Ways to Write a Milestone: Approaches to Operationalizing the Development of Competence in Graduate Medical Education

Andrea N. Leep Hunderfund, MD, MHPED, Darcy A. Reed, MD, MPH, Stephanie R. Starr, MD, Rachel D. Hayser, MD, Tara R. Lang, MD, and Suzanne M. Norby, MD

Abstract

Purpose
To identify approaches to operationalizing the development of competence in Accreditation Council for Graduate Medical Education (ACGME) milestones.

Method
The authors reviewed all 25 “Milestone Project” documents available on the ACGME Web site on September 11, 2013, using an iterative process to identify approaches to operationalizing the development of competence in the milestones associated with each of 601 subcompetencies.

Results
Fifteen approaches were identified. Ten focused on attributes and activities of the learner, such as their ability to perform different, increasingly difficult tasks (304/601; 51%), perform a task better and faster (171/601; 45%), or perform a task more consistently (123/601; 20%). Two approaches focused on context, inferring competence from performing a task in increasingly difficult situations (236/601; 29%) or an expanding scope of engagement (169/601; 28%). Two used socially defined indicators of competence such as progression from “learning” to “teaching,” “leading,” or “role modeling” (271/601; 45%). One approach focused on the supervisor’s role, inferring competence from a decreasing need for supervision or assistance (151/601; 25%). Multiple approaches were often combined within a single set of milestones (mean 3.9, SD 1.6).

Conclusions
Initial ACGME milestones operationalize the development of competence in many ways. These findings offer insights into how physicians understand and assess the developmental progression of competence and an opportunity to consider how different approaches may affect the validity of milestone-based assessments. The results of this analysis can inform the work of educators developing or revising milestones, interpreting milestone data, or creating assessment tools to inform milestone-based performance measures.

The shift in medical education from content- to competency-based curricula has fueled efforts to assess learners based on how well they perform various aspects of a physician’s expected role in society.1–3 The developmental progression of competence can be described using milestones, the framework selected by the Accreditation Council for Graduate Medical Education (ACGME) for reporting educational outcomes.4 To prepare for the ACGME Next Accreditation System, specialty groups developed milestones for subcompetencies relevant to their disciplines.4–6 The resulting milestones provide insights into how competence and its development are conceptualized in graduate medical education.

Competence is a complex construct that can be viewed through many different lenses.1 Likewise, the development of competence can be operationalized in different ways.5,6 The approach selected for a particular set of milestones influences its validity and, hence, may have far-reaching downstream effects on the instruction and assessment of learners, evaluation of training programs, and, ultimately, the confidence educators and society can place in milestone-based measures of competence.6 In undertaking this study, we aimed to identify approaches to operationalizing the development of competence in ACGME milestones and compare approaches used for different core competencies.

Method

Data source
We downloaded the 25 specialty-specific “Milestone Project” documents available on the ACGME Next Accreditation System Web site10 on September 11, 2013; these were allergy, aerospace medicine, colon and rectal surgery, dermatology, diagnostic radiology, emergency medicine, family medicine, general surgery, internal medicine, medical genetics, neurological surgery, neurology, nuclear medicine, occupational medicine, ophthalmology, orthopedic surgery, pathology, pediatrics, physical medicine and rehabilitation, plastic surgery, psychiatry, public health and general preventative medicine, radiation oncology, transitional year, and urology.

The Milestone Project documents of these 25 specialties contained 601 subcompetencies (median 23, range 10–41). Of these, 37% pertained to patient care (PC, 221/601), 19% to medical knowledge (MK, 112/601), 12% to systems-based practice (SBP, 72/601), 12% to professionalism (71/601), 11% to practice-based learning and improvement (PBLI, 64/601), and 10% to interpersonal and communication skills (ICS, 61/601). The origin and organization of these subcompetencies (and associated
milestones) have been previously described. A,11,12

We selected subcompetencies as the unit of analysis to identify how the development of competence was operationalized within the associated milestones.

