Program Director Perceptions of Usefulness of the Accreditation Council for Graduate Medical Education Milestones System for Urology Resident Evaluation

Elisabeth M. Sebesta, Kimberly L. Cooper, and Gina M. Badalato

OBJECTIVE To assess the application and perceived usefulness of the Accreditation Council for Graduation Medical Education Milestones system for resident evaluation among urology program directors (PDs).

MATERIALS AND METHODS We conducted an online survey of 133 urology PDs. The survey addressed several domains: (1) demographic information, (2) logistics and implementation of the faculty Clinical Competency Committee meetings, and (3) perceived overall effectiveness and usefulness of the Milestones assessments.

RESULTS Eighty-eight responses were obtained (66% response rate). A total of 42/88 programs (48%) described the Milestones as very or somewhat unhelpful in resident evaluation, with a comparable proportion (44%) responding Milestones assessments never or almost never accurately distinguished between residents. Respondents felt higher scores on all domains of the Milestones were completely or somewhat uncorrelated to higher in-service examination scores (58%), with a smaller fraction (49%) deeming they were not predictive of board passage rates. Overall, 30% of respondents answered neutrally as to whether they felt the Milestones format has led to better resident formative feedback, and 35% were neutral as to the implications of this system toward promoting professional development.

CONCLUSION The Accreditation Council for Graduation Medical Education Milestones system for resident evaluation was initiated to create a uniform competency-based assessment system; however, a sizable proportion of urology PDs in our cohort did not find the Milestones system helpful or accurate in assessing residents or predicting future successes. Given the Milestones system is still in its infancy, the utility of this system within urology has yet to be fully assessed. UROLOGY 124: 28–32, 2019. © 2018 Elsevier Inc.

In 1999, the Accreditation Council for Graduate Medical Education (ACGME) first introduced 6 domains of clinical competency and by 2009 began the process of restructuring the accreditation system to be competency based.[1] The result was the Next Accreditation System which was introduced to be implemented by 2013, and relied on a competency-based Milestones system to streamline resident evaluation. The Milestones are organized into 6 domains of clinical competency including patient care, systems-based practice, practice-based learning and improvement, interpersonal and communication skills, professionalism, and medical knowledge (MK).

The Next Accreditation System was initially adopted by 7 of the 26 ACGME-accredited core specialties, with urology included as 1 of the 7 early adopters.[2] The urology Milestones include 34 sub-competencies across all 6 competencies, and was initially implemented in July 2013.[3] Each sub-competency is evaluated for each resident in a program twice per year by the residency Clinical Competency Committee (CCC) consisting of faculty members within the program. The resident obtains a score on a 1-5 level scale, with level 4 Milestones corresponding to the minimum performance expected by the time of urology residency graduation.

Although the Milestones system of evaluation is now widely used among specialties, the specific benefits and perceptions of usefulness to residency Program Directors (PDs) and resident education, as well as barriers to implementing this system have yet to be addressed in the urology literature. In this study, we aimed to assess the application and the perception of usefulness of the Milestones system among PDs in urology resident evaluation.

From the Department of Urology, NewYork-Presbyterian Hospital/Columbia University Medical Center, New York, NY

Address correspondence to: Gina M. Badalato, M.D., Department of Urology, Columbia University Medical Center, 161 Fort Washington Avenue, 11th Floor, New York, NY 10032. E-mails: gmb2107@cumc.columbia.edu; gmb2107@columbia.edu

Submitted: July 27, 2018, accepted (with revisions): October 11, 2018

https://doi.org/10.1016/j.urology.2018.10.042
0090-4295

MATERIALS AND METHODS

An anonymous 26-question online survey was sent to 133 urology residency PDs in 2017 and 2018. The survey addressed several domains: (1) demographic information, (2) logistics and implementation of the faculty CCC meetings, and (3) perceived overall effectiveness and usefulness of the Milestones (Appendix) assessments. All responses were pooled for analysis and descriptive statistics were performed. All statistical analysis was performed in Microsoft Excel.

