The Milestones Guidebook

Laura Edgar, EdD, CAE
Sydney McLean, MHA
Sean O. Hogan, PhD
Stan Hamstra, PhD
Eric S. Holmboe, MD

Version 2020

TABLE OF CONTENTS

Preface

The Milestones have become an important formative component of the ACGME current accreditation model for graduate medical education (GME) in the United States. This accreditation model, previously dubbed “the Next Accreditation System” was part of the educational community’s response to public and policy makers’ concerns regarding the need to improve GME (Nasca et al. 2012). It more fully embraces the outcomes-based principles that started with the release of the Core Competencies in 1999, and the launch of the Outcome Project in 2001 (Batalden et al. 2002; IOM 2014). However, the ACGME and GME programs struggled to operationalize the Core Competencies and create meaningful outcomes-based assessments. Recognizing these challenges, the ACGME’s transition to the current model included two important new components to the accreditation process; the Milestones and the Clinical Competency Committee (CCC), both of which are designed to monitor and iteratively improve educational outcomes, and by extension, clinical outcomes, at the level of the individual learner and the program.

This Milestones Guidebook is designed to be informational and practical. The beginning sections explain the grounding principles of competency-based medical education (CBME) and development of the Milestones. The remaining sections include practical guidance on how to effectively use the Milestones. “Practical Tips” boxes that offer quick summaries are provided in certain sections. Finally, appendices provide a list of useful information on assessment systems and additional CBME resources that may help programs with implementation of the Milestones.

Other guidebooks are available in the Milestones section of the ACGME website, including a Milestones Guidebook for Residents and Fellows (written by and for residents and fellows), a Clinical Competency Committee Guidebook, and the newest addition, a Milestones Implementation Guidebook. All of these and other resources are available at https://www.acgme.org/What-We-Do/Accreditation/Milestones/Resources.

Feedback on this second edition of the *Milestones Guidebook* is invited and welcomed. Send comments to milestones@acgme.org.

Competency-Based Education and Assessment and the Rationale for the Educational Milestones

A brief historical timeline of the move toward competency-based education and assessment provides the context and rationale for use of the educational Milestones in the ACGME's accreditation model (Table 1). Key dates include the approval of the Core Competencies in 1999, the launch of the Outcomes Project in 2001, and the transition of the first phase of accredited specialties to the ACGME's Next Accreditation System in July 2013 (Batalden et al. 2002; Nasca et al. 2012).

Table 1: Key Dates in Educational Milestones History

Dates	Structure
1999	The six Core Competencies endorsed by ACGME and American Board of Medical Specialties (ABMS)
2001	The Outcome Project formally launched
2009	ACGME approves structure of NAS, including inclusion of the Milestones
2013	First seven specialties implement NAS, including Milestones reporting
2014	Remaining accredited specialties and subspecialties implement NAS, including Milestones reporting
2015	All specialties and subspecialties begin to report Milestones data
2018	Work begins on Milestones revisions, called Milestones 2.0

Competency-based medical education (CBME) serves as the foundation for the ACGME's accreditation model, which is also grounded in a continuous quality improvement and innovation philosophy (Nasca et al. 2012; Weiss, Bagian, and Nasca 2013). Before examining the role of the Milestones in assessment and programmatic improvement, it is useful to summarize the history of CBME.

Overview: Competency-Based Medical Education (CBME)

Competency-based educational models are not new. In other fields, this is often called competency-based education and training (CBET), a term transformed to CBME in medicine. What is CBET? As Sullivan notes (1995):

> "In a traditional educational system, the unit of progression is *time* and it is *teacher-centered.* In a CBET system, the unit of progression is *mastery* of specific knowledge and skills and is *learner-centered*."

The earliest conception of competency-based training arose in the United States during the 1920s as educational reform became linked to industrial and business models of work that centered on clear specification of outcomes and the associated knowledge and skills needed. However, the more recent conception of CBET had much of its genesis in the teacher education reform movement of the 1960s (Elam 1971).

This interest was spurred by a US Office of Education National Center for Education Research grant program. In 1968, 10 universities developed and implemented new teacher training models that focused on student achievement (outcomes). Carraccio and colleagues noted that some sectors in medical education explored competency-based models in the 1970s. Elam laid down a series of principles and characteristics of CBET in 1971 (Table 2).

Table 2: Principles and Characteristics of Competency-Based Educational (CBE) Models (Elam 1971)

Principles	Characteristics
1. Competencies are role-derived (e.g., physician), specified in behavioral terms, and made public 2. Assessment criteria are competency-based and specify what constitutes mastery level of achievement 3. Assessment requires performance as the prime evidence, but takes knowledge into account 4. Individual learners progress at rates dependent on demonstrated competence 5. The instructional program facilitates development and evaluation of the specific competencies	1. Learning is individualized 2. Feedback to the learner is essential 3. Emphasis is more on the exit criteria (i.e., outcomes) than on the admission criteria (i.e., selection) 4. CBE requires a systems approach to manage a training program 5. Training is modularized 6. Both the learner and the program have accountability

From these beginnings, interest within medical education began to grow (Sullivan 1995). Competency-based models for medical education were soon promoted for wide use by McGaghie and colleagues as part of a report to the World Health Organization in 1978. In that report, the authors defined CBME as:

> "The intended output of a competency-based programme is a health professional who can practise medicine at a defined level of proficiency, in accord with local conditions, to meet local needs." (McGaghie and Lipson 1978)

A group of international educators worked to "modernize" the definition of CBME and lay out the theoretical rationale for a CBME system. This group defined CBME as: (McGaghie and Lipson 1978)

> "an *outcomes-based* approach to the design, implementation, assessment and evaluation of a medical education program using an *organizing framework* of competencies."

Put simply, under CBME, graduation requirements and curricula would be based on standardized outcomes, while learning exercises and formative feedback would be personalized (Achike, Lakhan, and Yakub 2019). Carraccio and colleagues (2002) compared the elements between the structure/process-based educational approach and the outcomes-based approach (Table 3).

While momentum was building for the principles and promises of CBME, there was also consensus that wide-spread acceptance would depend on addressing questions about:

- developing conceptual frameworks and language around CBME that would become well established and widely understood (Englander et al. 2017; Ferguson et al. 2017; Frank et al. 2010)
- designing learning outcomes, and with them, frameworks for assessment and evaluation (Gordon et al. 2017)
- preparing faculty member to apply CBME principles in the learning environment (Tannenbaum et al. 2020)

- developing evidence that CBME produces better practitioners than the conventional approach (Ferguson et al. 2017; Whitcomb 2016)

Table 3: Structure/Process-Based versus Competency-Based Programs

Educational Program Approach		
Variable	**Structure/Process**	**Competency-Based**
Driving force for curriculum	Content-knowledge acquisition	Outcome-knowledge application
Driving force for process	Teacher	Learner
Path of learning	Hierarchical (Teacher→Student)	Non-hierarchical (Teacher↔Student)
Responsibility for content	Teacher	Student and Teacher
Goal of educational encounter	Knowledge acquisition	Knowledge application
Typical assessment tool	Single measure focused	Multiple measures
Assessment tool	Proxy	Authentic (mimics real tasks of profession)
Setting for evaluation	Removed (gestalt)	"In the trenches" (direct observation)
Evaluation	Norm-referenced	Criterion-referenced
Timing of assessment	Emphasis on summative	Emphasis on formative
Program completion	Fixed time	Variable time

Adapted from Carraccio, et al. 2002.

