

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Supervisors' perspective on medical thesis projects and dropout rates: survey among thesis supervisors at a large German university hospital
<b>AUTHORS</b>	Can, Elif; Richter, Felicitas; Valchanova, Ralitsa; Dewey, Marc

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Daniel Bauer Universität Bern Institut für Medizinische Lehre (IML) Abteilung Unterricht und Medien (AUM) Schweiz/Switzerland
<b>REVIEW RETURNED</b>	28-Jun-2016

<b>GENERAL COMMENTS</b>	<p>Thank you very much. The supervisors' perspectives were certainly something we knew very little about so far in the context of German medical dissertations.</p> <p>It seems that the present study, as well as that from Jüttemann and colleagues that the authors reference for comparison, are both somewhat relying on self-assessments. Unfortunately, there is quite a substantial body of literature (e.g. Kevin Eva, 2004 in Advances in Health Sciences Education) that finds the competency of self-assessment to be quite low, even in professionals, and even in health professionals. This might limit the interpretations: i.e., feeling better prepared hardly predicts performance in your thesis project and the notion of giving adequate support in stats hardly predicts actually doing it. And that is not even comparing it to the students' view who might be, individually and as a group, thinking on a completely different scale of "adequateness". Depending on if/how the authors anchored the items on their questionnaire, this is something they could choose to discuss.</p> <p>How did the authors determine the items on their questionnaire? By including some items or not, they might or might not have grasped all the reasons for dropouts. Also, they write the survey had 12 sections yet only report on four (motivation to supervise, supervisors'/candidates' reasons for dropouts and stats support). Were any of the suggestions offered that are relevant for this study?</p> <p>Some minor ideas – was there a reason you did not consider co-supervisors (Co-Betreuer) for this survey as well? It seems that the supervisors themselves (Betreuer, vulgo Doktorvater/Doktormutter) do only some of the actual supervision and rely sometimes quite a lot on their staff when it comes to doctoral candidates in Germany. Yet co-supervisors are themselves even less prepared to facilitate research and additionally lack the experience of senior researchers.</p>
-------------------------	--

	In the discussion's second paragraph you mention how a "vast majority" of supervisors see thesis supervision as a core responsibility or core duty. But the results state its only 37%, which hardly constitutes a "vast majority"?
--	---

<b>REVIEWER</b>	Dr. Marianne Giesler Center of Competency fo Evaluation in Medicine Students Deanary Medical Faculty University Freiburg Germany
<b>REVIEW RETURNED</b>	01-Jul-2016

<b>GENERAL COMMENTS</b>	<p>In what follows you will find some comments on the submitted article: Supervisors' perspectives on medical thesis projects and dropout rates: ...</p> <p>Page 2: The first sentence of the objectives section of the abstract refers a) to the identification of causes for failure of medical thesis projects and b) to the identification of causes for the high dropout rate in Germany.</p> <p>ad a) The dropout rate is not high, see explanations below ad b) Data were collected at the medical faculty Charité only, therefore no reliable conclusion can be drawn for Germany in general.</p> <p>The second sentence of the objectives section (the German research based system of earning a doctoral degree has come under attack) does specify an objective and should therefore be dropped.</p> <p>It is doubtful whether the mentioned survey with doctoral students was a longitudinal survey. Following the reference given by the authors it appears more likely that a cohort study has been conducted.</p> <p>The authors state, that 87% of the supervisors did not feel well prepared for thesis supervision. This result does not correspond to the result mentioned on page 7 und figure 2 (39%)</p> <p>Page 4: The authors report that the dropout rate from doctoral theses is high and also report, that almost 80 percent successfully completed their doctoral thesis in a short period of time (2013 until Jan/March 2014). Compared to the usual reported statistics [1,2,3] this dropout rate is rather low and vice versa the completion rate is a really high. Furthermore, it is advisable to reflect the use of the term dropout. It should be taken into account that termination of one project does not necessarily mean that candidates will not start another project.</p> <p>Page 5: The authors asked 3653 supervisors to participate but only 1069 (29%) answered. Even if one were to judge this response rate as acceptable the representativeness of the sample should be examined and discussed. The results of the survey are compared to the results of a survey with</p>
-------------------------	---

	<p>doctoral candidates. Therefore this comparison sample should be described in more detail.</p> <p>Page 7: Furthermore there is a lack of clear and accurate description of sample size used for analyses! The authors refer to a sample of 718 supervisors of thesis projects at the time of the survey (see figure 1). Why do they then refer to 948 in figure 2?</p> <p>Page 8: It is completely unclear how the absolute frequencies and percentages given for the number of supervisors on page 8 were determined. (n = 156 would be 25% of 624 supervisors and 344 supervisors would be 44% of 718 supervisors. Please explain. The hint on page 19 figure 1 (“decision questions”) is insufficient. The same criticism applies to the number of 107 supervisors given in table 1.</p> <p>Page 11 ff: As a consequence of the above listed shortcomings the discussion section should be revised regarding the respective points.</p> <p>Literature 1. <a href="http://www.his.de/pdf/22/medizinerreport_2012.pdf">http://www.his.de/pdf/22/medizinerreport_2012.pdf</a> 2. <a href="http://www.dzhw.eu/pdf/pub_fh/fh-201310.pdf?bcsi_scan_E872BC5C0E0115D2=0&amp;bcsi_scan_filename=fh-201310.pdf">http://www.dzhw.eu/pdf/pub_fh/fh-201310.pdf?bcsi_scan_E872BC5C0E0115D2=0&amp;bcsi_scan_filename=fh-201310.pdf</a> 3. <a href="http://www.egms.de/static/pdf/journals/zma/2016-33/zma001007.pdf">http://www.egms.de/static/pdf/journals/zma/2016-33/zma001007.pdf</a></p>
--	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Daniel Bauer

Institution and Country

Universität Bern  
Institut für Medizinische Lehre (IML)  
Abteilung Unterricht und Medien (AUM)

Schweiz/Switzerland

Please state any competing interests or state ‘None declared’:  
None declared

Please leave your comments for the authors below  
Thank you very much. The supervisors’ perspectives were certainly something we knew very little about so far in the context of German medical dissertations.

Comment: 1.1:  
It seems that the present study, as well as that from Jüttemann and colleagues that the authors reference for comparison, are both somewhat relying on self-assessments. Unfortunately, there is quite a substantial body of literature (e.g. Kevin Eva, 2004 in Advances in Health Sciences Education) that finds the competency of self-assessment to be quite low, even in professionals, and even in

health professionals. This might limit the interpretations: i.e., feeling better prepared hardly predicts performance in your thesis project and the notion of giving adequate support in stats hardly predicts actually doing it. And that is not even comparing it to the students' view who might be, individually and as a group, thinking on a completely different scale of "adequateness". Depending on if/how the authors anchored the items on their questionnaire, this is something they could choose to discuss.

Response:

Thank you for drawing our attention to this potential limitation. We agree that self-assessment may be quite unreliable and subjective. However, we also feel that it might not be that negative for our specific concerns and research interest (as we are not dealing with the assessment of performance in our survey). Rather, we were precisely interested in this subjective perspective on the situation of medical thesis supervision in Germany, and specifically dropout rates. We think that it is exactly how the situation of thesis supervision is experienced by supervisors and students and possible discrepancies can reveal underlying shortcomings, which will then help us find remedies. Knowing each others' perception of the situation and expectations can help overcome the gulf that appears to be at the basis of the problem. Our survey is about the perception of a shared situation which two parties engage in for different reasons. And this is where subjective perception (even if it misrepresents actual behavior) can indeed help us to solve the problem our study is about.

Our statistical results suggest that drop-out rates correlate with the self-assessments made by the respondents. A further investigation into this is beyond the scope of this work but would be interesting to help confirm whether the hypothesis holds up after further research and shed further light on the extent to which and the mechanisms whereby the factors mentioned determine the dropout rate. Furthermore, it makes intuitive sense for dropout rates to be correlated with the perception of competence (i.e., self-assessments), even if perception of competence is not that reliable an indicator of actual confidence.

Therefore, we have added a short explanation in the Methods section on p. 5 and in the Discussion on p.11. We included this aspect also in the conclusion (pages 2, 14). Kindly let us know if you feel that further, more extensive changes should be made regarding this concern.

Comment 1.2:

How did the authors determine the items on their questionnaire? By including some items or not, they might or might not have grasped all the reasons for dropouts. Also, they write the survey had 12 sections yet only report on four (motivation to supervise, supervisors'/candidates' reasons for dropouts and stats support). Were any of the suggestions offered that are relevant for this study?

Response:

Regarding the first part of your question, we included a sentence on the methods used to identify the questions (such as expert advice) including references on page 6. And we followed the recommendations for developing questionnaires for educational research (1,2).

We are aware that there are two selection processes that affect the results: the first selection takes place while setting up the questions and the second in selecting specific results for presentation. All questions were analysed and examined for this evaluation.

We selected those points supplementing the results of the survey among students, specifically all questions relating to the drop-out rate and items concerning supervisors' perception of the current situation. That is also the reason why we picked the four mentioned sections, and not all 12 sections. We excluded for example questions about time management and daily routine, because we think they would be more relevant in a national context and do not add to the recommendations made here to reduce drop-out. We briefly outline the methods used to select the questions on page 6 and 7.

1. Gehlbach H, Brinkworth ME. Measure twice, cut down error: A process for enhancing the validity of survey scales. *Rev Gen Psychol.* 2011;15:380–387.

2. Artino AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. *Med Teach.* 2014;36(6):463-474

Comment 1.3:

Some minor ideas – was there a reason you did not consider co-supervisors (Co-Betreuer) for this survey as well? It seems that the supervisors themselves (Betreuer, vulgo Doktorvater/Doktormutter) do only some of the actual supervision and rely sometimes quite a lot on their staff when it comes to doctoral candidates in Germany. Yet co-supervisors are themselves even less prepared to facilitate research and additionally lack the experience of senior researchers.

Response:

Sorry for not being more specific in this regard and causing this confusion. We did not make an explicit distinction between supervisors and co-supervisors. We use the term “supervisors” to refer to all those who actually supervised a thesis project during the period specified (whether in the official function or in the actual role). This was because we were interested in opinions based on the experience of supervision.

Note, however, that we only asked those employees with a university degree, so that we could assume a minimum level of research experience.

Comment 1.4:

In the discussion’s second paragraph you mention how a “vast majority” of supervisors see thesis supervision as a core responsibility or core duty. But the results state its only 37%, which hardly constitutes a “vast majority”?

Response:

Thank you for pointing out this mistake. We wanted to express that a vast majority of supervisors were highly motivated. We deleted the sentence.

Reviewer: 2

Reviewer Name: Dr. Marianne Giesler

Institution and Country

Center of Competency fo Evaluation in Medicine  
Students Deanary  
Medical Faculty  
University Freiburg  
Germany

Please state any competing interests or state ‘None declared’:

there is no conflict in interest

Please leave your comments for the authors below

In what follows you will find some comments on the submitted article: Supervisors’ perspectives on medical thesis projects and dropout rates: ...

Comment 2.1:

Page 2:

The first sentence of the objectives section of the abstract refers a) to the identification of causes for failure of medical thesis projects and b) to the identification of causes for the high dropout rate in Germany.

2.1. a) The dropout rate is not high, see explanations below.

2.1. b) Data were collected at the medical faculty Charité only, therefore no reliable conclusion can be

drawn for Germany in general.

Response:

Ad 2.1. a: Thank you. However, this is a matter of perspective. Compared to other fields of study, the completion-rate may be very high. It is also noteworthy that a large number of academics in the humanities (see your own references) never start a doctoral project in the first place.

The research done by doctoral candidates' accounts for a third of all publications of the Charité – Universitätsmedizin Berlin, and this is an important motivation to find out why almost a third of doctoral students cancel their projects.

We are especially interested in reducing the drop-out rate, which is unchanged at Charité despite the measures already taken to better prepare medical students for their thesis projects. We included a short explanation on p. 4. To clarify this, we replaced “rather high” by “constantly high” throughout the text (pages 2,4,14).

Ad 2.1.b: Our survey was not meant to be representative for Germany or Europe, we mentioned this aspect twice in the limitations section (p.3;14) :

Rather, we are aware of the special status of the Charité – Universitätsmedizin with its scientific staff of 1069 from all over the world and the highest publication rate in Germany. The more important it is to identify the reasons for the unchanged dropout rate, which is an indicator for process quality.

Comment 2.2:

The second sentence of the objectives section (the German research based system of earning a doctoral degree has come under attack) does specify an objective and should therefore be dropped.

Response:

Thank you, we dropped that sentence.

Comment 2.3:

It is doubtful whether the mentioned survey with doctoral students was a longitudinal survey. Following the reference given by the authors it appears more likely that a cohort study has been conducted.

Response:

In a longitudinal study, the same empirical study is performed at multiple time points. We administered the same questionnaire twice at two different timepoints. Therefore, we think it is adequate to refer to the student survey as a longitudinal study.

Comment 2.4:

The authors' state, that 87% of the supervisors did not feel well prepared for thesis supervision. This result does not correspond to the result mentioned on page 7 und figure 2 (39%)

Response:

13% of the respondents felt well prepared for their role. The remaining 87% did not, to variable degrees, i.e. the 87% comprise the following answer options: not at all (39%), moderately (32%), poorly (16%).