RESULTS

A total of 88 responses were obtained for a total response rate of 66% of all 133 United States accredited urology programs. Respondents represented programs in all American Urological Association sections (Table 1). Of all respondents, 19% (N = 17) were from the Western section, 18% (N = 16) from Southeastern section, 15% (N = 13) from New York section, and 13% (N = 11) from New England section, South Central section, and North Central section each, therefore programs from all across the country were well-represented. The majority of respondents were from programs located in major urban settings (68%), with suburban and small or mediumsized community programs less represented. The number of full-time faculty members in programs was 3-33, mean 14 (median 13).

Characteristics of CCC

When asked about the CCC at their individual programs, in most cases PDs stated the CCC consist of mean 5.6 faculty members (median 5), ranging 3-11 faculty members (Table 1). The CCC was most commonly chaired by the PD at 47% of programs, or another key faculty member (28%) or associate program director (24%). Only 1 (1.1%) CCC was chaired by the department chair. When asked whether CCC members receive any training or faculty development prior to involvement in the CCC, 31% of PDs replied all CCC members are offered training at their institution and 29% replied the PD or CCC chair is offered training, however only 10% of PDs replied training is required at their institution. In contrast, however, 25% of PDs replied that no faculty member receives training at their program. When asked specifically about what training and faculty development included at their program, responses varied widely. The majority of respondents stated training included assessments of residents through direct observation in a clinical setting (65%) and education about the ACGME general competency domains (65%). Fewer respondents stated CCC members received training on giving feedback (52%), on competency-targeted evaluation of trainees (34%), or clinical skills training (25%).

Assessment Processes and CCC Meetings

When questioned regarding the logistics of CCC meetings, the majority of programs reported to meet biannually (83%) over a single session (85%) which takes less than 2 hours (38%) or 2-4 hours (46%) to complete. Most CCCs evaluate each resident individually (77%) and makes evaluation decisions by all CCC members reviewing resident assessment data (58%). It was less common for evaluations to be completed by 1 member who then presents recommendations to the CCC (16%) or by discussing each resident based only on member knowledge (21%). Overwhelmingly, respondents stated the CCC placed the highest value in the patient care competency (69%), followed by MK (14%) and professionalism (9%).

Milestones as an Evaluation Tool

In terms of the usefulness of the Milestones in resident evaluation and the ability to accurately predict future success, the responses of PDs in our survey varied widely (Fig. 1). Of all respondents, 47% reported feeling somewhat or very confident in the judgments being made by the CCC. However, when asked regarding the overall perception of helpfulness of the Milestones for resident evaluation, 48% stated they felt the Milestones were very or somewhat unhelpful, and 38% reported they felt neutral. Additionally, 44% stated they felt the Milestones never or almost never accurately distinguished between residents. Likewise, responses varied in how the Milestones evaluations correlated to other measures of resident success. Respondents felt higher scores in all Milestones domains were completely or somewhat uncorrelated to higher in-service examination scores (58%) and passing the board examinations for graduating residents (49%). Overall, 30% of respondents answered neutrally as to whether they felt the Milestones format of evaluation has led

Table 1. Demographic information on programs and CCCs of all survey respondents

		Total Respondents (N = 38)		
		No. (%)	Mean ± SD	Range
AUA section	Western section	17 (19.3)		
	South Central section	11 (12.5)		
	South Eastern section	16 (18.2)		
	North Central section	11 (12.5)		
	Mid-Atlantic section	5 (5.7)		
	New York section	13 (14.8)		
	New England section	11 (12.5)		
	North Eastern section	4 (4.5)		
Setting of institution	Major urban	60 (68.2)		
	Suburban	12 (13.6)		
	Small- or mediumsized community	15 (17.0)		
	Rural	1 (1.1)		
Number of full time faculty on staff			14.1 ± 7.1	3-33
Number of CCC members			5.6 ± 1.6	3-11
CCC chair	PD	41 (46.6)		
	Associate or assistant PD	21 (23.9)		
	Department chair	1 (1.1)		
	Other key faculty member	25 (28.4)		

AUA, American Urological Association; CCC, Clinical Competency Committee, PD, program directors.