Ongoing work is being done in response to those challenges. Englander and colleagues (2017) published a glossary of key terms and schematics depicting the relationships between key concepts such as "competency," "entrustable professional activities," and "milestones." Similarly, Van Melle and colleagues (2019) outlined five core components for CBME, how practice should be individualized and organized, the principles of good practice, and a core conceptual framework to justify them (Table 4). They derived this approach through Dephi method feedback mechanisms during the design of institution-wide implementation of CBME at Queen's University.

A distinguishing feature of CBME is that learners could progress through the educational process at different rates: the most capable and talented individuals would be able to make career transitions earlier, while others would require more time (to a limit) to attain a sufficient level of knowledge, skills, and attitudes to enter unsupervised practice. It is important to note that experience and time still matter in a CBME program, but time should not be treated as an intervention; rather, as a resource that should be used wisely and effectively. No one would argue that a certain quantity of experience is **unimportant** (Ten Cate 2014). Equally important are real system constraints in the United States that translate into the reality that the vast majority of graduate medical education (GME) programs would work in "hybrid models" of CBME – using competency-based educational principles in the context of fixed years of an educational program. A second key feature is the increased emphasis on assessment, especially ongoing, longitudinal assessment that enables faculty members to determine more accurately the developmental progress of the learner, as well as to help the learner through frequent feedback, coaching, and adjustments to learning plans (Englander et al. 2017; Ferguson et al. 2017; Holmboe et al. 2010; Kogan and Holmboe 2013). This is consistent with Anders Ericsson's work in expertise and deliberate practice, which demonstrates the need to tailor the educational experience to continually

challenge the learner with experiences that are neither too easy nor overwhelming (too difficult) (Ericsson 2007). Recent scholarship has borne out that frequent, actionable feedback about observable behaviors enable struggling residents to make improvements (Bonnema and Spencer 2012; Ross et al. 2018).

Table 4: Van Melle et al.'s Core Components of CBME: An Organizing Framework

Core Components				
Outcome competencies are required for practice and are clearly articulated	Competencies and their developmental markers are sequenced progressively	Learning experiences facilitate the developmental acquisition of competencies	Teaching practices promote the developmental acquisition of competencies	Assessment practices support and document the developmental acquisition of competencies
Practice: What the Core Component Should Look Like in Practice				
Required outcome competencies are based on a profile of graduate and or practice-based abilities	Competencies are organized in a way that leads to a logical developmental sequence across the continuum of medical education or practice	Learning takes in settings that model practice, is flexible enough to accommodate variation in individual learner needs, and is self-directed	Teaching is individualized to the learner, based on abilities required to progress to the next stage of learning	Learner progression is based on a systematic approach to decision making, including standards, data collection, interpretation, observation, and feedback
Principle: How the Core Component is Supposed to Work in Practice				
Specifications of learning outcomes promotes focus and accountability	A sequential path supports the development of expertise	Learning through real-life experiences facilitates membership into the practice community and development of competencies	Development of competence is stimulated when learners are supported to learn at their own pace and stage	Programmatic assessment systems allow for valid and reliable decision making
Conceptual Frameworks: Why the Core Component Should Work According to Theories, Models, or Best Practices				
• Social accountability • Outcome-based education • Backwards design • Job task analysis	• Expertise theory • Entrustable professional activities • Surface and deep approaches to learning • Mastery learning	• Situated learning • Deliberate practice • Workplace based learning • Professional identity formation	• Zone of proximal development • Constructive friction • Learner-centered apprenticeship • Coaching theory • Growth mindset	• Programmatic assessment • Formative assessment • Learning analytics

Source: Van Melle et al. 2019.

While defining the "competencies" was an important and necessary step, operationalizing and implementing them in practice prior to the Milestones proved to be challenging. Program directors and faculty members struggled since the launch of the Outcome Project

to understand what the Competencies meant and, more importantly, what they should "look like" in practice. This lack of shared understanding (i.e., shared mental models) hampered curricular changes, as well as development and evolution of better assessment methods. The challenges to operationalizing the Competencies was not restricted to the United States, and during the last 18 years several notable advancements have emerged in an effort to enable more effective implementation of CBME.

Carraccio and colleagues (2002) described a four-step process for implementing CBME: 1) identification of the competencies (in the United States the six ACGME/ABMS Core Competencies); 2) determination of competency components and performance levels (e.g., benchmarks and milestones); 3) competency assessment; and 4) overall evaluation of the process. Similarly, Crawford and colleagues (2020) noted that individual programs would need to gain acceptance of their faculty members for CBME principles, offer faculty training in implementing CBME, and develop systems to assess trainee performance. Faculty members would need to develop skills in delivering timely and meaningful feedback to learners, and learners would need to assume "ownership" of their learning and familiarity with CMBE.

The consensus in current scholarship adds that the adoption of CBME practices increases when programs provide opportunities for stakeholder engagement and adaptation throughout the process. Adoption will take root in an organization when it is built upon a sound theory of what is to be accomplished, a clear connection between proposed practices and goals, and frequent opportunity for feedback, and course correction (Hall et al. 2019; Hamza, Ross, and Oandasan 2020; Oandasan et al. 2020). Hall et al. (2019), describe the initial identification of outcomes and design of assessment as a "sprint," while the long-term stakeholder engagement, learner buy-in, frequent evaluation, and modifications is the "marathon." In moving from implementation to adoption, Hall's program incorporated three-month and six-month reviews to ensure "fidelity" to the conceptual plans, and to enable faculty member and learner involvement.

Caverzagie and collaborators (2017) noted that buy-in and sharing of concepts would need to happen beyond individual programs. Wide-spread adoption would depend on aligning regulatory bodies around concepts of CBME; ensuring cooperation from programs, training locations, and health systems; and establishing methods of mutual accountability among the GME system and its stakeholders. Examples of such self-regulatory adoptions include the ACGME Milestones and community created entrustable professional activities (EPAs). These concepts approach competence as a developmental process and rely heavily on positivist behavioral theory.

