 – REVIEW

REVIEWER	Geoffrey Wells Department of Medical Education Brighton and Sussex Medical School United Kingdom
REVIEW RETURNED	21-Jan-2020

GENERAL COMMENTS	Firstly, many thanks for conducting such a valuable review. This is such an important area and will be a great addition to the steadily increasing literature that is supporting palliative care education, and shaping future curricula. I have made several points below which are purely for your consideration, and are based on my own experience of publishing a systematic review. Some are issues relating to wording and sentence structure, others are seeking clarification and greated depth of information regarding your choice and interpretation of tools you have used, and suggestions on how your search could be a little more inclusive (and therefore systematic). I have numbered my comments in the hope it will make things easier to follow. I hope you find them helpful. Format and editing: Abstract: Very first sentence. When you state 'most experienced doctors' do you mean those doctors with the most experience, or do you mean the majority of experienced doctors? If the former – then what about those doctors with lesser experience. If you mean the latter – then why not all experienced doctors? I would argue the point that palliative care is everyone's business irrespective of level of seniority and experience. You do go on later to state (page 4) 'medical students and doctors require the appropriate knowledge, skills and attitudes to care for patients who have advance and incurable illness', followed by 'The ability to care for, and communicate appropriately with all these patients and their families is an essential skill for all doctors. 1. I would consider amending the first sentence of your abstract for continuity and to reflect/support your later quotes. 2. Page 3 : Strengths and limitations: 'This was a rigorously conducted systematic review including "grey" literature and evaluating the quality of the individual included studies'.

	Sentence does not read well – consider restructuring. 3. Background: (Page 4) '40 patients who will die'. I wonder if this would better read as '40 actively dying patients'. I am nit-picking here, but as all patients/people will die eventually, it may just read better the other way around. Methods: Search terms used included: 'palliative care', 'medical student', 'teaching' What about the use of the search term 'medical undergraduates' (or med* adj3 undergrad*)? I would imagine there could be articles which may be entitled 'How are medical undergraduates being taught palliative care?, or 'How effective is palliative care teaching in undergraduate medicine?' or 'Does undergraduate medical curricula effectively teach palliative care?' Also what about the use of the search term 'education'?, again an article could read 'The effectiveness of palliative care teaching within medical education' These may well have been included within your MeSH headings – but may be worth just checking and clarifying (or making this more explicit) 4. There is the potential for some articles to have been missed by not including the above search terms and I wonder if you would consider re-running your search to include these items. I suspect it will not add many (if any) articles to your review, but would definitely prove to be a more robust and systematic search strategy. Eligibility: 5. Could the authors clarify (or perhaps give an example of) what objective measure they may hoped to have seen from qualitative studies, given the statement that studies with qualitative outcomes were excluded. I was unclear about what was being looked for. If qualitative outcomes were to be excluded, why were they included within the search criteria? 6. Length of follow up: No restrictions. Did you find this potentially lead to bias in some studies as a longer time period between intervention and follow up may affect the effectiveness of the results due to loss of knowledge / acquisition of new knowledge from other sources in the meantime. You may have mentioned this – but I couldn't see where. Quality assessment: (p7) and Quality assessment (p8) 7. Needs greater description of what the MMAT is for those who haven't heard of it What is it scored out of, how is this interpreted? The same applies to the Cochrane risk of bias tool. What is it? How does it score? How should readers interpret it? Why have you chosen these two tools specifically? Given MMAT gives a score between 0-100%, this is not recorded on the data extraction tables. You have stated – passed all points. What does this mean? Please consider documenting the actual percentage scores for those studies where the MMAT was applied. Did they all really achieve 100%?

