

Milestone Myths and Misperceptions

WALLACE A. CARTER JR, MD

In 1999 the Accreditation Council for Graduate Medical Education (ACGME) introduced 6 clinical competencies.¹ The longitudinal integration of these competencies into educational curricula and assessment of residents and fellows was intended to also generate outcomes assessments-based information on the educational effectiveness of training programs. In 2008, to further this process, ACGME Chief Executive Officer Thomas J. Nasca, MD, announced his vision for “The Next Step in the Outcomes-Based Accreditation Project,”² which became the Next Accreditation System (NAS). This vision was notable for 2 elements: development of specialty-specific, competency-based Milestones, along with design and implementation of Milestone assessment tools. The ultimate goal was to further transform graduate medical education from a process focus to an outcomes-based focus.

In July 2013, 7 specialties that had volunteered to be early adopters of the NAS began their journey down this exciting and relatively uncharted path.³ As a program director and member of the committee who created the Emergency Medicine Milestones, I have been struck by the incredible diversity of approaches during the Milestone rollout process.^{4,5} Given this diversity, inevitably some applications of the Milestone framework may not result in the lofty aspirational goals that we hoped for. The Milestones, along with other elements of the NAS, are intended to promote program improvements via self-study driven by objective data. The Next Accreditation System emphasizes trainee assessment based on observable behaviors, using stable, reproducible methods that eliminate that old educational saw, “I know it when I see it.” The community designing the Milestones aimed to make the assessment processes transparent and to stimulate creativity in the graduate medical education community with this paradigm shift. As the Milestones are beginning to be used, there are some misperceptions about the Milestones and misinterpretations in how they should be used (BOX). This perspective seeks to offer added clarity and address these misperceptions.

Using Milestones as Evaluation Tools

As residency programs have transitioned to the NAS and use of the Milestones, some have implemented the Milestones in ways quite different than envisioned. One of the primary unanticipated applications involved adopting the concept that Milestones are a de facto replacement for a global rating scale. Some have suggested that the Milestones should be used as the primary evaluation tool for assessing residents in a clinical environment. The contention is that since Milestones use a common language, they would control for interobserver variability and allow evaluators to completely assess a resident’s attainment of a given Milestone, sometimes during a single clinical encounter.

Nothing could be further from what Milestones were envisioned to accomplish. Can certain sections of individual Milestones be used to evaluate certain portions of a resident experience? Certainly. But most Milestones are complex, multifaceted, and sometimes fairly dense descriptions of a level of attainment on the road to competence for unsupervised practice in the specialty that do not lend themselves to being used as a stand-alone tool. More importantly, Milestones often are meant to be assessed by using multiple modalities—some of which exist and many of which still need to be created. Milestones were intended to be the antithesis of the one-size-fits-all assessment strategy. Many review committees developed Milestones that were intentionally broad enough to clarify an aspirational arrival point (readiness for practice), but require multiple strategies and methods to do so.

Some have also suggested that using Milestones as a global rating scale will limit grade inflation. When applied properly, Milestones should improve grading accuracy. But Milestones would be as vulnerable to grade inflation as other tools, if used as a single-evaluator global rating scale, because it is nearly impossible for a single observation to yield enough information to accurately assess each of the level anchors. What is more likely to occur is the “good guy syndrome,” where a trainee who is doing a good job will be given the benefit of the doubt that his or her good performance, as observed, will naturally spread across all of the domains of that Milestone, including those not observed. This phenomenon can also occur in the negative direction, producing a “reverse halo” effect. Milestones will limit grade inflation (inaccuracy) only if they are used as designed, and if programs use a combination of multiple evaluators with multiple assessment modalities to examine Milestone performance for the given trainee.

Wallace A. Carter Jr, MD, is Program Director, Emergency Medicine Residency, New York Presbyterian, Associate Professor of Emergency Medicine, Weill Cornell Medical College, and Associate Professor of Clinical Medicine, Columbia University College of Physicians & Surgeons.

Corresponding author: Wallace A. Carter Jr, New York Presbyterian Hospital, Weill Cornell Medical College, 525 East 68 Street, Mailbox 301, New York, NY 10021, 212.746.0892, wac9010@nyp.org

DOI: <http://dx.doi.org/10.4300/JGME-D-13-00438.1>

BOX COMMON MISINTERPRETATIONS AND MISPERCEPTIONS ABOUT MILESTONES

- Milestones are a global rating scale for all specialties
- Milestone levels accurately correspond to year of training
- The Clinical Competency Committee is a brand new process
- The assessment methodologies that are currently in place have no utility in Milestones
- The assessment methodologies for the Clinical Competency Committee are yet to be designed
- Milestones will eliminate grade inflation
- Milestone reporting will be so onerous that it will interfere with training
- Milestones should not be shared with residents

Consider the example of assessing a resident's ability to perform endotracheal intubation. Observing a resident doing an uncomplicated intubation, even on numerous occasions, would not give the 360-degree evaluation that would result if multiple evaluation methodologies were used. Even after observing multiple uncomplicated procedures the best that you could say about that resident, based on observations coming from that single context, is that the resident was technically adequate. It might be difficult, if not impossible, to assess the trainee's understanding of pharmacology, appropriate contraindications, or proficiency in airway rescue techniques. That level of analysis can only occur when the resident is evaluated across the spectrum of the available Milestone levels to arrive at the set of descriptors that best fits that trainee's performance.