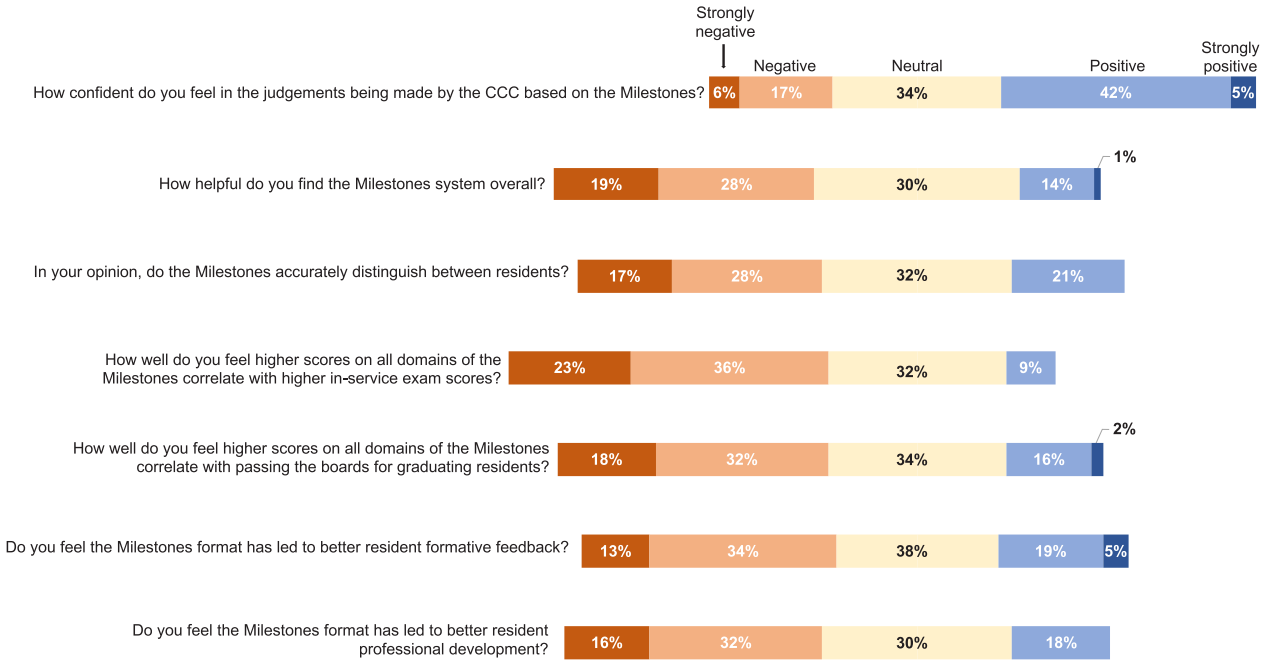

Figure 1. Responses to survey questions on a Likert scale regarding the Milestones format for resident evaluation. (Color version available online).

to better resident formative feedback; 35% were neutral as to the value toward improving professional development.

Implementing Educational System Changes Based on the Milestones

Despite these results and the fact that the Milestones system has been in place for 3 years in urology, 56% stated they have not made any changes to the resident evaluation process to adapt to the Milestones and 80% stated they have not made any changes to the resident curriculum based on the Milestones.

Feedback

Study participants were asked for any additional comments, either positive or negative, on the Milestones. Many comments reflected the fact that the Milestones are felt to be difficult to translate into practical clinical use and are in fact difficult for even the CCC members to understand.

> "We look at the milestones as a requirement to complete for the ACGME. This is not felt to correlate with much in our program, with the exception of discovering a resident that is truly struggling."

> "The majority of the milestones are cumbersome, repetitive, too numerous and, in some instances, somewhat difficult to understand."