	What was the level of agreement between reviewers of study quality? It would be good to see a measure of inter-rater variability, and a calculation of a Cohen's kappa score as this will add to the rigour of this review. Results: 8. In the interests of transparency and consistency, I wonder if it is worth stating which author(s) undertook which parts of the screening process, in the same way you did for data extraction. Who screened the titles, the full text? 9. You mention a range of participant numbers between studies. It may be worth considering stating what the publication year range is as well in this paragraph. 10. Page 19. Different assessment methods. OSCE used for first time in the main text. Needs writing in full. Even though it is a commonly used abbreviation, so is MCQ (which you have written in full). Synthesis of results: 11. I wonder if authors would consider including within the body of the text some of the statistical significance of the results, especially when referring to studies that have demonstrated overall improvement in knowledge scores. P values may be helpful for readers to put things into context. At the moment the reader has to keep going back to the table and working out which of the studies demonstrated statistical significance, no statistical significance, or simply did not apply stats to their results. This wouldn't take long to do and would allow the context of the results section to be more readily apparent, with an easier flow. 12. Again, when talking about the interactive eLearning course reporting equivalence in increased knowledge scores when compared with small-group teaching…what was the level of statistical significance to support this? Outcomes and constructive alignment considerations: I felt this was very well written, with very clear examples of what is meant by constructive alignment. Just a few wording/syntax issues. 13. Page 21 line 49. Advance care planning, not 'advanced'. 14. Page 22 line 13. Should it read: 'it is not possible to know whether constructively aligned learning outcomes, teaching, and assessment are important to effective palliative care teaching'? 15. Page 22 line 51. Should this read 'dispel these' rather than 'dismiss this'? 16. Page 23, line 3. I generally avoid beginning sentences with 'And' as it does not read well. Please consider rewording this. 17. Page 23, line 31. I wonder if this should read 'demonstrated improved knowledge amongst these students'

REVIEWER	Jie Chen Center for Global Health, Zhejiang university, CHINA
REVIEW RETURNED	CHINA 11-Feb-2020

GENERAL COMMENTS	I have to say this is a well-written paper for discussing palliative care teaching for medical students. The authors tried a lot to find the relationship between education and its potential influence on behaviours while the final results are not up to their expectations. This team followed a strict guideline and did the review systematically with the help of a clinical librarian. After finishing reading this paper, I still have the following questions: I checked the registration on PROSPERO(CRD42018115257) and noticed that there were two different versions. Furthermore, both of them were registered after this systematic review started (including Preliminary searches, Piloting of the study selection process, Formal screening of search results against eligibility criteria, Data extraction, Risk of bias (quality) assessment and Data analysis) Taiwan is part of China. Please correct it in both tables and paragraphs. In part 'comparator' of the 5th one of Table 2, it is not clear that who is the comparator. Please add more information about the reason for records exclusion in figure 1. Synthesis of results: Larger teaching interventions, line 5, it is unclear that who is compared to didactic teaching alone and what is the difference between its intervention and didactic teaching. DISCUSSION: Please discuss more the value of this article. Impact of Teaching Interventions, line 16, in the sentence that 'Studies suggest there are many misconceptions by lay and healthcare professionals of what palliative care is/hospices are…', what is the studies? Page 24 of 34, line 3, the logic that 'Both small amounts of specific teaching and larger scale interventions improved knowledge' can support ' institutions should investigate integrating some level of teaching palliative care, even if small, as these can prove beneficial to the knowledge base for students.' is doubtful. Page 24 of 34, line 8, the detailed meaning of 'effective instruction' is not clear. Hope you these comments are useful to you.

VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:	
Reviewer: 1	

Firstly, many thanks for conducting such a valuable review. This is such an important area and will be a great addition to the steadily increasing literature that is supporting palliative care education, and shaping future curricula. I have made several points below which are purely for your consideration, and are based on my own experience of publishing a systematic review. Some are issues relating to wording and sentence structure, others are seeking clarification and greated depth of information regarding your choice and interpretation of tools you have used, and suggestions on how your search could be a little more inclusive (and therefore systematic). I have numbered my comments in the hope it will make things easier to follow. I hope you find them helpful.	Thank you. We appreciate your very helpful review and support for this article and the important points you raise for our consideration
Format and editing:	
Abstract: Very first sentence. When you state 'most experienced doctors' do you mean those doctors with the most experience, or do you mean the majority of experienced doctors? If the former – then what about those doctors with lesser experience. If you mean the latter – then why not all experienced doctors? I would argue the point that palliative care is everyone's business irrespective of level of seniority and experience. You do go on later to state (page 4) 'medical students and doctors require the appropriate knowledge, skills and attitudes to care for patients who have advance and incurable illness', followed by 'The ability to care for, and communicate appropriately with all these patients and their families is an essential skill for all doctors. 1. I would consider amending the first sentence of your abstract for continuity and to reflect/support your later quotes.	Thank you. We agree, this was confusing. We have amended the first sentence of the abstract to: "all clinical doctors". This keeps the abstract more in line with the main text.
2. Page 3 : Strengths and limitations: 'This was a rigorously conducted systematic review including "grey" literature and evaluating the quality of the individual included studies'.	We agree. We have amended to: "This was a rigorously conducted systematic review, including "grey" literature, which evaluated the quality of included studies."

