Author’s response to reviews

Title: Vertically integrated medical education and the readiness for practice of graduates

Authors:

Marjo Wijnen-Meijer (m.wijnen-meijer@lumc.nl)
Olle ten Cate (T.J.tenCate@umcutrecht.nl)
Marieke van der Schaaf (M.F.vanderSchaaf@uu.nl)
Chantalle Burgers (c.h.m.burgers@gmail.com)
Jan Borleffs (j.c.c.borleffs@med.umcg.nl)
Sigrid Harendza (harendza@uke.de)

Version: 2 Date: 16 September 2015

Author's response to reviews:

Dear Mr. Ulep,

Thank you for giving us the opportunity to revise our manuscript “Vertically integrated medical education and the readiness for practice of graduates”. Below please find a point-by-point reply to all editor’s and reviewers’ comments and suggestions. Changes in the manuscript are marked in yellow. We hope that our manuscript will now be suitable for publication in BMC Medical Education. Thank you very much for your reconsideration.

With kind regards

Sigrid Harendza

Editor’s comments:

Ethics Statement

Please place the section Ethics within the Methods section.

Reply: We have placed the ethics statement within the methods section.

Consent statement:

Please state in the Methods section whether written informed consent for participation in the study was obtained from participants or, where participants are children, a parent or guardian.

Reply: Written informed consent for participation was obtained from all participants. We added this in the methods section.

Acknowledgement

Please detail funding in the Acknowledgements section and remove the section Funding Sources.
Reviewer 1: Simon Watmough
Reviewer's report:
I do think the authors have tried an interesting study here in a worthwhile area – it is essential for all of us to know whether one curriculum or another could better prepare graduates to work as doctors. However, I do have a number of doubts about the paper which I have listed below which I would file under major compulsory revisions.

Reply: Thank you for your interest in our study. We revised our manuscript according to your suggestions below.

Abstracts
The results and conclusion should be re written, they don't really give the reader a clear summary of any take home messages from the paper.

Reply: We rewrote the results and conclusion section of the abstract to give the reader a clear summary and take home message.

Introduction
Sets the scene for the paper fairly well, but does have some areas for improvement/clarification. The paper would benefit from saying why the specific research questions i.e. unfamiliar clinical situations was chosen.

Reply: We specified the choice of the research question. Unfamiliar clinical situations were chosen because they most closely resemble the work situation of a young resident on a ward in a hospital.

There is an error, though saying reference 9 is a questionnaire based study.

Reply: We changes this sentence into: Three of these studies [1, 8, 10] are questionnaire-based analyses of student or faculty perceptions. The fourth study [9] is a focus group study.

It is not clear at all what the final paragraph in the introduction is trying to say and what its purpose is.

Reply: We rewrote the final paragraph of the introduction.

Methods
Much more is needed in the method about what the readiness tool is and its development and about how the standardized patients were chosen and what their conditions etc were etc? How were the distracting tasks arranged/decided upon and what exactly were they? Who are the supervisors exactly? Were they all debriefed about this/taught how to be involved in this research? Participants – why those students? Exactly how were they recruited, what percentage of the
cohorts were these numbers?

Reply: Because of the length of the manuscript, it is unfortunately not possible to include all this information in the manuscript. Therefore, we added in the methods section a reference to a separate earlier validation study of this assessment where detailed information is provided about the selection of participants and instruments.

Results

Some of the results i.e. looking at the different scoring groups from the two different countries should have been in the methods section in the relevant part about the statistical tests. The results though are clearly presented overall.

Reply: We moved this information to the methods section.

Discussion

Again, not all of those papers referenced are questionnaire studies – and there are now more recent papers published which look at reasons for preparedness for practice and these should be included in the discussion.

Reply: We included and discussed more recent references.

The number of students is very small and this isn’t adequately discussed in the limitations, neither is the voluntary nature of the project. Why did those who signed up take part? I know it is discussed in the limitations the fact the students are from different countries, tied in with the small numbers involved does concern me.