> "My problem with the milestones at this point is that it is a lot of jargon dancing around some truth but is very hard to really get a sense of where the resident is at. . .."

> "Milestone is an artificially made up concept that is completely disconnected from the real world practice of medicine"

However, despite the drawbacks, many respondents also stated they feel the Milestones system is an overall improvement from prior methods of resident evaluation.

> "The Milestones are a fair evaluation tool as it forces us to make assessments on the different facets of a urology resident's skills, knowledge, and communications styles."

> "The feedback that I receive from faculty during the CCC meetings, as we complete the Milestones for each resident, has been dramatically more robust and useful for individual feedback. Forcing the faculty to convene twice per year to evaluate the residents has been very high yield for quality qualitative feedback."

> "The one advantage of the Milestones is that it provides very concrete examples and these examples often spur faculty to come up with other concrete examples of what residents do well or do poorly. . .Our residents do seem to like the examples."

COMMENT

Urology was 1 of the initial 7 specialties to adopt the Milestones system in 2013, and this is the first internal evaluation of this system from our specialty in the literature to date.[1] The ACGME Milestones system for resident evaluation was initiated to create a uniform competency-based, specialty-specific assessment of residents. In this survey-based study, feedback was garnered regarding the implementation and perception of usefulness of the Milestones system by urology residency PDs. Overall the survey provided many positive comments that the Milestones system has increased the overall quality of feedback obtained from faculty members due to its strict requirements. Nevertheless, a relatively sizable proportion of urology PDs in our cohort did not affirmatively find the Milestones system overall helpful in assessing residents or predicting future successes on in-service examination or board examinations.

The findings of this investigation are novel for the field of urology, although comparable studies pertaining to the Milestones have been performed in other surgical specialties. Specifically, a survey- and interview-based study from the contemporary neurosurgical literature asked program leadership and residents to reflect on their experience thus far using the Milestones system.[4] Their study, likewise, indicated the adjustment to competency-based evaluation is an ongoing process for many programs. Similar to our study, this survey revealed there is a wide range of training provided to CCC members on resident evaluation, with 36% of programs receiving no training whatsoever. Overall, comments from study participants also reflected that, while the Milestones are potentially an improvement on prior, less-structured methods of obtaining faculty feedback, there is still work to be done in order for faculty to better integrate these competency-based assessments into practical feedback and thus for residents to better understand how to clinically translate the results of Milestones assessments. Lastly, a 2017 survey-based study of PDs from the general surgery literature, similar to our results, afforded mixed reviews of Milestone utility.[5] In this investigation, 55% of PDs felt this system was an improvement on prior methods of evaluation, however 50% also felt the Milestones were unable to distinguish accurately between clinically successful and underperforming residents.

Despite the uncertainty that PDs expressed in such studies and echoed in our study, there is some evidence in the literature that the Milestones scores may accurately represent resident performance. Previous studies across numerous specialties, including data from our own institution, have indicated that the Milestones scores appropriately correlate between CCC and resident self-evaluations, and appropriately increase for increasing resident level.[6-10] Additionally, contrary to the perceptions by PDs in our study, there is evidence in other specialties that there may be some correlation between higher Milestones scores and other objective measures of resident success. A large study of 21,000 United States internal medicine residents correlated the Milestones assessments with American Board of Internal Medicine (ABIM) scores and found a higher MK competency score correlated with higher ABIM scores.[11] Additionally, residents who failed the ABIM had significantly lower mean MK scores than those who passed. Likewise, a recent study from a single general surgery program found some correlation between American Board of Surgery In-Training Examination percentage correct scores and all Milestone sub-competency scores, with the highest agreement between higher American Board of Surgery In-Training Examination scores and higher MK competency scores.[12]

The ACGME recognizes the Milestones are still imperfect and requires continuous re-evaluation. The ACGME has stated via their website they plan for increased involvement from the public in the creation of the next iteration of the Milestones, both via committees and via specialty-specific online surveys.[13] While no urology-specific surveys are currently available, this is a potential avenue for all PDs surveyed in this study to express their concerns regarding the Milestones system in urology directly to the ACGME to help improve the Milestones in the future.