Sentence does not read well – consider restructuring.	
3. Background: (Page 4) '40 patients who will die'. I wonder if this would better read as '40 actively dying patients'. I am nit-picking here, but as all patients/people will die eventually, it may just read better the other way around.	We have amended to: "Approximately 40 patients who will die under their care. "
Methods:	
Search terms used included: 'palliative care', 'medical student', 'teaching' What about the use of the search term 'medical undergraduates' (or med* adj3 undergrad*)? I would imagine there could be articles which may be entitled 'How are medical undergraduates being taught palliative care?, or 'How effective is palliative care teaching in undergraduate medicine?' or 'Does undergraduate medical curricula effectively teach palliative care?' Also what about the use of the search term 'education'?, again an article could read 'The effectiveness of palliative care teaching within medical education' These may well have been included within your MeSH headings – but may be worth just checking and clarifying (or making this more explicit) 4. There is the potential for some articles to have been missed by not including the above search terms and I wonder if you would consider re-running your search to include these items. I suspect it will not add many (if any) articles to your review, but would definitely prove to be a more robust and systematic search strategy.	Thank you for picking up this potential gap in our search strategy. We have re-run the searches in all databases to include undergraduates and education, and where possible use the MeSH term Education, Medical, Undergraduate. We have cross-referenced with the original search and re-screened new articles, in duplicate, both at abstract and at full text stage. As you suspected, this did not lead to any extra articles for inclusion, but nevertheless has significantly strengthened our search strategy. We have updated the PRISMA diagram accordingly.
Eligibility:	

5. Could the authors clarify (or perhaps give an example of) what objective measure they may hoped to have seen from qualitative studies, given the statement that studies with qualitative outcomes were excluded. I was unclear about what was being looked for. If qualitative outcomes were to be excluded, why were they included within the search criteria?	Thank you for asking us to clarify. To make it clearer we have added: "e.g. a test score" We agree and have removed qualitative.
6. Length of follow up: No restrictions. Did you find this potentially lead to bias in some studies as a longer time period between intervention and follow up may affect the effectiveness of the results due to loss of knowledge / acquisition of new knowledge from other sources in the meantime. You may have mentioned this – but I couldn't see where.	We did not want to exclude on Length of follow up and 3 studies followed up over a long time period. We discuss this under Impact of Teaching Interventions and limitations
Quality assessment: (p7) and Quality assessment (p8) 7. Needs greater description of what the MMAT is for those who haven't heard of it What is it scored out of, how is this interpreted? The same applies to the Cochrane risk of bias tool. What is it? How does it score? How should readers interpret it? Why have you chosen these two tools specifically? Given MMAT gives a score between 0-100%, this is not recorded on the data extraction tables. You have stated – passed all points. What does this mean? Please consider documenting the actual percentage scores for those studies where the MMAT was applied. Did they all really achieve 100%? What was the level of agreement between reviewers of study quality? It would be good to see a measure of inter-rater variability, and a calculation of a Cohen's kappa score as this will add to the rigour of this review.	Thank you for raising this. We have added to methods: The MMAT is a critical appraisal tool developed to evaluate studies using both qualitative and quantitative data.15 MMAT was used in line with its original purpose, to appraise mixed methods research and to evaluate non-randomised quantitative research. Two screening questions are asked, before progression to more detailed analysis: 1. Are there clear research questions?; and 2. Do the collected data allow to address the research questions? In this review, the answer to both of these questions had to be 'yes' for a study to qualify for inclusion. Evaluation using MMAT subsequently focuses most heavily on appraising methodology, assessing five core criteria for each study type. These core criteria can be reviewed in detail, with additional usage guidance, using the 2018 iteration of the MMAT tool.15 To aid interpretation of what was meant by the core quality criteria, the research team referred to this expanded guidance. A summary of the core criteria for mixed methods research and nonrandomised quantitative research, the ways in which the MMAT was used in this work, are listed in table 2. We have also added Table 2: Summary of MMAT core quality criteria for mixed-methods and non-randomised quantitative research.