Since adoption, the Milestones have generated more than 350 scholarly publications. These papers have described, among other things, the challenges and advantages programs and residents/fellows experience in operationalizing and implementing the Milestones (Sangha and Hamstra 2020). One of the guiding principles of the Milestones project was the recognition that revision would be both necessary and desirable (Edgar, Roberts, and Holmboe 2018). It was not long after their initial use that four specific Competencies (interpersonal and communication skills, practice-based learning and improvement professionalism, and systems-based practice) were analyzed on how the milestones in these areas were being operationalized across specialties. This systematic research evaluated for redundancy across Competencies, how the subcompetencies and associated Milestones were conceptualized within and across specialties, and where important common themes existed. Subsequently milestones in these four Competency domains were

streamlined, or “harmonized” (Edgar et al. 2018). This harmonizing effort foreshadowed a more substantive revision called Milestones 2.0. Several specialties have already developed new Milestones using the Milestones 2.0 process. The Milestones continue to be an essential component of the the ACGME’s accreditation model, and this guidebook hopefully provides helpful information and direction in most effectively using the Competencies and the Milestones.

Milestones

Milestones are simply a significant point in development. They can enable the learner and the program to determine individual trajectories of professional development in narrative terms.

What Are Milestones?

In general terms, a milestone is simply a significant point in development. The Milestones in GME provide narrative descriptors of the Competencies and subcompetencies along a developmental continuum with varying degrees of granularity. Simply stated, the Milestones describe performance levels residents and fellows are expected to demonstrate for skills, knowledge, and behaviors in the six Core Competency domains. They lay out a framework of observable behaviors and other attributes associated with a resident's or fellow's development as a physician.

It is essential to recognize that milestones, based on the concept of stages of professional development, are designed to be criterion-based and agnostic to the actual PGY level of the resident or fellow. Programs should judge each resident or fellow based on the actual level of performance as described in the Milestones, not in relation to peers or others. However, Figure 1a provides some guidance about where a resident of fellow should be developmentally on the Milestone levels during the educational program. Figure 1b provide a description of the general anatomy of a milestone.

The Milestones describe the learning trajectory within a subcompetency that takes a resident or fellow from a novice in the specialty or subspecialty, to a proficient resident or fellow, or resident/fellow expert. Milestones are different from many other assessments in that there is an opportunity for the learner to demonstrate the attainment of aspirational levels of the subcompetency, and just as importantly allows for a shared understanding of the expectations for the learner and the members of the faculty. The Milestones can provide a framework for all GME programs that allows for some assurance that graduating residents and fellows across the US have attained a high level of competence.

It is also important to recognize what the Milestones are not. First and foremost, they do not describe or represent the totality or a complete description of a clinical discipline. They represent the important *core* of a discipline, meaning programs will need to use good judgment to fill in the gaps in curriculum and assessment. Second, it is essential that the Milestones are not thought of as curriculum in and of themselves, but rather that they should guide a thoughtful analysis of curriculum to identify strengths and gaps. Even for those specialties that developed more general subcompetencies, there was an understanding that the Milestones would not cover all areas essential to the unsupervised practice of medicine. Third, they are not tools designed to negatively affect program accreditation. The Milestones are intended for formative purposes to help learners, programs, and the Review Committees improve educational, assessment, and accreditation processes.

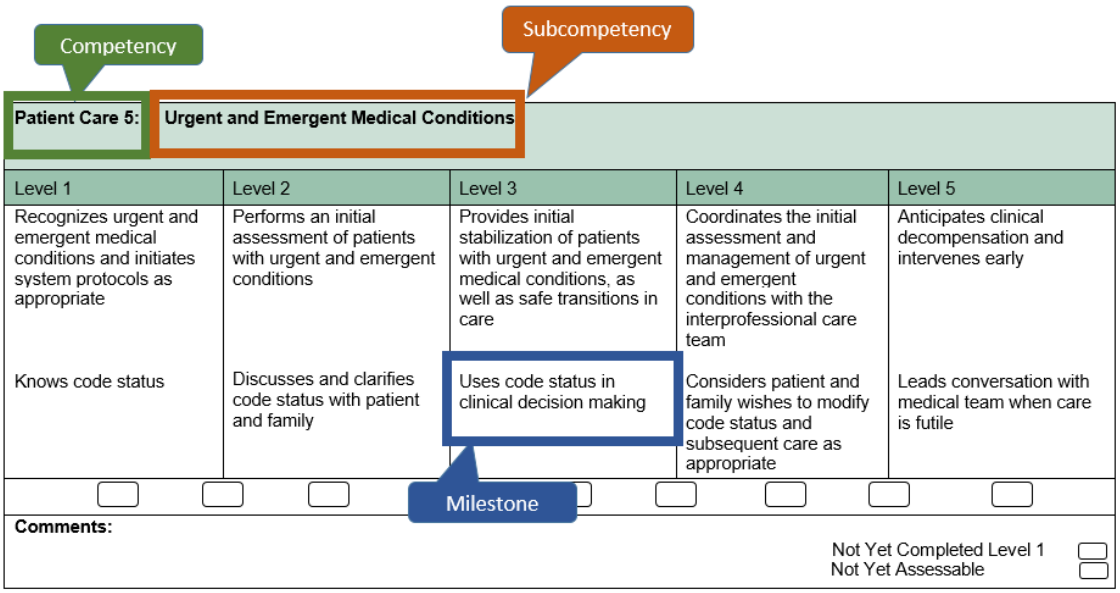
The *entire* Milestones document (set) used for reporting to the ACGME was also never intended to serve as a regular assessment tool, especially for short rotations (e.g., two to eight weeks in duration). The Milestones, and specifically the subcompetencies, do not contain enough detail or levels of performance on a developmental trajectory to facilitate an

accurate determination of the knowledge, skills, or abilities of an individual learner over a short period of time. In addition, the Milestones must not be used as the only set of assessment tools. Instead, the Milestones should inform the use and development of assessment tools aligned with the curricular goals and tasks. As stated previously, the Milestones are not inclusive of all areas of competency, and to limit the assessments to the Milestones would indicate that regular assessment is not occurring in the many other areas of learning.

Figure 1a: General Description of Milestone Levels Related to Stage of Education

Competency: Subcompetency				
Level 1	Level 2	Level 3	Level 4	Level 5
Novice Resident/Fellow Brand new to the specialty	Advanced Beginner Resident/Fellow Performs some tasks with limited autonomy	Competent Resident/Fellow Performs common tasks with autonomy	Proficient Resident/Fellow Target for graduation (not a requirement)	Resident/Fellow Expert Exceeds their peers
☐	☐	☐	☐	☐ ☐ ☐ ☐

Figure 1b: Example of the Basic Anatomy of a Milestone



How Were the Milestones Developed?

