Author's response to reviews

Title: Using a Formative Simulated Patient Exercise for Curriculum Evaluation

Authors:

Dr David J Solomon PhD (dsolomon@msu.edu)
Heather S Laird-Fick MD (Heather.LairdFick@ht.msu.edu)
Carole W Keefe PhD (ckeefe@msu.edu)
Margaret E Thompson MD (Peg.Thompson@ht.msu.edu)
Mary M Noel PhD (Mary.Noel@ht.msu.edu)

Version: 2 Date: 22 Mar 2004

PDF covering letter
Dear Dr. Gadd,

Thank you for providing feedback on our manuscript “MS: 3463773812802879 Using a Formative Simulated Patient Exercise for Curriculum Evaluation”. My co-authors and I have tried to address the comments and suggestions provided by the three reviewers. A description of how they were addressed is given below.

Reviewer 1

Discretionary Revisions (which the author can choose to ignore)

It would be helpful for the reader not familiar with participatory decision making to include a sentence or two that describes the uniqueness of this approach by physicians when communicating with patients.

[Response] The background section was expanded to provide more information concerning PDM.

In the Methods section, Intervention subsection, it would be helpful to refer back to the Keefe work to clarify for the reader that it is what the authors are referring to in their description of the intervention. For example, "As described by Keefe et al, students are trained to use a variation of a model...."

[Response] A more in depth description of the Braddock model was added to the “Intervention” section.

Reviewer 2

The question of whether training in PDM impacts behavior is important. One thing that needs to be clarified for the interpretation of this research to be clearer is whether the students in the IM rotation taking the SPE were told that the SP case would be rated based on PDM skills. Or stated another way, when the case was presented to the student, did they understand that PDM skills would be an important dimension on which they would be evaluated?

[Response] The students were not told the case would be rated on PDM skills. They were told the case involved teaching and working with a patient to make a treatment decision. This is stated in the last sentence of the Study Design Section.
As the Methods section in the abstract conveys, this study had 3 groups: 1.) No PDM (Cell 2); 2.) PDM 4wks ago (Cell 3 & 4.)  3.) PDM 12 weeks ago (Cell 5). Yet the analysis doesn’t seem to reflect this. Shouldn’t Cell 3 & 4 be one group (PDM 4 wks ago)? Also in Cell 1, was there an IM rotation? If so, could that be combined with Cell 2?

[Response] First, I want to apologize. In the final revision of Figure 1, I inadvertently switched the definitions for Cells 4 and 5. The students in Cell 5 had completed the PDM training during the second rotation, 4 weeks prior to taking the SPE. This has been corrected.

The reason students who completed the internal medicine rotation during the second and third rotation who received PDM training four weeks prior to the SPE where not combined into a single group was due to the possible (actually I believe quite likely) impact of different levels of training. This is also the reason students who completed the internal medicine the first rotation of the year (cell 1) and did not have PDM training were not combined with cell 2, students in the second rotation of the year who did not have PDM training for the purposes of the study.

There is significant overlap in what is learned across clinical rotations and it is generally understood that students’ history taking and physical examination skills improve markedly over their first few clinical rotations of the third year of medical school. We felt it would be quite likely an additional eight weeks of clinical training in another clerkship would impact on the students’ performance on the SPE whether or not this training involved specific training in PDM skills. This is the reason we limited comparisons to among students in the same internal medicine clerkship rotation with the same total amount of experience in clinical clerkships.

Figure 2 & 3 should have better labels on the horizontal (No PDM – 4 wk PDM – 12 wk PDM – or something?). Otherwise the reader has to go back and figure out what Cell 1, 2, 3 & 4 represent to make sense of this.

[Response] This is a good point however they are labeled in the legend of the figure which is included at the end of the text of the review manuscript. When published, the legends will be under the figures. In retrospect, as noted by the reviewer below, I wonder if these graphs are really that useful and would like to drop them from the paper if this is acceptable.

When assigning subjects by pre-existing groups, there is always a confounding of the treatment and everything else that defines the group (ie – common clerkship experiences). This has led methodologist to suggest group be the unit of analysis. That is not possible here, but confounding needs to be considered here. I can understand that subjects may have been somewhat randomly assigned to groups, and to the degree this is not strictly true it may not be a major worry, however the group they belong to has other common group experiences, and that is where confounding comes in.

[Response] I totally agree with the reviewer’s comments. However, conducting studies with true experimental designs and using classrooms or in this case rotations as the unit of analysis is often just is not practical in the real world. This is clearly a quasi-experimental design as noted in the manuscript and subject to
potential confounding. After discussing how our students are assigned to rotations with the administrator who handles the process at our institution, I doubt there is any significant bias in the assignment to rotations that would impact on the results however I can’t prove it. The experimental unit/unit of analysis issue noted by the reviewer in my view is probably more serious however there is no practical solution. As the reviewer notes, this is a chronic issue that plagues educational research studies.

The Figures 2 and 3 on my copy are a bit confused. For instance Cell 3 doesn’t have a mean bar and the top and bottom confidence interval is missing.