	The Cochrane risk of bias tool was used to appraise any randomised trial studies; as it is the gold-standard for such evaluation.14 ,The Cochrane risk of bias tool which has more stringent appraisal criteria, focusing on evaluating the presence of several types of bias: selection bias; performance bias; detection bias; attrition bias; reporting bias; and other bias. The plausible bias within studies deemed 'low risk' are unlikely to seriously alter results and can therefore be accepted. Studies at medium risk of bias imply 'some confidence that the results represent true… effect'. Despite medium risk, the issues with these studies are 'not sufficient to invalidate results' and, therefore, these studies can be included in our review unproblematically.16 Studies rated as high risk of bias should be considered sceptically. Additionally, we have further reviewed the mixed methods studies each and double checked our evidence based on MMAT 2018. In previous, earlier versions, the MMAT has been scored using a percentage value based upon the number of core quality criteria met by the study. However, in the most recent version of the MMAT, the 2018 iteration, scoring is explicitly discouraged, with qualitative description instead recommended. This work has payed heed to the five core criteria, explaining whether or not all criteria have been met. It was agreed that in any cases where all criteria were not met, qualitative explanation would be provided describing the shortcomings of the study We have added to methods: Disagreement was resolved by consensus and/or with a third reviewer (either A.D. or M.B.).
Results:	
8. In the interests of transparency and consistency, I wonder if it is worth stating which	In methods we say: Titles/abstracts and full-text papers were independently screened against

author(s) undertook which parts of the screening process, in the same way you did for data extraction. Who screened the titles, the full text?	pre-defined eligibility criteria (table 1) by two reviewers (J.B. and either A.D./M.B.). Disagreement at all stages was resolved by consensus and/or with a third reviewer (either J.B., A.D./M.B.). We hope this clarifies.
9. You mention a range of participant numbers between studies. It may be worth considering stating what the publication year range is as well in this paragraph.	Good point. We have added: Publication dates were between 2002 and 2018
10. Page 19. Different assessment methods. OSCE used for first time in the main text. Needs writing in full. Even though it is a commonly used abbreviation, so is MCQ (which you have written in full).	Thank you. We have defined the acronym OSCE.
Synthesis of results:	
11. I wonder if authors would consider including within the body of the text some of the statistical significance of the results, especially when referring to studies that have demonstrated overall improvement in knowledge scores. P values may be helpful for readers to put things into context. At the moment the reader has to keep going back to the table and working out which of the studies demonstrated statistical significance, no statistical significance, or simply did not apply stats to their results. This wouldn't take long to do and would allow the context of the results section to be more readily apparent, with an easier flow.	This is a good point and we tried to do this, but as in the results we are pooling studies for reporting, eg "Six of the seven included studies showed improved knowledge assessment outcomes" adding for each would be unclear. We have added statistically significant improvements in knowledge scores to each of the sections, to clarify that these studies did indeed reach statistical significance.
12. Again, when talking about the interactive eLearning course reporting equivalence in increased knowledge scores when compared with small-group teaching…what was the level of statistical significance to support this?	Although we can add in for individual studies, this would then seem unbalanced. This is covered under the statistical significance as per above comment.
Outcomes and constructive alignment considerations:	
I felt this was very well written, with very clear examples of what is meant by constructive alignment. Just a few wording/syntax issues.	Thank you

13. Page 21 line 49. Advance care planning, not ‘advanced’.	Thank you – we have amended
14. Page 22 line 13. Should it read: ‘it is not possible to know whether constructively aligned learning outcomes, teaching, and assessment are important to effective palliative care teaching’?	Thank you – we have amended
15. Page 22 line 51. Should this read ‘dispel these’ rather than ‘dismiss this’?	Thank you – we have amended
16. Page 23, line 3. I generally avoid beginning sentences with ‘And’ as it does not read well. Please consider rewording this.	Thank you – we have amended
17. Page 23, line 31. I wonder if this should read ‘demonstrated improved knowledge amongst these students’	Thank you – we have amended
Reviewer: 2	
I have to say this is a well-written paper for discussing palliative care teaching for medical students. The authors tried a lot to find the relationship between education and its potential influence on behaviours while the final results are not up to their expectations. This team followed a strict guideline and did the review systematically with the help of a clinical librarian. After finishing reading this paper, I still have the following questions:	Thank you
I checked the registration on PROSPERO(CRD42018115257) and noticed that there were two different versions. Furthermore, both of them were registered after this systematic review started (including Preliminary searches, Piloting of the study selection process, Formal screening of search results against eligibility criteria, Data extraction, Risk of bias (quality) assessment and Data analysis)	We can only see 1 version, but indeed this version has been updated as the review has progressed. It was originally registered in 2018 but updates have occurred up until 7.5.2019 as searches, screening, extraction and analysis were completed. This is in the revision note.
Taiwan is part of China. Please correct it in both tables and paragraphs.	We agree and have amended to china in both tables and paragraphs. However, as the authors of the papers (e.g. Tsai) have specifically mentioned both China and Taiwan, we think it would be disingenuous to not honour this.