Limitations of this study include its survey-based design, which introduces selection bias in those who chose to respond to the anonymous online survey. Additionally, we only polled PDs, therefore the opinions do not represent that of all members of the CCC at a given program. We also did not include resident opinions in this survey, which could have added valuable information. Finally, by keeping the survey anonymous, we were unable to control for multiple survey responses from a single PD, although we believe the incidence of this event to be unlikely. Future investigations on this topic might involve correlating the actual Milestones ratings amongst different programs nationally to determine consistency in their use for residents of similar postgraduate levels of training. This information may afford more objective data on how much standardization to the resident evaluation system the Milestones have afforded at this point.

CONCLUSION

The ACGME Milestones system for competency-based resident evaluation in urology is still in its early stages of implementation. This system requires continued internal review within our subspecialty nationally to assess efficacy and avenues for potential improvement. This data suggests there are still gains to be made to optimize the utility and applicability of this mandated rubric for resident evaluation and assessment.

SUPPLEMENTARY MATERIALS

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.urology.2018.10.042.

References

1. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366:1051–1056.
2. Swing SR, Beeson MS, Carraccio C, et al. Educational milestone development in the first 7 specialties to enter the next accreditation system. *J Grad Med Educ*. 2013;5:98–106.
3. Coburn M, Amling C, Bahnson RR, et al. Urology milestones. *J Grad Med Educ*. 2013;5(suppl 1):79–98.
4. Conforti LN, Yaghmour NA, Hamstra SJ, et al. The effect and use of milestones in the assessment of neurological surgery residents and residency programs. *J Surg Educ*. 2018;75:147–155.
5. Drolet BC, Marwaha JS, Wasey A, Pallant A. Program director perceptions of the general surgery milestones project. *J Surg Educ*. 2017;74:769–772.
6. Yao A, Massenburg BB, Silver L, Taub PJ. Initial comparison of resident and attending milestones evaluations in plastic surgery. *J Surg Educ*. 2017;74:773–779.
7. Watson RS, Borgert AJ, CT OH, et al. A multicenter prospective comparison of the accreditation council for graduate medical education milestones: clinical competency committee vs. resident self-assessment. *J Surg Educ*. 2017.
8. Goldman RH, Tuomala RE, Bengtson JM, Stagg AR. How effective are new milestones assessments at demonstrating resident growth? 1 year of data. *J Surg Educ*. 2017;74:68–73.

9. Lyle B, Borgert AJ, Kallies KJ, Jarman BT. Do attending surgeons and residents see eye to eye? An evaluation of the accreditation council for graduate medical education milestones in general surgery residency. *J Surg Educ*. 2016;73(6):e54–e58.
10. Ross FJ, Metro DG, Beaman ST, et al. A first look at the Accreditation Council for Graduate Medical Education anesthesiology milestones: implementation of self-evaluation in a large residency program. *J Clin Anesth*. 2016;32:17–24.
11. Hauer KE, Vandergrift J, Hess B, et al. Correlations between ratings on the resident annual evaluation summary and the Internal Medicine Milestones and Association With ABIM certification examination scores among US internal medicine residents, 2013-2014. *JAMA*. 2016;316:2253–2262.
12. Kimbrough MK, Thrush CR, Barrett E, Bentley FR, Sexton KW. Are surgical milestone assessments predictive of in-training examination scores? *J Surg Educ*. 2018;75:29–32.
13. Milestones Engagement. 2018. https://www.acgme.org/What-We-Do/Accreditation/Milestones/Engagement.

EDITORIAL COMMENT

This topic is extremely relevant to the Urology Graduate Medical Education (GME) community at large, as it shines a light on, not only the process, but the *progress* of educating urology residents in the era of Milestones. It is critical to review what we are doing and to gain perspective in order to determine the effectiveness of any intervention. The authors do just that by obtaining feedback from program directors about the usefulness of Milestones for urology residents. What the authors fail to do, however, is point out that imperfections are *expected* with an introduction of a fairly new intervention, like the Milestones.

