Reviewer’s report
Title: Science of Health Care Delivery Milestones for Undergraduate Medical Education

Version: 0 Date: 30 Mar 2017

Reviewer: Stanley Hamstra

Reviewer's report:
First of all, I want to congratulate the authors for taking on a difficult task. The authors undertook to develop SHCD sub-competencies for use in assessing medical students. The authors should be commended for reviewing the Milestones language across 25 specialties. However, part of the challenge of doing work in this area is the sheer variability and complexity across specialties within the field of medicine. For example, while 25 specialties certainly represent the "core" specialties, there are another 90 or so sets of Milestones for sub-specialties that reflect the greater complexity of residency training in the field. So in this sense, the research described here has some limitation in terms of generalizability.

The problem with this, in terms of interpretation of these results, is that the iteration of the Milestones language that this team worked with probably represents too much unnecessary complexity. This has to do with the history of the process and the enormous "change management" project that was undertaken to initiate the Next Accreditation System. The initial vision for this is laid out nicely in Nasca et al. 2012 NEJM 366:11:1051-6.

The change management aspect refers to the decision to let each specialty (and sub-specialty), along with their Boards, develop and write their own set of sub-competencies and Milestones language to reflect the presence of the 6 "core" competencies in the context of their own specialty.

The ACGME recognizes that the current formulation of the Milestones likely represent a degree of variability between specialties and a complexity that could be greatly reduced without loss of the larger effort to move forward in the Next Accreditation System, i.e. to move GME along the path towards competency-based medical education (and thus, away from a pure time-based model).
All this to say that there is currently an effort by the ACGME and its specialty partners to undertake a massive revision of the Milestones language within and across specialties to both continue to reflect the variance that should exist between specialties and at the same time harmonize language where appropriate (which would significantly aid the current group in doing this type of work in the future).

1. So from that perspective, I believe this work may be even more influential if done after "Milestones 2.0" rolls out. In the meantime, I believe the work represented here should be published, pending the following itemized revisions:
   a. Response required: the authors should update their Introduction and Discussion material to identify that since Sept 2013, Milestones have been developed for the full complement of all ACGME accredited specialties - both residency programs and fellowships. And that in this sense, the findings here might be limited to only the so-called "core" specialties.

2. L6: The authors acknowledge in several places that these constructs and the particular language used to assess them will undergo periodic revision.
   a. Response required: Please describe briefly what recommendations/vision you have for ongoing revision.
   b. Also, please describe if/how they are generalizable across specialties? Have the authors checked with representatives from other specialties?

3. L16: This is really excellent work. Even in the changing landscape of Milestones language at the GME level and EPAs at the UME level, I believe SCHD content must be refined and integrated into GME and UME competency categories wherever possible and highlighted, given their importance in contributing to effective, conscientious, and high quality care.
   a. Response required: I congratulate the authors for taking on this initiative and only hope that they create some kind of "institutional memory" or other means of memorializing this work to ensure continuation as the landscape changes. In other words, even if not a perfect fit for GME Milestones and UME EPAs today, it provides a blueprint for ongoing revision and fine-tuning, which will need to be attended to in any case.
4. L18: The authors state "This set of milestones is applicable to all specialties…"
a. Response required: Based on what criteria? It would be helpful if this claim were made more concrete, perhaps by a specific reference.

5. L19: The authors state that this work "…may help facilitate students' transition from medical school to residency."
a. Response required: This is a great motivation and rationale for doing this work, but really, how will we know if you don't propose some way of evaluating the impact of this work on that question? I don't really expect the authors to develop a complementary program evaluation or research question directed at this question, but I would like to see them describe how one might go about this, perhaps as part of their concluding remarks for Future Directions in the Discussion.

6. L31: I'm sure the case is made completely for the need for this work. The entire rationale for the paper rests on a single reference [2]… from which the authors conclude that this work is needed.
a. Response required: So I would ask the authors… based on what criteria? And is this only the case for Internal Medicine (i.e. the focus of reference [2])? or is it broader? A reasonable counter-argument would be that these competencies are covered in the current formulation of ACGME Milestones. It would help if the authors provided a bit more specific rationale for doing this work, and - perhaps audaciously - layout the implications for this work in terms of improvements in training and ultimately, patient care.

7. L89: I think it's a mistake to use the negative language that was adopted by Internal Medicine. It distracts faculty raters who might otherwise be likely to rate a learner at "Level 1" if the language were more descriptive of positive instances of behavior. In other words, given the culture of assessment in the hands of otherwise inexperienced raters, this is a mechanism for encouraging grade inflation and/or range restriction in the resulting data.
a. Response required: It would help if the authors would either modify the negative language by removing all reference to "critical deficiencies", or make the case for why this is essential in
crafting Milestones language within or across specialties. I'm looking for a rationale for this beyond "this is what people are used to".

8. L98-100: Further to the previous point, here is some text that could be considered more fully to think about how to address the problem of negative language in the Milestones:
   a. Response required: Could the authors reframe this (L98) in positive language? The following point (L100) about "uses technical jargon" is a good example of using positive language to describe where the learner "is" at the current moment, i.e. instead of "what they must become", which may be, in principle at least, much harder for different raters to agree on.

9. L115: re: "help facilitate transition from UME to GME"
   a. Response required: How do you know this is an improvement over the existing status quo? Also, more generally... do we know this is a problem? Based on what? Perhaps a bit more discussion of cases in which medical students have had trouble adjusting and why.

10. L133-4: Age-old problem of fighting for space in a crowded curriculum.
    a. Response required: Could this be framed in more objective terms, i.e. evidence base for priority of some elements of the proposed curriculum over existing one? If not, this could either be interpreted by those outside of SHCD as conceit or perhaps political. I think it would help the authors to reframe this and flesh it out a little with independent evidence for the need.

11. L146: The authors are correct to point out a limitation of this study as being comprised of members from a single institution. If we were to consider Neurology as sub-specialty of Internal Medicine (as it is elsewhere in the world), we might argue that the authors also represent at most, two medical disciplines: IM and Peds. Surely, some of the hospital-based and procedural specialties were not present in the team.
    a. Response required: I think the authors should discuss the potential limitations of forming their approach from the perspective of a single specialty (or if they don't agree with my characterization of what constitutes a "medical discipline", then at least address the lack of representation from a procedural discipline).
12. Supplementary Material - Additional File 1:

a. Finally, I'm not sure the authors have done a complete mapping. While I did not go through and try to repeat their mapping exercise, I did find a few places where they missed the connections between their conceptualization and the existing Version 1.0 Milestones - e.g. ACGME IM_ICS1 clearly states "Engages patients in shared decision making in uncomplicated conversations" compared with their "Additional File" item #2. E.g. in the second table, under "Level 4", it appears that "IM ICS1" is missing. This language clearly shows up in the ACGME Milestones for IM ICS1.

b. The first table makes reference to "PS ICS", but I could not readily identify this theme in the Plastic Surgery Milestones for ICS. Perhaps this is miscoded? (There is similar language in PS PROF1.)

c. In the second table, under "Level 5", "Ethics" is not a category for ACGME competencies. Perhaps the authors are referring to PS PROF1?

i. Response required: Please correct the above deficiencies in coding and/or provide justification. Also, please assure the reviewer the whole set has been double-checked for accuracy.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.
Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.
Yes

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.
Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.
I am able to assess the statistics
Quality of written English
Please indicate the quality of language in the manuscript: Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?
4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?
5. Do you have any other financial competing interests?
6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

'I declare that I have no competing interests'

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal