

MAY 2020

**Use of Individual Milestones Data by External Entities for High Stakes Decisions –
A Function for Which they Are not Designed or Intended**

Abstract

The Milestones are an educational and formative assessment methodology designed to help promote improvement in every specialty and subspecialty graduate medical education (GME) program in the United States. The Milestones were not designed or intended for use by external entities, such as state medical licensing boards or credentialing entities, to inform or make high stakes decisions. The ACGME is concerned that GME programs may artificially inflate individual Milestones assessment data if the Milestones are used for high stakes decisions. Their value would risk being lost as an honest and valuable assessment tool for continuous improvement and professional development.

The Milestones

The Milestones are an attempt to create a common language of professional development of resident and fellow physicians in each medical specialty and subspecialty. In July 2013, they were first implemented in residency programs by the ACGME in seven specialties, and subsequently they have been incorporated into accredited residency and fellowship programs in all specialties and subspecialties in the United States.¹

The primary goal of the Milestones is to drive improvement in educational experiences and assessment of residents and fellows in diverse clinical teaching settings across the country.

The Milestones are narrative descriptions of the development of resident and fellow abilities in each of six Core Competencies defined by the ACGME and the American Board of Medical Specialties (ABMS):

- Professionalism
- Patient Care and Procedural Skills
- Medical Knowledge
- Practice-based Learning and Improvement
- Interpersonal and Communication Skills
- Systems-based Practice

¹ In Academic Year 2018-2019, there were approximately 140,500 residents and fellows in 11,700 residency and fellowship programs in the United States.

Although these six domains of clinical competency are common to all specialties and subspecialties, the Milestones' developmental narratives are tailored to each specialty and subspecialty.

Residents and fellows are periodically assessed on the Milestones as they progress from the beginning to completion of a residency or fellowship program. The results serve as one of many guides for program personnel to chart the educational course of each resident and fellow.² To be effective in this regard, the assessments must be rigorous, accurate, and honest.

Currently, over 150 sets of specialty and subspecialty Milestones are in use in all ACGME-accredited residency and fellowship programs in the United States. The ACGME estimates that, to date, over 900 physicians and other experts throughout the United States have contributed over 27,000 volunteer hours in the development of the specialty and subspecialty Milestones.

The Milestones have been recognized by the public and the physician community in the United States as a promising approach to transforming GME. As highlighted in the 2014 National Academy of Medicine report, *Graduate Medical Education that Meets the Nation's Health Needs*:

The ACGME is currently implementing its "Next Accreditation System" (NAS) for all specialties. The new system was specifically developed to enhance the ability of the accreditation process to promote the training of physicians for practice in the 21st century. Assessments of educational outcomes and the clinical learning environment are key components of the NAS and are based on six core competencies—patient care, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism and systems-based practice.³

A 2017 ACGME study demonstrated that programs are using the Milestones to improve curriculum and assessment by using the data as an important part of program evaluation and improvement. This study showed that residents feel they receive more regular, organized, formalized feedback that is both richer and offers more clarity for expectations for progression through training.⁴ A 2019 ACGME study demonstrated

²Chappell, Kathy, Eric Holmboe, and Jacqueline Remondet Wall. 2018. "The Role of Health Care Profession Accreditors in Promoting Health and Well-Being across the Learning Continuum." *NAM Perspectives* 8 (11). <https://doi.org/10.31478/201811a>.

³Eden, Jill, Donald Berwick, and Gail Wilensky. 2014. *Graduate Medical Education That Meets the Nation's Health Needs*. *Graduate Medical Education That Meets the Nation's Health Needs*. National Academies Press. <https://doi.org/10.17226/18754>.

⁴Conforti, Lisa N., Nicholas A. Yagmour, Stanley J. Hamstra, Eric S. Holmboe, Benjamin Kennedy, Jesse J. Liu, Heidi Waldo, and Nathan R. Selden. 2018. "The Effect and Use of Milestones in the Assessment of Neurological Surgery Residents and Residency Programs." *Journal of Surgical Education* 75 (1): 147–55. <https://doi.org/10.1016/j.jsurg.2017.06.001>.

that the Clinical Competency Committees in different Sponsoring Institutions may vary in how they are assigning Milestone levels, with some becoming more stringent over time.⁵ Much work remains to be done to determine the extent to which the Milestones are useful and valid tools for residency and fellowship programs. However, based upon what has been learned, drafting is about to begin on the next version of the Milestones, with the expectation that revisions to all specialty and subspecialty Milestones will be completed by 2023.

Intended Use of the Milestones in Residencies and Fellowships

The Milestones provide a framework (i.e., a frame of reference or rubric) for a required periodic assessment of a resident or fellow in relation to a developmental description of attainment of specific, more granular sub-competencies over the course of the educational program curriculum. They guide the judgment of the program and the faculty members evaluating the residents and fellows in their respective programs; they do not and were not intended to represent (1) the totality of a specialty or subspecialty discipline, (2) complete assessment of all knowledge, skills, and attitudes, (3) norm-based reference criteria (i.e., comparison with other learners), or (4) a complete overall determination of a resident's or fellow's abilities. Moreover, they are tools used to provide an interim identification of progress in competency areas toward that necessary for unsupervised practice.

There is currently no "expected" or established rate of resident or fellow progression in Milestones achievement. Different residents and fellows learn different skills and concepts in different orders and at different rates. This is explicitly recognized in a position statement of the Federation of State Medical Boards from 1998 and still in effect today:

According to the ACGME, today there is wide variation in the timing and sequence of the various training elements among the 7000+ residency programs in the United States, and it is therefore impossible for state medical boards to discern, prior to completion of postgraduate training, which applicants for licensure have achieved appropriate training that qualifies them for a full and unrestricted license to practice medicine.⁶

ACGME-accredited residency and fellowship programs prepare the next generation of physicians, and the program directors of these programs are expected to attest to the preparedness of those who successfully complete their programs to serve the public independently in their respective specialty or subspecialty.

⁵ Hamstra, Stanley J., Kenji Yamazaki, Melissa A. Barton, Sally A. Santen, Michael S. Beeson, and Eric S. Holmboe. 2019. "A National Study of Longitudinal Consistency in ACGME Milestone Ratings by Clinical Competency Committees: Exploring an Aspect of Validity in the Assessment of Residents' Competence." In *Academic Medicine*, 94:1522–31. Lippincott Williams and Wilkins. <https://doi.org/10.1097/ACM.0000000000002820>.

⁶ Position of the Federation of State Medical Boards in Support of Postgraduate Training and Licensure Standards, adopted as policy by the Federation of State Medical Boards in 1998, page 3.

In the Milestones framework, everything else prior to the program director's final judgment of readiness or non-readiness for unsupervised practice is interim; the responsibility for the final judgment rests with the program director and supersedes all interim assessments, including Milestones ratings. The ACGME's accreditation requirements recognize the centrality of the program and program director's overall judgment relating to an individual resident's/fellow's readiness or non-readiness for unsupervised practice.

The judgment of the program, using a comprehensive and multifaceted approach to assessment, is paramount in determining the readiness of a resident or fellow to enter practice. The ACGME is sometimes asked whether a resident's Milestones data supersede a program director's judgment of readiness or non-readiness for independent practice. In fact, the reverse is true. As stated above, a program director's final, holistic, overall judgment at the end of the residency or fellowship program supersedes all interim assessments, as it represents the summative evaluation of the knowledge, skills, attitudes, and behaviors of the graduating resident or fellow at the time of graduation.

As noted previously, the ACGME has begun a process to revise each specialty and subspecialty set of Milestones. The rationale for revising the Milestones is an expression of the ACGME's commitment to continuous quality improvement on behalf of its stakeholders, and relates to their primary purposes, i.e., to facilitate the improvement of GME programs and to guide more effective professional development of residents and fellows. Once they have been revised, a new cycle of evaluation and research will begin to study the validity of the revised Milestones.

Non-Intended Use of Milestones by State Medical Licensing Boards

The ACGME does not have evidence that individual Milestones data can be validly used in any other context beyond provision of individual resident and fellow feedback, especially for any higher stakes decisions.⁷ In recognition of this, the following disclaimer appears at the beginning of the published Milestones in each specialty and subspecialty:

The Milestones are designed only for use in evaluation of resident physicians in the context of their participation in ACGME-accredited residency or fellowship programs. The Milestones provide a framework for the assessment of the development of the resident physician in key dimensions of the elements of physician competency in a specialty or subspecialty. They neither represent the entirety of the dimensions of the six domains of physician competency, nor are they designed to be relevant in any other context.

⁷ Aggregate Milestones data (as opposed to individual Milestones data) are used to facilitate national improvement efforts in curriculum and program design.

“Nor are they designed to be relevant in any other context” is intended to preclude the use of the Milestones in the context of physician licensure, or any other higher stakes use.

Consistent with this, the 26 ABMS member certifying boards and the certifying boards of the American Osteopathic Association do not use individual Milestones data for the purpose of assessing physician applicants for specialty board certification. Although ACGME accreditation requirements provide for residency and fellowship use of the Milestones, the ACGME does not review identified individual Milestones data for accreditation purposes. Instead, it views the data in aggregate, using the program as the unit of analysis.

The ACGME assumes that most state medical licensing boards (including osteopathic medical boards) heed the ACGME declaration that the Milestones are not designed for any non-residency/non-fellowship use.⁸ Nevertheless, the ACGME has learned of instances in which several state medical licensing boards have requested and used individual Milestones data in their decision on an individual physician’s license.

This is a non-designed and non-intended use of Milestones data. For licensure decisions after completion of a GME program, it ignores the program director’s judgment of readiness for unsupervised practice upon completion of the program. For licensure decisions before and after completion of a GME program, (1) it ignores the disclaimer for this use included at the beginning of each set of Milestones; (2) it is inappropriate to compare one specialty’s Milestones assessments against another, as specialty programs have different content and different durations; and (3) all states grant general licenses rather than specialty licenses, and the Milestones are specialty-specific. For licensure decisions before completion of a GME program, it ignores the fact that each Milestones assessment is against the entirety of the curriculum, GME programs in the same specialty do not necessarily order the curriculum in the same way, and the same GME program may alter the order of its curriculum from year to year as part of its improvement process.

Potential Negative Consequences of Non-Intended Use of Individual Milestones Data:

State Medical Licensing Boards

One consequence of this non-designed and non-intended use of the Milestones by state medical licensing boards might be an adverse licensure decision being reversed on administrative review if Milestones data were used as part of the decision. This is a real possibility since the ACGME, as an original and continuing developer of the Milestones, clearly declares that Milestones data is not designed or intended for that purpose.

⁸ In addition, on a state by state basis, state statutes and/or case law may protect individual Milestones data from submission to the state medical licensing board, as well as from production in litigation and under state public records acts.

But there is a second and more universal consequence. Milestones assessment occurs within a learning context. Residency and fellowship programs use the Milestones to guide a learning course on a per-resident/per-fellow basis, as each individual physician learner progresses to clinical independence in the specialty or subspecialty, according to a uniform set of criteria within that specialty/subspecialty. Integrity in the assessment process is critical to the function of the learning process.

The ACGME is concerned about the possibility that the GME community might artificially inflate Milestones assessment data were the Milestones to be used, or perceived to be used, for high-stakes decisions such as the granting of individual licenses by state medical licensing boards. There is ample scientific evidence that validates this concern, in particular, “Campbell’s Law”⁹ demonstrates that “any indicator monitoring something anyone believes to be important is a candidate for being influenced.”¹⁰ The value of the Milestones as an accurate and honest assessment tool would risk being lost.

Conclusion

The Milestones are a framework of assessments for the six Core Competencies, intended as one among many tools to inform and guide the learner and the members of the faculty as the learner progresses through the residency or fellowship curriculum. They are not designed or intended to supplant the overall judgment of the program director as to the ability of an individual learner to perform particular clinical tasks during the residency or fellowship, or to enter the independent practice of medicine upon completion of the residency or fellowship.

It is important that individual Milestones assessments be used and maintained within each GME program to preserve them as robust and accurate tools in the learning process. Without such limitation of use, the residency or fellowship program might be tempted to artificially assess the individual more positively for the consumption of a state medical board, and thus jeopardize the Milestones as a learning and teaching tool.

⁹ Campbell, Donald T. 1979. “Assessing the Impact of Planned Social Change.” *Evaluation and Program Planning* 2 (1): 67–90. [https://doi.org/10.1016/0149-7189\(79\)90048-X](https://doi.org/10.1016/0149-7189(79)90048-X).

¹⁰See also: Sharon Nichols Assistant Professor, by L, and David C Berliner Regents. 2005. “The Inevitable Corruption of Indicators and Educators Through High-Stakes Testing.” <https://files.eric.ed.gov/fulltext/ED508483.pdf>