The initial introduction of the Milestones was never meant to be the final process by which we assess residents. Perhaps we have forgotten the Outcomes Project, introduced by the Accreditation Council for Graduate Medical Education in 1999, nearly 20 years ago. This iteration of assessing residents provided us with the 6 core competencies, which were admittedly too broad to be effective in evaluating residents in all specialties. But, in my humble opinion, the Outcome Project did exactly what it was supposed to do at the time, which was to introduce to the GME community to the Competencies which were used to create the next iteration of resident assessment, the Milestones.

Historically, developing behavioral interventions involve an iterative process to build a robust body of evidence that initially supports feasibility, proves efficacy and effectiveness, and subsequently involves translation, implementation, and cultural change. The process of change occurs over several decades, and, even then, typically yields under 20% of the evidence being integrated into practice.

In my view, we need to question what we are doing. We need to be skeptical, as these authors are, of the process. Success in evaluating residents can only be measured when we are skeptical because it forces us to look closer at what we are presently doing—and to think about what we should be doing or what we could be doing better.

In the final analysis, though, assessing residents will take us truly believing in the process so that we can *earnestly* educate ourselves, our faculty, Clinical Competency Committees, and trainees about the goals of the Milestones. Milestones cannot teach residents, provide them with timely verbal feedback or written evaluations. This is the task of the faculty. But, faculty needs education of and dedication to the process in order for this fairly new intervention to be critically evaluated—and possibly successful.

Milestones provide good residents with feedback that can provide them assurance that they are moving in the right direction. But, in general, Milestones are not meant for good residents. Milestones—and all forms of assessment, are made for the early detection of residents who are struggling in a given specialty so that they can be remediated early enough to prevent a catastrophe and to guide them to the right place.

Although Milestones in and of themselves may not be the *apotheosis* of assessing residents, it is a large part of our foundation for creating an effective workforce. And, certainly, if we are not mindful of how to assess residents—and do our part to train them, it might be the *antithesis* to our profession.

Byron D. Joyner, MD, MPA, Children's Hospital and Regional Medical Center, Division of Pediatric Urology, Division of Urology, Seattle, WA 98105

https://doi.org/10.1016/j.urology.2018.10.043
UROLOGY 124: 32, 2019.

AUTHOR REPLY

We appreciate the editorial comments regarding our survey-based study. We agree the topic is extremely relevant to the urology community, and we hope our efforts can help initiate some of the necessary conversations to continue to refine the urology Milestones moving forward. Our intentions in creating this survey were purely to gain a better understanding of how the Milestones are currently viewed within the urology education community. We in no way mean to imply that we do not support the mission of the Accreditation Council for Graduation Medical Education Milestones project, and agree with the reviewers that the Milestones are a work in progress. We also agree with the sentiments that the Milestones are only useful in conjunction with feedback. In fact, as our survey indicates, most urology Program Directors have found that the Milestones have been successful in obtaining better feedback from faculty and useful in identifying residents who are struggling.

As we stated in the manuscript, the Milestones are still in the early stages of implementation, and we hope information, such as that obtained by our survey, will continue to spark internal review within our subspecialty in order to optimize the system. We agree with the reviewer's statements that educational and behavioral interventions are an iterative process. As we now have several years of information on the current Milestones, we believe the present is an appropriate time to start a critical review of the process in order to continue to improve our resident education and evaluation for the future.

Elisabeth M. Sebesta, Kimberly L. Cooper, Gina M. Badalato, MD, Department of Urology, NewYork-Presbyterian Hospital/Columbia University Medical Center, New York, NY; Columbia University Medical Center, 161 Fort Washington Avenue, 11th Floor, New York, NY 10032. *E-mail:* gmb2107@cumc.columbia.edu (G.M. Badalato).

https://doi.org/10.1016/j.urology.2018.10.044
UROLOGY 124: 32, 2019.