Comment 2.5:

Page 4:

The authors report that the dropout rate from doctoral theses is high and also report, that almost 80 percent successfully completed their doctoral thesis in a short period of time (2013 until Jan/March 2014). Compared to the usual reported statistics [1,2,3] this dropout rate is rather low and vice versa

the completion rate is a really high. Furthermore, it is advisable to reflect the use of the term dropout. It should be taken into account that termination of one project does not necessarily mean that candidates will not start another project.

Literature

1. [http://www.his.de/pdf/22/medizinerreport\\_2012.pdf](http://www.his.de/pdf/22/medizinerreport_2012.pdf)
2. [http://www.dzhw.eu/pdf/pub\\_fh/fh-201310.pdf?bcsi\\_scan\\_E872BC5C0E0115D2=0&bcsi\\_scan\\_filename=fh-201310.pdf](http://www.dzhw.eu/pdf/pub_fh/fh-201310.pdf?bcsi_scan_E872BC5C0E0115D2=0&bcsi_scan_filename=fh-201310.pdf)
3. <http://www.egms.de/static/pdf/journals/zma/2016-33/zma001007.pdf>

Response:

You are right - compared to other disciplines, the completion rate may be very high indeed. Here it may be noted that a large number among human science absolvents (see your own references) never start a doctoral project. Please also see our response to the question No. 2.1.a.

We also included a short explanation on p. 4 of the manuscript.

Regarding your other concern, we are aware that many students drop out and start again, which is an interesting aspect. However, in the present study this was not our focus. We wanted to examine why one would drop out in the first place and not how often students try again.

We thank you for drawing our attention to additional literature. We have included your article entitled "Importance and benefits of the doctoral thesis for medical graduates" (p.4).

Comment 2.6:

Page 5:

The authors asked 3653 supervisors to participate but only 1069 (29%) answered. Even if one were to judge this response rate as acceptable the representativeness of the sample should be examined and discussed.

The results of the survey are compared to the results of a survey with doctoral candidates. Therefore this comparison sample should be described in more detail.

Response:

Thank you very much for pointing out this confusion. We made an effort to improve our description and clarify this issue. The response rate of 29% is rather high compared to other online surveys without postal memory (references 27 and 28 on page 11). We have elaborated on this point. In the "Participants" subsection of Methods (p. 5), we give a brief description of the comparison sample.

Comment 2.7:

Page 7:

Furthermore there is a lack of clear and accurate description of sample size used for analyses! The authors refer to a sample of 718 supervisors of thesis projects at the time of the survey (see figure 1). Why do they then refer to 948 in figure 2?

Response:

For the Question how well supervisors felt prepared for their role, it is not important to avoid recall bias. Therefore, this question was not restricted to those who supervised a thesis project during the last 12 months and the number of respondents was larger (p. 6). We made a change to the legend of Fig. 2.

Because of the large number of questions, we have chosen a flexible method of investigation. We used decision questions that led to that not every participant was able to answer the same questions. This results in different sample sizes. This approach was chosen in order not to deter potential participants by the large total number of questions and explains the variation in sample sizes for different items on the questionnaire. We explained that in more detail on p. 6.

Comment 2.8:

Page 8:

It is completely unclear how the absolute frequencies and percentages given for the number of supervisors on page 8 were determined. (n = 156 would be 25% of 624 supervisors and 344 supervisors would be 44% of 718 supervisors. Please explain. The hint on page 19 figure 1 (“decision questions”) is insufficient. The same criticism applies to the number of 107 supervisors given in table 1.

Response:

We are sorry that we were not very clear regarding this point. We changed that part (p.8) to make the data more transparent: in our survey 25% of supervisors (156 of 624) stated that a total of 208 of medical thesis projects were discontinued, while 44% of supervisors (344 of 782) stated that a total of 598 projects had been completed successfully. Fifty-eight doctoral projects were prematurely ended by 11% of supervisors (47 of 436), while 28% of supervisors (109 of 386) reported a total of 150 thesis projects terminated by the candidates themselves. Because of the large number of questions, we have chosen a flexible method of investigation. We used decision questions that led to that not every participant was able to answer the same questions. This results in different sample sizes. This approach was chosen in order not to deter potential participants by the large total number of questions and explains the variation in sample sizes for different items on the questionnaire. We made changes in the text to be more clear about this, see Methods section and on page 8 as well as in Table 1.

Comment 2.9:

Page 11 ff:

As a consequence of the above listed shortcomings the discussion section should be revised regarding the respective points.

Response:

We followed your advice and made changes to the Discussion to the effect that we provide more details relevant for the focus of the present study.