In part 'comparator' of the 5th one of Table 2, it is not clear that who is the comparator.	It was 26 Small group sessions on palliative/end of life care. (vs the eLearning)
Please add more information about the reason for records exclusion in figure 1.	We have included the reasons for exclusion by full text based on the main areas of eligibility: We currently have: population, intervention and outcome
Synthesis of results: Larger teaching interventions, line 5, it is unclear that who is compared to didactic teaching alone and what is the difference between its intervention and didactic teaching.	Thank you for pointing this out. We have added: mandatory participation in a clinical palliative care module compared to didactic teaching alone
DISCUSSION: Please discuss more the value of this article.	We have added: All types of teaching intervention (small- and large-scale teaching, clinical and eLearning) improved knowledge scores for medical students. No method appeared to be superior in improving knowledge. Few studies explored knowledge retention, skills or attitudes. No studies explored the impact of teaching on clinical care for patients.
Impact of Teaching Interventions, line 16, in the sentence that 'Studies suggest there are many misconceptions by lay and healthcare professionals of what palliative care is/hospices are…', what is the studies?	We have added references to misconceptions: Smith T, Temin S, Alesi E, Abernethy A, Balboni T, Basch E, et al. American Society of Clinical Oncology Provisional Clinical Opinion: The Integration of Palliative Care Into Standard Oncology Care. *Journal of Clinical Oncology* (JCO). 2012;30(8):880-7. Hui D, Bruera E. Integrating palliative care into the trajectory of cancer care. *Nature Reviews Clinical Oncology*. 2016;13(3):159-71.
Page 24 of 34, line 3, the logic that 'Both small amounts of specific teaching and larger scale interventions improved knowledge' can support ' institutions should investigate integrating some level of teaching palliative care, even if small, as these can prove beneficial to the knowledge base for students.' is doubtful.	We can see how the logic of this recommendation might seem flawed. However, this suggestion is grounded in the results of our review. As both small-scale and large-scale palliative care teaching interventions were both found to improve student knowledge, it can be inferred that an intervention of any size may be useful in improving palliative care knowledge. As such, we suggest that small-scale interventions, which may be more readily and practically implemented by medical schools, could be of use. We believe this to be an important finding of this work, that medical educators reviewing

	already ‘full’ curricula could find of use, when considering provision of palliative care teaching. We have also made recommendations for future research to evaluate small- and large-scale palliative care teaching interventions alongside one another in a more easily comparable way, to assist in evaluation as to benefits and disadvantages to each approach. Given all this, we have not altered the meaning of this sentence, as we believe it a valid finding, grounded in evidence, that we build upon in suggestions for future research.
Page 24 of 34, line 8, the detailed meaning of 'effective instruction' is not clear.	We have changed 'effective instruction' to ‘teaching’.

VERSION 2 – REVIEW

REVIEWER	Dr Geoffrey Wells Brighton and Sussex Medical School, Brighton, England. UK
REVIEW RETURNED	20-Apr-2020

GENERAL COMMENTS	I think the revision has been done to a high standard with the vast majority of the issues addressed where possible. I think this will add much to existing literature, and is a very interesting and important read. Just a couple of very minor points the authors may want to change (but not enough to be deemed a 'minor revision'): Point 3 - You have made changes to this as previously highlighted, however I still don't think it reads well. When I read the newly structured sentence, to me it now sounds like the patients die because they are under the care of the FY1. Am sure this is not the message the authors want to portray. I still think it would read better if simplified to just '40 dying patients'. I will leave this to the discretion of the authors. Three very minor grammatical suggestions: Page 10 Patient and Public Involvement. I think a word is missing here. May I suggest: 'No patients were involved in this systematic review' or even 'Patients and public were not involved in this systematic review' Page 8 and page 10. MMAT. You first mention MMAT on page 8. It should spell out the name in full 'The mixed methods appraisal tool', rather than 'The mixed methods tool'. After this you can simply refer to MMAT. (see page 10 where you have written it in full again under Quality appraisal, this can just be the abbreviation I would think) Thankyou for making all the other suggested changes and clarifying points. Really good work - writing a SR is not easy, and you have done this well in my opinion.

VERSION 2 – AUTHOR RESPONSE

Thank you for the opportunity to respond to the helpful and constructive comments of the reviewer of this paper. Addressing their concerns has strengthened the paper.

We thank the reviewer for their kind words and final review of the manuscript and useful suggestions. We have taken all of these into account and have modified the manuscript based on these; it now reads better.