The process of Milestones development was unique for each specialty. Early development of the Milestones began with internal medicine in 2007. The American Board of Internal Medicine began working on the project very soon after the idea was first conceptualized. The ACGME began to formally bring specialties together in 2009 to start the process and determine the best course for development. By 2011, the formation of a Work Group for each of the core specialties was fully developed. That same year, the decision was made to

include five levels within the Milestones, guided by the Dreyfus Model of expertise development (Batalden et al. 2002). It was determined that Level 4 was to be considered the graduation target (*not a requirement*) and Level 5 would be for aspirational milestones. (See Figure 1a for an explanation of each level). Specialties that had already started the process were allowed to continue as they had been (i.e., fewer levels, levels with different descriptions, different graduation targets). Several changes have been made for Milestones 2.0. There are changes to the Work Groups that develop the content, as well as to the structure and format of the Milestones. Finally, there is more harmonization across the non-patient care and medical knowledge Milestones.

Work Group Formation

Each Work Group was composed of representatives, including: an appointed member(s) of the relevant ACGME Review Committee; the ABMS, through the individual certification boards; the American Osteopathic Association (AOA); and relevant program directors' groups. Each Work Group also included up to five members selected through a Call for Volunteers, at least one resident and/or fellow, and a public member. Each group was quite diverse and included representation of various sizes and types of programs (e.g., academic medical centers, rural hospitals, military hospital), subspecialty representation (e.g., neurologic surgery had representation for each of the eight primary subspecialties), and time in practice (e.g., junior and senior faculty members, program directors). For specialties into which medical school graduates enter directly (e.g., internal medicine, surgery), representatives from the Association of American Medical Colleges (AAMC) and American Association of Colleges of Osteopathic Medicine (AACOM) were included to ensure a more realistic expectation of incoming graduates.

Harmonized Milestones

A set of Harmonized Milestones was developed for the Core Competencies of interpersonal and communication skills, practice-based learning and improvement, professionalism, and systems-based practice (Edgar, Roberts, and Holmboe 2018). These Milestones were developed by four interdisciplinary, interprofessional groups and distributed for public comment. The intent was to have a common set of subcompetencies that allow each specialty to tailor the language to fit its distinct needs. For example, in the subcompetency of Patient- and Family-Centered Communication, the specific outcomes for internal medicine, surgery, and pathology vary based on the needs of the specialty.

Meeting Structure

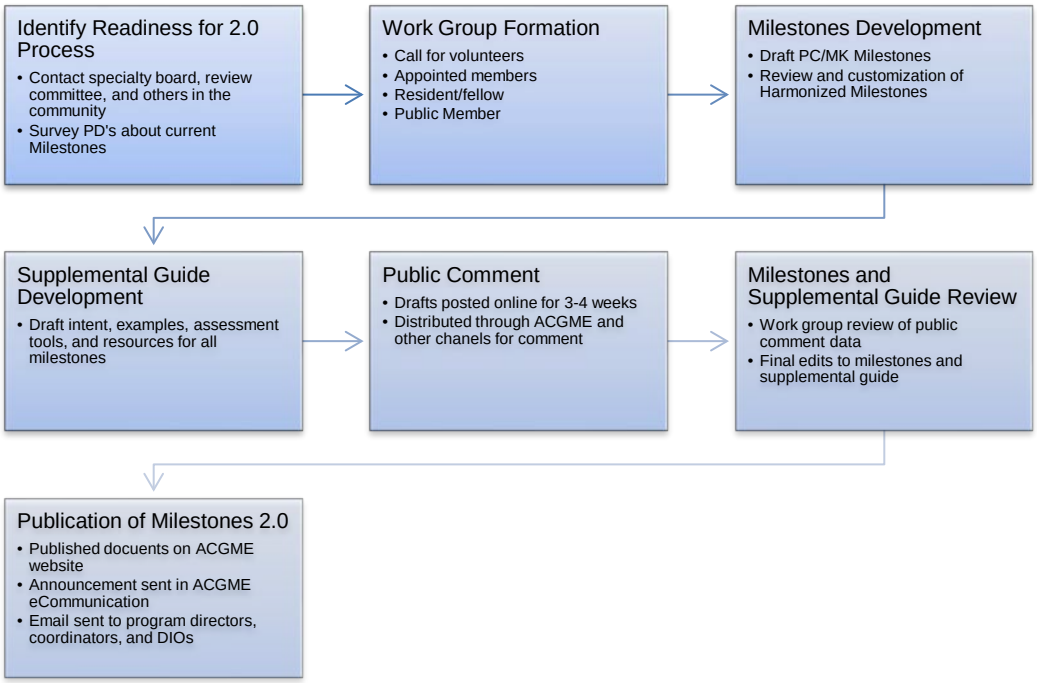
Each Work Group met two or three times to complete the process, which included a review of published documents, including the Program Requirements, certification blueprints, competency statements, shared curricula, and other literature. Each group also reviewed national data that had been reported to ACGME and results from a program director survey regarding the Milestones. Before identifying the subcompetencies, groups created a shared mental model around the educational frameworks used to develop the Milestones. These elements were taken into consideration while selecting the subcompetencies for Milestones 2.0. The discussion of what knowledge, skills, and attitudes would be most important was enthusiastic and complete. In many cases, the groups were able to select the most important topics for patient care and medical knowledge within a few hours. In some cases, the decision regarding which subcompetencies were most important took more than one full meeting due to the need to dissect the specialty and identify what is truly considered core, and the work of development started later.

Supplemental Guide

After several rounds of editing, a Supplemental Guide was created for each Set of Milestones. The Supplemental Guide serves as a companion document that describes the intent of each subcompetency, provides concrete examples, identifies potential assessment models, and offers notes and resources for faculty members and learners alike. The Supplemental Guide is intended to help programs understand the subcompetency, and can help the CCC form its own shared mental model for local implementation. More information on the Supplemental Guide is provided later in this guidebook.

After the Milestones and Supplemental Guide were drafted, they were made available for public comment on the ACGME website. Emails were sent to the specialty program directors and coordinators, and to the designated institutional officials (DIOs), with links to the drafted Milestones and Supplemental Guides. Those who received the emails were asked to share the information with the faculty members, residents, and fellows. Program director organizations were also asked to share information through their channels (i.e., listserv, emails). Those responding to the surveys were asked about the Milestones and the Supplemental Guide. The Work Group used the outcomes of the survey and the feedback received to edit and finalize the documents. Some specialties repurposed drafted Milestone sets that had been considered either duplicative or too elementary and published them in an appendix that could be used as a remediation or learning tool; these are sometimes referred to as “non-reportable Milestones.”

Figure 2. Milestones 2.0 Development Process



Why Milestones?

First and foremost, the Milestones are designed to help all residencies and fellowships produce highly competent physicians to meet the 21st century health and health care needs of the public. Second, as noted above, programs have struggled to operationalize the six Core Competencies since their introduction in 1999 (Batalden et al. 2002). The Milestones, along with the related concept of entrustable professional activities (EPAs), were developed to provide descriptive language that can facilitate a deeper, shared understanding among programs regarding the competency outcomes of interest within and across disciplines. The Milestones also enable the movement away from an overreliance on high stakes medical knowledge testing and use of numeric rating scales on evaluation forms, which faculty members have historically found very difficult to use effectively. Third, the Milestones can serve as a guide and "item bank" to create more meaningful assessments. Fourth, as learners' gaps are identified, there is the ability to provide individualized coaching to help them progress to the next level. Finally, the Milestones provide a critical framework for CCC deliberations and judgments.

The Milestones play a number of important roles depending on the constituent or stakeholder. Table 5 provides an overview of the purposes and functions of the Milestones related to each key stakeholder (Holmboe et al. 2015).

There are several key aspects to highlight about the use of the Milestones. First, as noted above, the Milestones that are reported to the ACGME were not designed to be used as evaluation forms for specific rotations or experiences, especially short rotations less than three months in length. The Milestones are designed to guide a synthetic judgment of progress twice a year. However, utilizing language from the Milestones may be helpful as part of a mapping exercise to determine which Competencies are best covered in specific rotations and curricular experiences. Second, the Milestones can also be used for guided self-assessment and reflection by a resident/fellow in preparation for feedback sessions and in creating individual learning plans. Residents and fellows should also use the Milestones self-assessment in a guided feedback conversation with a faculty advisor, mentor, or program director. Residents and fellows should not judge themselves on the Milestones in isolation. As highlighted in the Feedback section below, Milestones feedback is most effective when it is performed in dialogue between a learner and faculty advisor. Third, the Milestones can be useful in faculty development. They can help faculty members recognize their performance expectations of learners, more explicitly assess the trajectory of skill progression in their specialty, and discern how best to assess a learner's performance. Finally, it is imperative that programs remember that the Milestones are not inclusive of the broader curriculum, and that limiting assessments to the Milestones could leave many topics without proper and essential assessment and evaluation.

Table 5: The Purpose and Function of Milestones

Constituency or Stakeholder	Purpose/Function
Residents and Fellows	• Provide a descriptive roadmap for education and training • Increase transparency of performance requirements • Encourage informed self-assessment and self-directed learning • Facilitate better feedback to learner • Encourage self-directed feedback-seeking behaviors
Residency and Fellowship Programs	• Guide curriculum and assessment tool development • Provide meaningful framework for CCC (e.g., help create shared mental model) • Provide more explicit expectations of residents and fellows • Support better systems of assessment • Enhance opportunity for early identification of under-performers • Enhance opportunity to identify advanced learners to offer them innovative educational opportunities
ACGME	• Accreditation – enable continuous improvement of programs and lengthening of site visit cycles • Public Accountability – report at an aggregated national level on Competency outcomes • Community of practice for evaluation and research, with focus on continuous improvement
Certification Boards	• Enable research to improve certification processes

Implementing and Using Milestones Effectively

While there is still much to learn, early research combined with solid educational theory does provide some useful guidance for programs.

Involving Residents and Fellows

Summary – Practical Tips

1. *Share and discuss the pertinent Milestones Set with residents/fellows at the beginning of the program. This helps them to gain a shared understanding of the goals of the program and the Milestones.*
2. *Share the* Milestones Guidebook for Residents and Fellows *at the beginning of the program.*
3. *Have residents/fellows complete individualized learning plans, using the Milestones as an important guide (ACGME 2020).*
4. *Consider having residents/fellows complete a self-assessment of the Milestones that they can compare and contrast, with a trusted advisor, to the Milestone judgments of the CCC every six months.*
5. *Enable residents/fellows to seek out assessment (i.e., self-directed assessment seeking), especially direct observation, from faculty members.*

Residents and fellows are primary stakeholders in the Milestones system. Education is always co-created and co-produced between teacher and learner (Bate and Robert 2006; Freire and Sangiorgi 2014; Fuchs 1968, 12; Sabadossa and Batalden 2014; Normann 2001; Ostrom 1996; Garn et al. 1976). The recognition of this need for active engagement seems to invite new attention in health professional development as the shared work of teacher and learner. Learners in a CBME system must be active agents co-guiding both the curricular experiences and the assessment activities.

Viewing medical education in these ways might invite consideration of the highly trained learner as a critical input into the health care system, rather than as an "output" of an isolated educational process (Sabadossa and Batalden 2014; Normann 2001). Sabadossa and Batalden (2014) described the importance of co-production in clinical care. They noted that such co-production requires "capabilities of the patient, family, and clinical professionals for the 'coproduction' of good care" (Sabadossa and Batalden 2014). Wagner, et al. (1996) described the importance of "activated patients" for the development of good care. Medical education-as-service is no different (Freire and Sangiorgi 2014).

What does it mean for residents and fellows to be "active agents" in their own learning and assessment? Learners must learn to be self-directed in seeking assessment and feedback (Molloy and Boud 2013), and thus residents and fellows should ideally:

1. be introduced to the content and purpose of the Milestones at the very beginning of the program through dialogue, with that dialogue continuing so as to deepen their understanding on an ongoing basis; simply e-mailing or providing a hard copy of the Milestones without explanation and discussion is insufficient;
2. read the *Milestones Guidebook for Residents and Fellows*;

3. direct and perform some of their own assessments, such as by seeking out direct observation, auditing medical records and/or case logs around quality and safety performance, creating an evidence-based medicine clinical question log, etc.;
4. perform a self-assessment in conjunction with the CCC report to help them identify areas of agreement (concordance) and disagreement (discordance); self-assessment in isolation is not effective, but self-assessment combined with external data (e.g., the CCC Milestones report) is a valuable and impactful activity (Sargeant et al. 2015);
5. develop personal learning plans that they revisit and revise at least twice a year;
6. actively seek out assessment and feedback on an ongoing basis; and,
7. provide systematic feedback to the program on their experience with the Milestones.

Faculty Members

Summary – Practical Tips

1. *Share and discuss the pertinent Milestones Set with faculty members as a group at the beginning of the academic year (at a minimum). This helps faculty members develop and use a shared understanding of the goals of the Milestones.*
2. *Observe, observe, observe! Faculty members' observation of key competencies is essential to effective feedback, coaching, and professional development of residents/fellows.*
3. *Embed observation in "what faculty members do" – clinic precepting, procedures, bedside rounds, discharge planning, joining part of an admission, and so on.*
4. *Participate in faculty development around the Milestones, assessment and observation, and feedback as core educator skills.*
5. *Help faculty members understand where their assessments map onto the pertinent Milestones related to their role in the program.*

Faculty members represent the essential educational core of any graduate medical education program. The conception of faculty members is also expanding to include others on the interprofessional health care team beyond physicians. Faculty members need, at a minimum, a basic understanding of the structure and purpose of the Milestones. However, not all faculty members necessarily need a deep understanding of all the subcompetencies and milestones. Faculty members "in the trenches" (e.g., who serve as preceptors and attendings) should focus on those subcompetencies and milestones most pertinent to their role, curricular activity, and site of education and training. This may mean that the program will need to revise the nature of the evaluation forms faculty members complete (more below). Assessment is a skill that needs ongoing practice and feedback. This is especially true of direct observation of clinical skills. The important implications for faculty members are that they should:

1. familiarize themselves with the overall Milestones;
2. focus on those subcompetencies and milestones pertinent to their attending or assessment role;
3. participate in faculty development, especially around assessment and feedback;
4. make a commitment to improving and refining their assessment skills;
5. provide feedback to the program on how to improve assessment approaches and feedback;

6. provide meaningful narrative assessment as part of direct observations and evaluation forms–it is this information that is often most helpful to program directors and CCCs; and,
7. provide ongoing feedback to learners, which is essential for good coaching and professional growth.

Program Leadership

Summary – Practical Tips

1. *Create a shared mental model of the Milestones using the Supplemental Guide and other resources.*
2. *Share and discuss the pertinent Milestone set with faculty members as a group at the beginning of the academic year (at a minimum). This helps faculty members develop and use a shared understanding of the goals of the Milestones.*
3. *Empower and facilitate direct observation of residents/fellows by faculty members. Faculty observation of key competencies is essential to effective feedback, coaching, and professional development.*
4. *Provide longitudinal faculty development around the Milestones, assessment, observation, and feedback. These are difficult skills, and single, one-time workshops are helpful, but insufficient. Assessment instruments are only as effective as the person using them.*
5. *Build "small aliquots" (e.g., 15-30 minutes) of faculty development into existing structures, such as section and department meetings, grand rounds, morning reports, noon conferences, and CCC meetings. Use the "practice makes perfect" principle through continued dialogue around the Milestones. This helps to deepen shared understanding.*
6. *Map the curriculum and assessment program against the pertinent milestones. This will help to identify curricular gaps and areas for opportunity, and ensure the most effective combination of assessments.*

The transition into the NAS and use of the Milestones has substantially affected the role and nature of work for program directors and other program leaders. Program directors represent the essential hub of the program. Institutions should actively support professional development for program leaders. The program director, associate program director, and program coordinator roles are vitally important to the overall medical education enterprise, with profound influences on learner and patient outcomes. As such, program leaders need ongoing professional development around the key roles and tasks now required of them. Key tasks for program leadership include:

1. conducting a crosswalk of the curriculum with the specialty Milestones and Supplemental Guide to ensure that learners have sufficient experience. For example, review the educational objectives and purpose of a rotation, then map the essential subcompetencies with the objectives, purpose, and goals of the rotation. For example:

Milestones	Curriculum Mapping (which rotation objectives meet this Milestone)	Assessment Tool/Method
Patient Care 1	Outpatient rotations	Direct observation tool; multisource feedback
Medical Knowledge 2	Inpatient rotations	Assessment of case-based discussion; journal club participation; assessment of presentation

2. developing a program of assessment that aligns with the Milestones and functions as an integrated, holistic package; assessment activities should tightly align with the actual education and/or training activity;
3. identifying and address gaps in assessment strategies to ensure meaningful and authentic Milestones judgments;
4. conducting ongoing program evaluation to assess what is working, for whom, in what circumstances, and why; do not be afraid to discontinue things that are not working – think of the Milestones as part of a continuous quality improvement process; logic models, the Kirkpatrick hierarchy, and other approaches to program evaluation can be very helpful; if the program has access to an education department or expertise, program leaders are encouraged to sit down with these individuals to explore what the best program evaluation strategy would be for their programs;
5. providing ongoing faculty development, especially around assessment; while workshops are clearly helpful, they are not enough, and program leaders should think of ways the program can build "small aliquots" of faculty development into section or department meetings, grand rounds, CCC meetings, etc.; taking just 15 minutes on a regular basis to review a few subcompetencies and their milestones, review and rate a short video tape performance, etc., can be very valuable;
6. building a team—program directors cannot do this alone And building a team that has deeper understanding of the Milestones and basic educational and assessment methods and theory is crucial; most specialties now have active program director associations or groups that provide excellent resources and training; it is equally important not to be afraid to reach across disciplinary boundaries; much good work is happening in some of specialties within institutions of which others in the same institutions are unaware—program directors should check with the institution's DIO and graduate medical education committee (GMEC) to learn what is happening in their local institution; and,
7. exploring the functionality of the electronic residency/fellowship management system with respect to linking items on assessment tools and methods to the Milestones to aid in curriculum review.

Assessment Program

As noted above, educational leaders need to build an assessment *program* (Schuwirth and Van der Vleuten 2011). No single assessment tool or method will be sufficient to judge all the Competencies necessary for 21st century practice. There is also no single "magic combination" – programs will potentially need to choose and develop a set of assessments that meet local needs and context. Basic common assessment methods are provided below as a simple guide, but this is not meant to be an exhaustive list.

The CCC is also a vital component of the assessment program and overall program system. Appendix B demonstrates a high performing assessment system. In conjunction with this

Milestones Guidebook, program directors and others are encouraged to review the *CCC Guidebook* available on the Resources page of the Milestones section of the ACGME website: https://www.acgme.org/What-We-Do/Accreditation/Milestones/Resources.

Additionally, for more information on assessment, review the new *Assessment Guidebook*, also available on the Resources page of the Milestones section of the ACGME website (link above).

Table 6: Common Assessment Methods for the Six Core Competencies

Core Competency	Common Assessment Methods
Patient Care	• Direct observation (live or video) • Rating scales/evaluation forms • Audit of clinical practice (e.g., quality performance measures) • Simulation (including standardized patients) • Case logs/registries
Medical Knowledge	• In-training examinations • Oral questioning methods (e.g., SNAPPS) • Direct observation (live or video) • Assessment of Reasoning Tool
Professionalism	• Multi-source feedback • Patient surveys (can be part of multi-source feedback) • Direct observation
Interpersonal and Communication Skills	• Multi-source feedback • Patient surveys (can be part of multi-source feedback) • Direct observation (live or video) • Simulation (including standardized patients)
Practice-based Learning and Improvement	• Audit of clinical practice (e.g., quality performance measures) • Evidence-based medicine logs • Case logs • Rating scales/evaluation forms • Reflective practice rubrics
Systems-based practice	• Quality improvement knowledge assessment test • Audit of clinical practice (e.g., quality performance measures) • Multi-source feedback • Rating scales/evaluation forms

Importance of Feedback

Feedback to residents and fellows is an essential and required activity of the Milestones assessment system. Research has clearly shown that feedback is one of the most effective educational tools faculty members and programs have to help residents and fellows learn and improve. The Milestones should be used to help residents and fellows develop action plans (i.e., individualized learning plans) and adjustments to their learning activities and curriculum. Feedback sessions should also be conducted in person. Research is clear that interpreting and understanding multi-source and multi-faceted performance data, as represented by the Milestones, should be facilitated and guided by a trusted advisor.

Five basic features of high-quality feedback are (Skeff and Stratos 2015):

1. **Timeliness**. Faculty members should always try to provide feedback in a timely fashion. The results of the CCC deliberations and Milestones determinations should also be shared in person with the resident or fellow soon after the meeting has occurred.
2. **Specificity**. The Milestones help to facilitate this criterion by providing descriptive narratives. Generalities (often called "minimal" feedback), such as "you're doing great," or, "should read more," etc., are not very helpful in promoting professional development, especially in the context of Milestones data. There may be a tendency to gloss over the high performing residents or fellows but remember that they will benefit from "stretch" goals.
3. **Balance reinforcing ("positive") and corrective ("negative") feedback**. It is important to include both in specific terms. An imbalance between too much reinforcing or conversely corrective feedback can undermine the effectiveness. The popular feedback sandwich (*positive-negative-positive*) is actually not very effective and not routinely recommended.
4. **Learner reaction and reflection**. It is very important to allow the resident or fellow to react and reflect on the feedback and Milestones data. Reaction and reflection help garner resident and fellow buy-in and development of action plans.
5. **Action plans**. Creating and executing an action plan after a Milestones review is critical to professional development and is often neglected in feedback. As Boud and Molloy (2013) argue, feedback hasn't occurred until the learner has actually attempted an action or change with the information. Feedback is more than just information giving and dissemination (Friedman et al. 2014).

Lessons Learned about the Milestones

ACGME Milestones staff members regularly attend program director and society meetings, and visit institutions. These encounters enable high-level conversations on the benefits and challenges of the Milestones and have helped to drive the changes in Milestones 2.0. Along with other more systematic and rigorous research, these conversations have provided clear signals and helped to guide next steps. In that spirit, Table 7 provides a topline summary.

Table 7: Perceived Benefits and Challenges of Milestones Implementation

Benefits	Challenges

• Milestones and CCC process can provide better feedback for residents and fellows
• Milestones system can catalyze feedback for residents and fellows (e.g., for many, this can be first time formal feedback given)
• Milestones provide useful language for assessment and feedback
• Milestones help faculty members develop shared mental model of competence
• Milestones have helped to identify curricular gaps
• Milestone mapping onto curricular activities has facilitated better assessment
• Milestones are facilitating earlier identification of residents and fellows in difficulty
• CCCs are a useful mechanism to facilitate working with residents and fellows in difficulty
• Milestones facilitate faculty development
• Milestones provide a continuous quality improvement philosophy of system
• The common framework of Milestones allows for more generalizability of medical education research on assessment in GME
	• Time and resources ("relative value units [RVUs] always win")
	  o Data entry burden
	• Synthesizing multiple assessments into a CCC developmental judgment
	• Misalignment of assessment forms and scales and Milestones judgments
	• Lack of assessment methods and tools
	• Use of Milestones as rotation evaluation form (problem of "cognitive load")
	• Need for faculty development
	• Assessment burden on faculty members
	• Increasingly short faculty attending periods (e.g., one to two weeks) in a number of specialties
	  o Insufficient faculty member exposure to properly perform assessment
	• Challenging to use a five-level Milestone rubric for one-year fellowships
	• Educational jargon and framing of language (select Milestones Sets)

Annotated Bibliography of Research

The ACGME Department of Research, Milestone Development, and Evaluation maintains a bibliography of research related to the Milestones and CCCs. The bibliography is updated approximately every six months and can be found at https://www.acgme.org/What-We-Do/Accreditation/Milestones/Research.

Use of Milestones by the ACGME

Milestones data is not shared with the Review Committees. The Review Committees are made aware of program compliance with submission of the data. Residents' and fellows' performance on the Milestones, aggregated at the national level, will become a source of specialty-specific data for the Review Committees to use in their continuous quality improvement efforts for facilitating improvements to program curricula and resident/fellow assessment. The critical concept is that the Milestones' primary purpose is to drive improvement in GME programs and enhance the resident and fellow educational experience. The Milestones will also be used by the ACGME to demonstrate accountability of the effectiveness of GME within ACGME-accredited programs in meeting the needs of the public over time.

As the transition to Milestones 2.0 continues, the ACGME will continue to learn through several mechanisms, including through its own research and evaluation activities, as well as collaborative research and evaluation with other stakeholders, through comments received through the Milestones mailbox (milestones@acgme.org), and ongoing outreach activities. The ACGME and ABMS will also work together to develop a revision process with the educational community and share learnings and research from this early phase. The exact date of implementation of "Version 2.0" of the Milestones for each specialty is still being determined – check the weekly ACGME *e-Communication* and Milestones page of the applicable specialty section of the ACGME website for updates. Additionally, when opportunities arise to volunteer for a Milestones 2.0 Work Group or comment on a draft, they will be posted on the Engagement page of the Milestones section of the website, at https://www.acgme.org/What-We-Do/Accreditation/Milestones/Engagement.

Data Security and Milestones

The ACGME is dedicated to protecting the data collected from programs and residents/fellows. There are four key components:

1. From a legal standpoint, the ACGME is subject to the Illinois state peer review statutes. These statutes are tracked very carefully and have successfully blocked discoverability of ACGME data.
2. The Review Committees will not review any identified individual resident or fellow Milestones data, but will instead view the data in aggregate, using the specialty and program as the unit of analyses for continuous quality improvement purposes.
3. The plan is to convert the resident/fellow identifier to the National Provider Identifier (NPI) to discontinue use of Social Security Numbers for this purpose.
4. The ACGME also uses state-of-the-art data security methods to ensure the safety of all data, including data related to the Milestones.

How will the ACGME Continue to Evaluate the Milestones?

Evaluation of the Milestones iteratively and longitudinally is essential in achieving the desired goals of the ACGME's accreditation model. Unlike traditional biomedical approaches to research, evaluation of the Milestones will require a predominantly practice-based, action research utilizing principles of complex interventions and program evaluation (Campbell et al. 2007; Medical Research Council 2014; Pawson 2013; Pawson and Tilley 1997; Rogers, 2011). Much has been learned since the initial implementation of the Milestones in 2013. Research on Milestones is collected and updated approximately every six months in a collated Milestones Bibliography, available on the Research page of the Milestones section

of the ACGME website, at https://www.acgme.org/What-We-Do/Accreditation/Milestones/Research.