Reply: We added the fact that the number of students was small. The issue regarding the voluntary nature is included in the final sentence of the paragraph about the limitations of the study.

Also, the take home message in the discussion for other medical educators doesn’t to me seem to be very clear. It seems to be more about the methods used than the results - it seems to say that this doesn’t add to the literature about VI curricula and preparedness for practice but other studies do.

Reply: We stated the take home message of our study more precisely.

Conclusion

Doesn’t really give a “take home message” – it suggest the tools works but not that the study has itself has given us any firm conclusions. The aim of the paper isn’t about the tool but about the results of the research aren’t they? Overall, there is a lot of use of the words “we” and “our” etc, these would be better written in the third person in a scientific article. Needs some language corrections before being published.

Reply: We added a take home message and changed writing mostly to the third person. Language corrections were also made.
Reviewer 2: Matthias Angstwurm

Reviewer's report:

The authors have addressed an important question regarding the integration of basic science and clinical patient care in education to become medical professionals. According to the different approaches used in different medical schools it is very difficult to answer the question about different outcomes due to different educational strategies. The research question is a clear question: Do students of different curricula show differences in readiness for clinical practice? Did they have different competencies to cope with unfamiliar clinical situations? To compare the knowledge of passing students the same international exam was used: USMLE Step 1 was more successfully passed in integrated curricula compared to non-integrated one whereas Step2 clinical skills and clinical knowledge were similar. To overcome the problem they developed a specific assessment instrument called assessment was called “Utrecht Hamburg Trainee Responsibility for Unfamiliar Situation Test” (UHTRUST). This well balanced assessment test typical situations in clinical settings on a ward. The raters were carefully scored and their results were comparable in standardized settings with video taped scenarios. Patients were standardized. The results are therefore as standardized as possible and the results can be compared between groups.

Reply: Thank you for your appreciation of our study design.

Comments:

Major revision

- The number of candidates was 60. It is not clear whether a previous power calculation was done to find significant differences.

Reply: No power calculation was done in advance, but in a comparable study [28] the power calculation indicated that 21 candidates per group should be sufficient. To check for possible differences in the two populations, the NEO-FFI and the knowledge test were used.

- In addition students from two Dutch universities took place Utrecht (n=23) and Groningen (n=7) as well as students from one university in Hamburg (n=29). Altogether one student is missing.

Reply: We corrected the number of German candidates in the manuscript.

- Why are two different Dutch universities choose? Are the curricula completely identical?

Reply: The curricula in Utrecht and Groningen are identical and both vertically integrated.

- In addition please give more details about the characteristics of candidates. It is stated that candidates had nearly graduated from medical schools. What means nearly?
Reply: The Dutch students were two weeks before graduation at the moment of the assessment, and the German students three months. This difference is based on the fact, that we did not want the assessment to be scheduled too close to the final knowledge exam in Germany, to prevent possible bias. We specified this in the revised manuscript.

- It is stated that candidates participated voluntarily. The total assessment time is more than 6 hours. What was the impetus to participate a whole day? Did candidates get feedback, money for participation etc.? Do the authors have arguments to state that selected voluntary candidates are comparable to non participating students?

Reply: The candidates got feedback about their scores. They did not receive money or other rewards. The issue regarding the voluntary nature is included in the final sentence of the paragraph about the limitations of the study.

- which funding was used? Did the assessors e.g. medical physicians, nurses as well as standardized patients did receive payment in similar amounts in all three universities?

Reply: Assessors and other personnel were only from two universities, Utrecht and Hamburg. In both universities, there was no specific funding. Assessors, physicians and nurses participated during their normal work hours. Their specific clinical duties were covered by colleagues. Standardized patients were paid according to their usual rates like for other tasks they perform in the Hamburg or Utrecht curriculum.