Data analysis

We divided the 25 Milestone Project documents among the six members of the study team, who represent different specialties and have 70 years of combined experience in medical education (including advanced degrees or fellowship training, education research experience, and institutional education leadership roles). One investigator (S.M.N.) reviewed the milestones of all 25 specialties, and other team members each reviewed the milestones of 5 to 8 specialties. Team members analyzed each set of milestones to identify approaches to operationalizing the development of competence. We did not analyze the milestones with preselected approaches in mind. Rather, we performed an inductive analysis, allowing different approaches to emerge naturally.13

Our team met approximately monthly throughout 2014 and early 2015 to compare and discuss our findings. During this process, we recognized that many of the approaches we were identifying could be mapped to existing conceptual frameworks related to the processes of learning and skill development, and these frameworks (listed in Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/A437) informed subsequent discussions. We reconciled differences through discussion until consensus was achieved regarding a final list of 15 approaches.

The primary investigator (A.N.L.) subsequently re-reviewed the milestones of all 25 specialties to classify which approaches to operationalizing the development of competence were used in the milestones associated with each subcompetency. During this process, no new approaches were identified beyond those initially defined. We then compared the approaches used for different core competencies (MK, PC, PBLI, SBP, ICS, and professionalism).

Results

Table 1 lists the 15 approaches we identified among 601 sets of ACGME milestones from 25 specialties. To enhance the usefulness of these approaches, we grouped them into four larger categories according to whether they focused on the learner, the context, social interactions, or the supervisor. Examples of milestones using each approach are provided in Supplemental Digital Appendix 1 at http://links.lww.com/ACADMED/A437.

The first and largest category of approaches focused on activities and attributes of the learner. Approaches in this category commonly inferred competence from a learner’s ability to perform different, increasingly difficult tasks (304/601; 51%) or to perform a given task better or faster (171/601; 28%). A related approach relied on progression from performing parts of a task to performing a whole task (74/601; 12%). For example, the emergency medicine milestones for emergency stabilization describe progression from “recognizing abnormal vital signs” to “performing a primary assessment on a critically ill or injured patient,” to “managing a critically ill or injured patient.”14

Another common learner-oriented approach inferred competence from the consistency with which a behavior or skill was demonstrated (123/601; 20%). For example, the colon and rectal surgery milestones used “rarely,” “occasionally,” and “consistently” to describe the development of competence in SBP, ICS, PBLI, and professionalism.15 The allergy and immunology milestones for emergency stabilization described progression from “recognizing abnormal vital signs” to “performing a primary assessment on a critically ill or injured patient,” to “managing a critically ill or injured patient.”14

Two learner-oriented approaches had an attitudinal dimension: progression from “unwilling” to “willing but unable” to “able” to perform a task (57/601; 9%) and progression from extrinsic to intrinsic motivation (100/601; 17%). The former approach is illustrated by a set of internal medicine milestones for PBLI that operationalized the development of competence as progression from “unwilling” to self-reflect (Level 1, critical deficiency) to “unable” to self-reflect (Level 2) to “able” to self-reflect with increasing consistency.17 The latter approach is demonstrated by a set of pediatrics milestones for professionalism that describe progression from help-seeking only in response to external prompts (Level 2) to help-seeking arising from an internalized personal value system where patient care supersedes any perceived value of physician autonomy (Level 4).18 Lack of insight into personal limitations (or professional expectations) was sometimes cited as an indicator of Level 1 performance for these approaches, akin to what has been described as “unskilled and unaware.”19

The remaining learner-oriented approaches inferred competence from increasing knowledge (150/601; 25%), progression from “awareness” to “knowledge” to “ability” (141/601; 23%), and progression from “knowing,” “applying,” “analyzing,” or “evaluating,” to “creating” (281/601; 47%), as evidenced by scholarly contributions (research projects, educational materials, clinical practice guidelines, etc.). A minority relied on more traditional approaches to assessment such as normative comparisons (e.g., “demonstrates level-appropriate knowledge for patient management on post-graduate year 2 rotations”20; 7/601; 1%), increasing experience or completion of special training (e.g., number of procedures performed, advanced cardiovascular life support certification; 16/601; 3%), or scores on a written test (7/601; 1%).

The second category of approaches focused primarily on the context. For example, nearly 40% of milestones inferred competence from the ability to perform a given task in increasingly difficult or infrequently encountered situations (236/601; 39%), which were often described as complex, rare, high-acuity, rapidly changing, adversarial, or ambiguous. Chart 1 shows a set of pediatrics milestones combining this context-oriented approach with a learner-oriented approach (ability to perform a given task better or faster). This example describes how a resident’s ability to gather essential and accurate information about a patient becomes more effective, more efficient, and more equipped to deal with complex or rare situations as their skill level increases.18
Table 1
Approaches to Operationalizing the Development of Competence in the Initial ACGME Milestones of 25 Specialties by Core Competency

<table>
<thead>
<tr>
<th>Initial ACGME milestones infer increasing competence from...</th>
<th>PC (n = 221)</th>
<th>MK (n = 112)</th>
<th>SBP (n = 72)</th>
<th>PROF (n = 71)</th>
<th>PBLI (n = 64)</th>
<th>ICS (n = 61)</th>
<th>Total (n = 601)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learner-oriented approaches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Performing different, increasingly difficult tasks</td>
<td>150 (68)</td>
<td>21 (19)</td>
<td>42 (58)</td>
<td>25 (35)</td>
<td>39 (61)</td>
<td>27 (44)</td>
<td>304 (51)</td>
</tr>
<tr>
<td>2. Performing a given task better or faster</td>
<td>92 (42)</td>
<td>6 (5)</td>
<td>12 (17)</td>
<td>19 (27)</td>
<td>13 (20)</td>
<td>29 (48)</td>
<td>171 (28)</td>
</tr>
<tr>
<td>3. Progression from performing parts of a task to performing a whole task</td>
<td>66 (30)</td>
<td>1 (1)</td>
<td>2 (3)</td>
<td>0 (0)</td>
<td>5 (8)</td>
<td>0 (0)</td>
<td>74 (12)</td>
</tr>
<tr>
<td>4. Increasing consistency with which a behavior or skill is demonstrated</td>
<td>29 (13)</td>
<td>7 (6)</td>
<td>16 (22)</td>
<td>37 (52)</td>
<td>15 (23)</td>
<td>19 (31)</td>
<td>123 (20)</td>
</tr>
<tr>
<td>5. Progression from “unwilling” to “willing but unable” to “able” to perform a task</td>
<td>6 (3)</td>
<td>0 (0)</td>
<td>5 (7)</td>
<td>22 (31)</td>
<td>17 (27)</td>
<td>7 (11)</td>
<td>57 (9)</td>
</tr>
<tr>
<td>6. Progression from extrinsic to intrinsic motivation and/or increasing drive to improve</td>
<td>16 (7)</td>
<td>4 (4)</td>
<td>7 (10)</td>
<td>42 (59)</td>
<td>29 (45)</td>
<td>2 (3)</td>
<td>100 (17)</td>
</tr>
<tr>
<td>7. Increasing knowledge</td>
<td>30 (14)</td>
<td>74 (84)</td>
<td>12 (17)</td>
<td>4 (6)</td>
<td>8 (13)</td>
<td>3 (5)</td>
<td>150 (25)</td>
</tr>
<tr>
<td>8. Progression from awareness to knowledge to ability</td>
<td>48 (22)</td>
<td>13 (12)</td>
<td>36 (50)</td>
<td>18 (25)</td>
<td>8 (13)</td>
<td>18 (30)</td>
<td>141 (23)</td>
</tr>
<tr>
<td>9. Progression from knowing, applying, analyzing, or evaluating to creating</td>
<td>87 (39)</td>
<td>87 (78)</td>
<td>29 (40)</td>
<td>26 (37)</td>
<td>27 (42)</td>
<td>25 (41)</td>
<td>281 (47)</td>
</tr>
<tr>
<td>10. Increasing experience or completion / maintenance of special training</td>
<td>10 (5)</td>
<td>2 (2)</td>
<td>3 (4)</td>
<td>1 (1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>16 (3)</td>
</tr>
<tr>
<td>Context-oriented approaches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Performing a task in increasingly difficult or infrequently encountered situations</td>
<td>137 (62)</td>
<td>14 (13)</td>
<td>10 (14)</td>
<td>35 (49)</td>
<td>5 (8)</td>
<td>35 (57)</td>
<td>236 (39)</td>
</tr>
<tr>
<td>12. Expanding scope of engagement</td>
<td>48 (22)</td>
<td>10 (9)</td>
<td>43 (60)</td>
<td>34 (48)</td>
<td>12 (19)</td>
<td>22 (36)</td>
<td>169 (28)</td>
</tr>
<tr>
<td>Socially oriented approaches</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Progression from learning to teaching, leading, role modeling, or being a consultant for others</td>
<td>87 (39)</td>
<td>22 (20)</td>
<td>37 (51)</td>
<td>54 (76)</td>
<td>30 (47)</td>
<td>41 (67)</td>
<td>271 (45)</td>
</tr>
<tr>
<td>14. Increasing interdependence with other members of the healthcare team</td>
<td>14 (6)</td>
<td>0 (0)</td>
<td>23 (32)</td>
<td>6 (8)</td>
<td>3 (5)</td>
<td>13 (21)</td>
<td>59 (10)</td>
</tr>
<tr>
<td>Supervisor-oriented approach</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Decreasing need for supervision or assistance</td>
<td>94 (43)</td>
<td>10 (9)</td>
<td>4 (6)</td>
<td>8 (11)</td>
<td>25 (39)</td>
<td>10 (16)</td>
<td>151 (25)</td>
</tr>
</tbody>
</table>

Abbreviations: ACGME indicates Accreditation Council for Graduate Medical Education; MK, medical knowledge; PC, patient care; PBLI, practice-based learning and improvement; SBP, systems-based practice; ICS, interpersonal and communication skills; PROF, professionalism.

*Accessed September 11, 2013. Specialties are allergy and immunology, aerospace medicine, colon and rectal surgery, dermatology, diagnostic radiology, emergency medicine, family medicine, general surgery, internal medicine, medical genetics, neurology, neurological surgery, nuclear medicine, occupational medicine, ophthalmology, orthopaedic surgery, pathology, pediatrics, plastic surgery, physical medicine and rehabilitation, psychiatry, public health and general preventative medicine, radiation oncology, transitional year, and urology.

*Numbers in each column do not add up to the total for that column because multiple approaches to operationalizing the development of competence were combined within one set of milestones.

*Sometimes evidenced by compliance with rules and regulations for extrinsic motivation, n = 26.

*As evidenced by the type of information known (e.g., its complexity, rarity, novelty; n = 131), amount of information known (n = 36), or scores on written tests (n = 7).

The third category of approaches used socially defined indicators of competence, including progression from “learning” to “teaching,” “leading,” “role modeling,” or “being a consultant for others” (271/601; 45%). For example, the internal medicine milestones described aspirational performance in ICS as “role models and teaches collaborative communication with the team to enhance patient care” and in PBLI as “is able to lead a quality improvement project.”

One approach in the final category focused on the role of the supervisor, inferring competence from a decreasing need for supervision or assistance (151/601; 25%). This approach invokes the concept of scaffolding21 and was particularly explicit in the medical genetics milestones, which used the words “with substantial guidance,” “with minimal guidance,” and “independently” to describe increasing levels of performance within many subcompetencies.22

Multiple approaches were often contained within the milestones associated with a single subcompetency (mean 3.9, SD 1.6). Chart 2 provides an example from the neurology milestones illustrating how three different approaches to operationalizing the development of competence were combined within one set of milestones.23

Approaches to operationalizing the development of competence varied by
core competency (Table 1). For example, milestones associated with the 221 PC subcompetencies often operationalized the development of competence as the ability to perform different, increasingly difficult tasks (150/221; 68%) or to perform a task in increasingly difficult or infrequently encountered situations (137/221; 62%). PC milestones were also the most likely to infer competence from a decreasing need for supervision or assistance (94/221; 43%).

Milestones associated with the 112 MK subcompetencies usually inferred competence from increasing knowledge (74/112; 84%) as evidenced by the type of information known (e.g., its complexity, rarity, novelty), amount of information known, or scores on a written test (usually a resident in-training examination). Test scores were defined in different ways, including a minimum percent correct, a minimum overall raw score, a scaled score predictive of passing a certification examination, a passing score, improvement in the percent correct compared with a previous score, or an “acceptable” percentile ranking.

Milestones associated with the 72 SBP subcompetencies frequently operationalized the development of competence as an expanding scope of engagement (e.g., from individual patients to groups of patients to populations or systems of care; 43/72; 60%) or increasing interdependence with other members of the health care team (e.g., progression from “dismissing” to “receiving” to “valuing and actively soliciting” input from interprofessional team members). SBP milestones were also the most likely to use progression from “awareness” to “knowledge” to “ability” (36/72; 50%), as illustrated by a set of pathology milestones describing the development of competence in “technology assessment” (Chart 3). Milestones related to the 71 professionalism and 64 PBLI subcompetencies commonly included an attitudinal dimension. For example, 59% (42/71) of professionalism milestones operationalized the development of competence as progression from extrinsic to intrinsic motivation or an increasing drive to improve, as did 45% (29/64) of milestones related to PBLI. Professionalism milestones were also the most likely to use progression from “unwilling” to “willing but unable” to “able” to perform a task (22/71; 31%) and increasing consistency with which a behavior or skill is demonstrated (37/71; 52%).

Chart 1

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Either gathers too little information or exhaustively gathers information following a template regardless of the patient’s chief complaint, with each piece of information gathered seeming as important as the next.</td>
<td>Clinical experience allows linkage of signs and symptoms of a current patient to those encountered in previous patients. Still relies primarily on analytic reasoning through basic pathophysiology to gather information, but has the ability to link current findings to prior clinical encounters allowing information to be filtered, prioritized, and synthesized into pertinent positives and negatives, as well as broad diagnostic categories.</td>
<td>Demonstrates an advanced development of pattern recognition that leads to the creation of illness scripts, which allow information to be gathered while simultaneously filtered, prioritized, and synthesized into specific diagnostic considerations. Data gathering is driven by real-time development of a differential diagnosis early in the information-gathering process.</td>
<td>Creates well-developed illness scripts that allow essential and accurate information to be gathered and precise diagnoses to be reached with ease and efficiency when presented with most pediatric problems, but still relies on analytic reasoning through basic pathophysiology to gather information when presented with complex or uncommon problems.</td>
<td>Creates robust illness scripts and instance scripts (where the specific features of individual patients are remembered and used in future clinical reasoning) that lead to unconscious gathering of essential and accurate information in a targeted and efficient manner when presented with all but the most complex or rare clinical problems. These illness and instance scripts are robust enough to enable discrimination among diagnoses with subtle distinguishing features.</td>
</tr>
</tbody>
</table>

Abbreviation: PC indicates patient care.

a Reprinted with permission from The Pediatrics Milestone Project. https://www.acgme.org/acgmeweb/Portals/0/PDFs/Milestones/PediatricsMilestones.pdf. Published January 2013. Milestone is subject to a copyright belonging to the American Board of Pediatrics, Inc. and the Accreditation Council for Graduate Medical Education.

b Reprinted from the American Medical Association. Vol. XX, No. X / XX XXXX

Copyright © by the Association of American Medical Colleges. Unauthorized reproduction of this article is prohibited.
those charged with revising ACGME milestones, developing new milestones (e.g., for undergraduate medical education or emerging competencies), interpreting milestone data (the ACGME and associated review committees), using assessment tools to inform milestone-based measures of trainee performance (training programs and clinical competency committees), and researchers gathering validity evidence for milestones-based assessments.

Grounding milestones in conceptual frameworks contributes to their content validity and may enhance the ease with which faculty use and understand milestones (and associated assessment instruments), hence bolstering their response process validity. Encouragingly, many of the approaches we identified can be mapped to established conceptual frameworks such as the Dreyfus model of skill acquisition (progression from “novice” to “advanced beginner” to “competent” to “proficient” to “expert”), Bloom’s taxonomy of objectives for the cognitive domain (progression from “knowing,” “applying,” “analyzing,” or “evaluating” to “creating”), Miller’s pyramid (progression from “awareness” to “knowledge” to “ability”), Ericsson’s definition of expert performance (performing a task better, faster, or with increasing consistency), Hatano and Inagaki’s description of adaptive expertise (performing a task in increasingly difficult or infrequently encountered situations), and Vygotsky’s zone of proximal development (progression from performing parts of a task to performing the whole task or a decreasing need for supervision or assistance). Basing milestones on conceptual frameworks, as was explicitly done by some specialty groups, will allow the education community to build upon previous work and gain a deeper understanding of how competence is developed.

Conceptual frameworks may also enhance the reliability of milestone-based assessments both within and across institutions. Approaches that align well with the cognitive schemas or “reality maps” of raters may be especially advantageous in this regard. For example, physicians are already accustomed to making entrustment decisions and, hence, may find it particularly intuitive to infer competence from a decreasing need for supervision or assistance. Use of concrete anchor statements also increases the likelihood that milestones will be interpreted similarly across different institutions (e.g., “>90% of the time” rather than “consistently”).

The approach selected for a set of milestones may also influence their construct and consequential validity. Specifically, milestones are intended to strengthen the public accountability of graduate medical education, enabling programs to make informed decisions about the degree of supervision required by each trainee to ensure high-quality care. Furthermore, graduation from an accredited training program should indicate to patients that a resident or fellow is competent and ready for unsupervised practice in their specialty (with the expectation that they will continue to develop further expertise over the course of their professional career). However, only 25% of ACGME milestones directly infer competence from a decreasing need for supervision or assistance (i.e., a trainee’s ability to perform a task independently). This complicates use of milestone data by the ACGME and specialty review committees and suggests that convergence of approaches across specialties may be desirable—especially for Level 4 milestones, which represent the graduation target. Particular consideration could be given to inferring competence from a decreasing need for supervision or assistance, as this approach most directly speaks to a trainee’s readiness for independent practice. Additional studies are needed to determine whether milestones generated by alternative approaches align with entrustment decisions.
The initial ACGME milestones have been referred to as “version 1.0,” anticipating the need for iterative improvements and revisions. Our findings provide milestones workgroups, whose initial efforts were somewhat siloed within specialties, with an opportunity to reexamine their milestones in light of approaches used by other specialties. In doing so, educators should consider which approaches to operationalizing the development of competence are most likely to support the reliability, validity, and intended purposes of milestone-based assessments and use this information to guide subsequent revisions.

The optimal approach for a given subcompetency will likely vary depending on the nature of the subcompetency. For example, progression from extrinsic to intrinsic motivation may be particularly well suited for operationalizing the development of competence within professionalism or PBLI, whereas increasing interdependence with other members of the health care team is well suited for ICS or SBP subcompetencies related to teamwork.

Another important consideration is the level of performance expected of incoming learners. For example, many SBP milestones operationalize the development of competence as progression from “awareness” to “knowledge” to “ability.” This approach assumes that most incoming residents will have little to no knowledge or skills related to SBP. However, the skill level of incoming residents in SBP will likely increase over time as medical schools incorporate education on health care systems and systems thinking into their curricula. As the knowledge and skills of incoming residents improve, approaches based on “performing a given task better or faster” or “performing a task in increasingly difficult or infrequently encountered situations” may become more appropriate. This highlights the dynamic nature of the milestones, which are expected to evolve and change over time.

Our findings demonstrate that multiple approaches to operationalizing the development of competence are often combined within a single set of milestones. This may add value, as different conceptual frameworks can offer complementary perspectives on the developmental progression of performance within a particular subcompetency. However, combining multiple approaches without thoughtful consideration risks confusing rather than clarifying the picture. Combining different approaches within a single set of milestones also complicates the assessment process for teaching faculty (who often already struggle to translate milestone language into directly observable behaviors) and clinical competency committees. Training programs may similarly find it challenging to implement evaluation systems capable of adequately assessing multiple dimensions of a trainee’s performance (the types of tasks performed, the situations in which they were performed, the amount of supervision required, the consistency of performance, etc.) across multiple subcompetencies. This consideration is particularly important given the complex logistical challenges associated with procuring longitudinal assessment data and providing faculty with multiple opportunities to assess trainee performance.

Some approaches may more readily lend themselves to workplace-based assessment than others. For example, many milestones inferred competence from a learner’s ability to perform a task in increasingly difficult or infrequently encountered situations. By their nature, such experiences may be rare and variable across learners, making assessment strategies difficult to implement. In such cases, training programs may instead turn to simulation; however, this can be resource intensive and may not predict performance in actual workplace settings.

Finally, we observed that some milestones continue to rely on approaches to assessment that have historically been more common, such as normative comparisons, increasing experience or completion of special training, and scores on a written test. While approaches like these may be appealing because of their familiarity, they are not well suited for a milestones framework. Normative comparisons, for example, compare trainee performance against that of other trainees rather than measuring performance against a competency standard. Approaches based on increasing experience (i.e., number of procedures performed) or completion of special training are likewise suboptimal because they measure educational processes rather than educational outcomes. Furthermore, specialty boards sometimes require different numbers to achieve competence for a given procedure, which undercuts the validity of this approach.

Limitations of this study include the performance of analyses by a select group of individuals from a single institution. Additionally, approaches to operationalizing the development of competence were coded at the subcompetency level rather than the individual milestone level. Further study is needed to determine whether approaches are combined in predictable ways across milestone levels or whether certain approaches tend to co-occur.

In conclusion, initial ACGME milestones operationalize the development of competence in many ways. This variety, both within and across specialties, provides an opportunity to investigate the relative utility, acceptability, and meaningfulness of different approaches. Milestone validity may be enhanced by using approaches that are grounded in conceptual frameworks, align with the cognitive schemas of raters, correlate with entrustment decisions, use criterion-referenced outcomes, and are informed by the nature of the subcompetency and skill level of incoming learners.

Funding/Support: This study was prepared with financial support from the American Medical Association as part of the Accelerating Change in Medical Education initiative, and all participating schools received grants through this initiative (see www.changemeded.org for further details) and from the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery. The content reflects the views of the authors.

Other disclosures: None reported.

Ethical approval: Reported as not applicable.

Previous presentations: A preliminary version of these data was presented at the University of Illinois at Chicago 16th Annual Masters of Health Professions Education Summer Conference on July 30, 2015; at the Academy for Professionalism in Health Care 4th Annual Conference on April 28, 2016; and as a poster at the Society of General Internal Medicine Annual Meeting on May 11, 2016.

A.N. Leep Hunderfund is assistant professor of neurology, Mayo Clinic, Rochester, Minnesota.

D.A. Reed is associate professor of medical education and medicine and senior associate dean of academic affairs, Mayo Medical School, Mayo Clinic, Rochester, Minnesota.
S.R. Starr is assistant professor of pediatric and adolescent medicine and director of science of health care delivery education, Mayo Medical School, Mayo Clinic, Rochester, Minnesota.

R.D. Haverty is assistant professor of medicine, Mayo Clinic, Rochester, Minnesota.

T.R. Lang is assistant professor of pediatric and adolescent medicine, Mayo Clinic, Rochester, Minnesota (now at Gundersen Health System, LaCrosse, Wisconsin).

S.M. Norby is associate professor of medicine, Mayo Clinic, Rochester, Minnesota.

References