[Response] The box and whisker plots were generated by SPSS. I suspect the unusual configuration is due to the small sample size, for example, I believe the median bar is not clearly visible because it is at the very top of the box. As noted above, I think it would be best to drop these figures from the manuscript.

When I look at these results I see the right side of Table 2 being most important (why don’t Ns agree on the left and right side? Why is N squared on the left side?)

[Response] The numbers differ due to missing data. The PDM score is a subset of the total score and on one of the rating scales in cell 5, an item was not marked by the faculty rater. Since it was not one of the items that reflected PDM skills we were able to include the student in the analysis for the PDM score but not for the total score.

The superscripted “2” by the “N” in the title was a reference in the legend (explaining why the numbers disagreed). To make this clearer, I changed the references superscripts to symbols instead of numbers.

No PDM = 15.2 – sd = 1.6 - n = 17
4 wk = 15.9 – sd = 1.55 - n = 23
12 wk = 14.7 – sd = 1.8 - n = 7

Above appears the most important summary of this research. There is about a .5 SD gain, but then after 12 weeks it shows a .33 SD loss. Given the small n for 12 week I am wondering if this condition should be deleted from the study. It appears to me that this research could be best presented with one Table and one Graph with confidence intervals using just those 3 conditions for which you have adequate numbers.

[Response] As noted above I would be very uncomfortable about combining data for the analysis from students who had completed different numbers of clinical rotations despite the obvious benefits of an increased sample size.

I agree with the reviewer, it would be useful to include a graph of the scores with confidence intervals. However, given the small sample sizes, I don’t think it would be appropriate to calculate confidence intervals. There are nowhere near enough observations for the central limit theorem to apply and the distribution of observations within the groups does not appear to be normally distributed.
I like this investigation and think the methodology should offer others instruction on how they might do useful research on these questions, but I think it needs to be reworked a bit. It should present just information for which there is some level of confidence and discuss the theoretical implication.

[Response] Thank you. We decided to use nonparametric statistical significance tests due to the small sample size. Despite the small sample sizes, we found statistically significant differences among the students in the third rotation who completed the PDM training 4 versus 12 weeks earlier. Differences in the second rotation approached but did not reach statistical significance. The statistical tests that were used are appropriate even with the small sample size. The only issue is whether the assumption of independence is met. As noted by the reviewer, this is questionable due to the use of individual students as the unit of analysis when they are not truly the experimental unit. This however, would still be an issue even with a larger sample size using either parametric or nonparametric tests.

Reviewer 3

-------------------------------------------------------------------------------
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

I have a few concerns about this study that revolve mainly around its clarity for someone who may not be very familiar with Participatory Decision-Making (PDM) and other unique aspects of this study. In the background section the authors mention the prior research into the use of PDM but do not address the findings of these works. More information is needed here. In the intervention section and latter in the statistical analysis section Braddock’s model is mentioned. This appears to be a critical component of the study. Some elaboration on the specifics of this model would be very helpful to the reader.

[Response] Per the reviewer’s suggestion, we have expanded the discussion of the research on PDM in the background section of the manuscript. We have also expanded the discussion of Braddock’s model in the Methods Section.

-------------------------------------------------------------------------------
Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Figure 1 is not labeled.

[Response] The title for Figure 1 is at the end of the text in the main manuscript per instructions for authors. I assume it will be placed above or below the figure in the final manuscript.

-------------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)

The last sentence of the measures section states that "no formal training was given in how to rate the students." It would be reasonable for the authors to note there sense as to
whether this may or may not have an influence on the results. It is a potential limitation of the study.

I found both the study design and results sections a little hard to follow. The timing of the various clerkships done in each rotation along with when the intervention occurred make for a somewhat confusing situation. The explanation in the paper was a bit hard to follow. The second paragraph of his section was the most difficult for me. The figure helps clear things up some. At the same time, I don’t have a suggestion of how to present this information any more clearly.

The last three paragraphs of the results section present the statistical analyses. My interpretation of these is that the first of the three paragraphs is presenting the overall significance testing. Finding significance the authors then pursued post hoc comparisons and these are presented in the last two paragraphs of this section. If so, it might be reasonable to state this explicitly in the paper.

[Response] This is a complex study design and I understand that the description is difficulty to follow. This is why we included Figure 1. I struggled with how to present it more clearly and like the reviewer, I can’t think of a way to make it clearer.

The wording in the statistical analysis section was changed to specifically state post-hoc tests were performed.

I was also curious why a comparison between cells 3 and 5 was not presented. I recognize they were the smallest groups. At the same time the differences between scores is the largest.

[Response] As noted above, we chose not to make comparisons between the students who completed the internal medicine clerkship in the second and third rotations. It was felt any differences would be confounded by the additional clinical experience the students in the third rotation.

I hope this response has adequately addressed the concerns of the reviewers. Please contact me if you feel additional changes are warranted.

Best wishes,

David J Solomon, PhD
Associate Professor
College of Human Medicine
Michigan State University