Some differences were found between vertically integrated and not integrated curricula:

- Table 1: Within the facets of competence by physicians active professional development. What is meant by this and what was the difference? It is surprising that standard deviation was greater in smaller means. Is this correct? Does this mean that there were two different populations?

Reply: The complete description of the FOC “active professional development” is included in the discussion section. We double checked the calculations and they are correctly presented.

- Table 2: facets of competence by nurses the categories "Knowing and maintaining own personal bounds and possibilities" and "Teamwork and collegiality" were different. Are these differences caused by different nursing practice and cultures in Holland and Germany? The nurses as assessors were not compared between countries.

Reply: As we discussed in the discussion section, we cannot completely certain that the differences we found are not caused by cultural differences.

- Table 3: EPAs "breaking bad news" as well as "solving a management problem" were different. Especially breaking bad news is a skill needed to be learning in
clinical practice and in seminars. Is the curriculum in both countries / in all universities similar? Again the SD was smaller in German candidates despite higher scores. This means that a stargazed course in Hamburg might have been visited which is lacking in Dutch students. Please comment. The second EPA is vice versa. Therefore a specific teaching unit might be present in Dutch curricula but lacking in Hamburg.

Reply: We did make a detailed comparison of the content of the curricula of the medical schools involved and we did not find major differences regarding these subjects.

- in limitations: The authors did not find huge differences between the curricula. Perhaps the differences in curricula are to small? How often and how many hours were Dutch students exposed to patients? Is the number of hours spent in teaching sessions similar between countries? How often are basic science teaching sessions during the clinical part of curricula?

Reply: In addition to the former comment: we are not completely sure that the exact amount of hours are the same. We are convinced, though, that it is more important, how and when things are learnt (related to the different kind of curricula) than the exact amount of hours for certain subjects.

Minor Revisions:
In Conclusion:
One conclusion might even be that the new assessment instrument is not able to detect differences in clinical skills at the end of studies. One might assume that early integration of patients might inspire students in clinical skills but not in better learning of basics. Other way round repeated basic science during clinical education might improve understanding and help to find pathophysiologically related solutions in problems. Both was not tested by the assessment instrument.

Reply: In this assessment we tested the application of these kinds of knowledge, for instance in the choice for specific blood tests (e.g. without knowledge about pathophysiology, ordering bloodtests which are specifically needed, will not be possible). In another study (Harendza et al., 2013), the differences in ordering laboratory and radiology tests between the two groups are described.

Reviewer 3: Cheri Bethune
Reviewer’s report:
This study utilizes EPA (Entrustable Professional Activities) in a way that I am not familiar. In my own (perhaps limited but not insignificant) knowledge and experience of EPAs as a assessment tool has been for the programmatic assessment of trainees who are witnessed and assessed in their competence in achieving that particular array of competencies articulated in the EPA. I have no knowledge or experience in using EPAs as they have in this study to ask assessors to extrapolate entrustment to non witnessed professional activities. As this struck me as a unique utilization of EPAs I conferred with two experienced
educators who are quite knowledgeable about EPAs in competency based assessment. Neither of them were familiar with this way of using EPAs.

Reply: We were a bit surprised about this comment. Our method section states: “The second questionnaire consisted of nine so called “Entrustable Professional Activities” (EPAs), tasks that are suitable to entrust to a trainee once sufficient capability is attained for unsupervised practice [23] (see Table 3). The physicians were asked to indicate on a 5-point scale how much supervision they think that the candidate would need for these EPAs, which were different from the actually observed activities (1 = he/she is not able to do this; 2 = he/she is able to do this under direct supervision; 3 = he/she is able to do this if supervision is available; 4 = he/she is able to do this independently; 5 = he/she is able to supervise others in performing this activity) [24].” This is one common way to conceptualize EPAs and entrustment decisions. Entrustment decisions are based on a variety of observations that enable inferences about required level of supervision for future executions of these activities. This study did just that. We have added a sentence in the results section for clarity and one to the discussion section including two recent references.