Milestones and Level of Training

Another area of concern and some misperceptions is whether the Milestone levels correlate to level of training and whether trainees should be assessed only for that particular level. This would suggest a strict correlation between the 5 Milestone levels and residents' chronologic progression in training. While Milestones are meant to give an anchored, common language approach to training evaluation, they are not, for many specialties, to be tied to a specific year of training. This framework is essential as there are differences in how curricula and clinical experiences are sequenced in different programs across the nation. An added intrinsic advantage of the framework is that while attainment of a specific Milestone might be what is generally expected of a trainee at a certain point in time, trainees are by definition a heterogeneous group and will reach different Milestones at different times.

The benefit of having all levels of proficiency available, to be assigned by evaluators and ultimately the Clinical Competency Committee (CCC), is that residents possessing advanced skills can be acknowledged as having met that level regardless of the year in training and then be free to continue their journey to expertise. Imagine how

frustrating it may be for a trainee who entered medical school after a long period as a midlevel provider or in another health-related profession, who is a chronologic postgraduate year 1, but in fact possesses a level of skills and abilities far in excess of that level. If Milestones were to be limited to a specific year we would then lose the intrinsic benefit of identifying trainees with an exceptional skill set who could be used as teachers and emerging experts.

Clinical Competency Committees: Role and Function

Of the changes evoked by the NAS and Milestones, the composition, role, and work of the CCC may have generated the most discussion.⁶ The CCC may also be the element of the NAS that has the largest beneficial impact on trainees' experience and learning. CCCs are tasked with reviewing all residents semiannually and making decisions on Milestone achievement, with aggregate program level data being reported to the ACGME.

Some have opined that 1 way to simplify the work of the CCC is to replace the evaluation forms with new ones that consist of the Milestones. If the role of the CCC was to simply average the submitted scores from an evaluation card, a staff member with a calculator could replace this group. Rather, a Milestone and its various descriptors across the rating spectrum are discrete points to be evaluated by using multiple methods, resulting in assignment of a Milestone level by the CCC. This allows the committee creative license to work with program leaders and educational experts to develop additional methods to generate responses to the various questions posed by the Milestones.

Furthermore, the Milestones serve as a common standardized language that helps to eliminate institutional and personal idiosyncratic interpretations of resident performance. The new process allows the CCC and faculty to look at the knowledge, skills, and abilities that surround each Milestone and define which strategies most efficiently and accurately assess a trainee. Many of these strategies will be currently in use; examples include the observed structured clinical examination or the mini-clinical evaluation exercise. Most excitingly, the CCC process places a premium on educational innovation and technology.

It may help to visualize the Milestones as the hub of the wheel. The spokes of the wheel are the various options of evaluation available to the faculty. The job of the CCC and program director is to decide which methods to use to assess each anchor in the various levels and, through a data-anchored, conversation-based iterative process, assign a trainee to a Milestone level. A review of the Emergency Medicine Milestones shows that many Milestones use as many as 6 assessment strategies for a given level. This level

of assessment data will provide an incredible amount of educational information for our trainees as well as an indication to our patients that physicians are being trained and assessed in a uniform manner across the country. In this context, a problematic misperception is that Milestone information should *not* be shared with residents, when this information offers them new, relevant data on their performance and progress in their formal education as physicians.

Conclusion

It has been said that the NAS and Milestones are akin to designing a plane in flight. It should come as a surprise to no one that we are definitely in midflight with this process. Aviators will tell you that the takeoff is much simpler than the landing. There can be no doubt that at some point we will need to “land” this process and report our findings to our trainees, patients, and the American public. Starting in July 2014 we will all be aboard this incompletely designed airplane that is now in midflight. We need to make sure that we design a plane that can in fact make a safe and gentle landing. I predict that the need to make NAS and Milestones a success will result in substantial national and international creativity to develop new ways of assessing trainees.

In closing, I offer a recent personal observation. We are still far away from realizing the dream of Milestones and the NAS anchoring our country’s graduate medical education in a stable, predictable, outcomes-based process that anticipates rather than reacts. However, in a very short time we have made spectacular progress in offering our trainees a much richer assessment environment. I recently walked into our CCC meeting, which was in the midst of a

spirited discussion regarding a particular resident who was having some difficulties. Around the table were 8 academic emergency physicians, who collectively had more than 75 years of clinical and medical education experience. I watched as this discussion, lasting more than 15 minutes, offered various observations and prescriptions for helping the resident get over his “speed bump.” A number of things struck me. First, the deep level of respect and investment that the faculty felt in this trainee’s education and the lengths they were willing to go to help. Second, that as a result of the robust evaluation information offered to them by multiple assessments they were able to be precise in pinpointing the potential problem and its solution. Furthermore, I was struck that, owing to Milestones, this resident was receiving the undivided attention of skilled educators whose mission was to help him deliver better care to his patients.

While I realize that there is still much for us to do, I am confident that we are on the right flight path.

References

- 1 Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach*. 2007;29(7):648–654.
- 2 Nasca TJ. The next step in the Outcomes Based Accreditation Project. *ACGME Bull*. 2008;May 2–4.
- 3 Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366(11):1051–1056.
- 4 Beeson MS, Carter WA, Christopher TA, Heidt JW, Jones JH, Meyer LE, et al. Emergency Medicine Milestones. *J Grad Med Educ*. 2013;5(suppl 1):5–13.
- 5 Beeson MS, Carter WA, Christopher TA, Heidt JW, Jones JH, Meyer LE, et al. The development of the emergency medicine milestones. *Acad Emerg Med*. 2013;20(7):724–729.
- 6 Accreditation Council for Graduate Medical Education. ACGME Common Program Requirements. <http://www.acgme.org/acgmeweb/Portals/o/PFAssets/ProgramRequirements/CPRs2013.pdf>. Accessed December 3, 2013.