One advantage of the Milestones, compared to some other assessment tools currently used by individual programs, is that assessment data is collected on thousands of residents and fellows, producing a sample that, over time, makes it possible to establish their reliability and validity on a national scale. This has enabled important validity research on a national scale. The Messick framework is a useful framework in understanding validity (Cook and Beckman 2006):

> **Content**: the assessment instrument items completely and appropriately represent the construct being assessed
> **Response process**: the relationship between the intended construct and the thought processes of subjects or observers (e.g., have the observers been trained?)
> **Internal structure**: acceptable reliability and factor structure of the assessment
> **Relations to other variables**: examining correlations with scores from another instrument assessing the same construct (e.g., medical knowledge, clinical skills)
> **Consequences (intended uses)**: how scores are used affects how the assessment instrument is used and how the data is interpreted

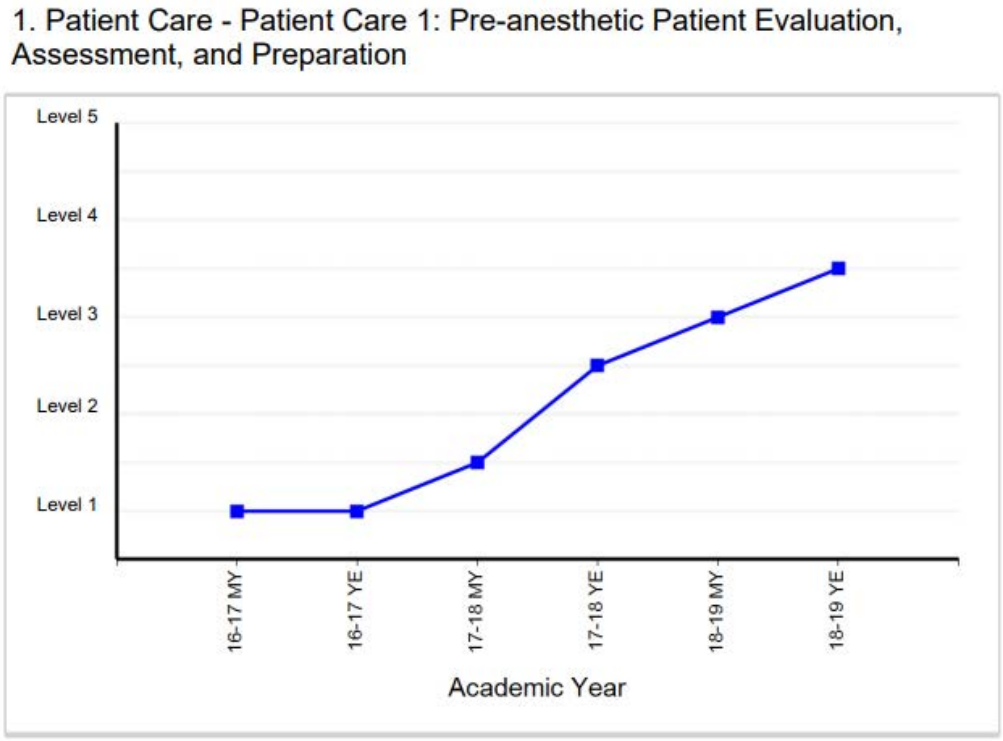
The important principle in validity frameworks is that validity is treated more as an argument that requires ongoing refinement and investigation. As noted above, the Milestones will need to be revised and refined over time, building from the “on-the-ground” experience of programs and rigorous research and evaluations.

Milestone Reports Available in the Accreditation Data System (ADS)

After the program director submits the Milestones evaluations twice each year, several reports can be downloaded. Available reports include individual resident/fellow reports, program reports, and specialty reports.

Resident Reports
The resident/fellow reports can be used as part of the resident/fellow semiannual evaluation. There is a space for signatures, should the program choose to use it. It is not required that programs print these reports; the ACGME does not require any further action after the Milestones data has been submitted. The individual detailed PDF documents of the reports will be available 10-14 days after the close of the reporting window. The examples below are from a third-year anesthesiology resident.

Report 1: Individual Milestone Trends
This report includes a graph showing the individual's progression for each subcompetency. Notice how the resident begins at Level 1 and steadily progresses to Level 3.5. The goal of the Milestones system is to support professional development, and these "growth curves" can help programs assess whether a resident or fellow is on the appropriate trajectory (see predictive probability values below).



Report 2: Individual Milestone Summary
This report provides a snapshot of the individual's most recent evaluation for each subcompetency. The example below shows that while the resident effectively communicates with patients and families, the resident could improve these skills with other professionals.

Interpersonal and Communication Skills

	Level 1 Not Yet Achieved	Level 1		Level 2		Level 3		Level 4		Level 5
a). Interpersonal Communication Skills 1: Communication with patients and families						●				
b). Interpersonal Communication Skills 2: Communication with other professionals				●						
c). Interpersonal Communication Skills 3: Team and leadership skills							●			

Report 3: Individual Milestone Evaluation
This report provides the text of the milestone level assigned for each subcompetency. When an individual's evaluation is between levels, the text for both levels are displayed, with the higher level test identifying that the resident has achieved certain, but not all of the requirements. In Patient Care 7, below, the resident is between Levels 4 and 5.

7 Patient Care	Patient Care 7: Acute, chronic, and cancer-related pain consultation and management

Dr. Name is between Level 3 and Level 4.

Formulates differential diagnoses of acute and chronic pain syndromes; identifies appropriate diagnostic evaluation.

Participates in complex procedures (e.g., thoracic ESI, medial branch blocks, radiofrequency procedures, sympathetic blocks) for alleviating acute, chronic, or cancer-related pain, under direct supervision.

Prescribes initial therapy for pain medication, and adjusts ongoing medication regimens with indirect supervision; uses ultrasound and fluoroscopy with direct supervision.

In addition, Dr. Name has achieved certain, but not all, elements of the competency level listed below:

Acts as consultant for acute pain management to junior residents and other health care providers with conditional independence.

Consults with non-anesthesiologist specialists regarding pain management as appropriate.

Recognizes treatment failures and obtains appropriate consultations, including with a pain medicine specialist.

Program and Specialty Reports

At the end of the academic year two additional reports are available in ADS. Both reports are box plots with one demonstrating the results at year end for the program, and the other a national report for the specialty. A key to understanding the box plots is included in the *Milestones National Report* published annually in the fall for the prior academic year. The *Milestones National Report* also includes other important data, including predictive probability values for evaluating if a resident is on track to graduate below Level 4 for a specific subcompetency. The *Milestones National Reports* can be found on the Research and Reports page of the Milestones section of the ACGME website: https://www.acgme.org/What-We-Do/Accreditation/Milestones/Research.

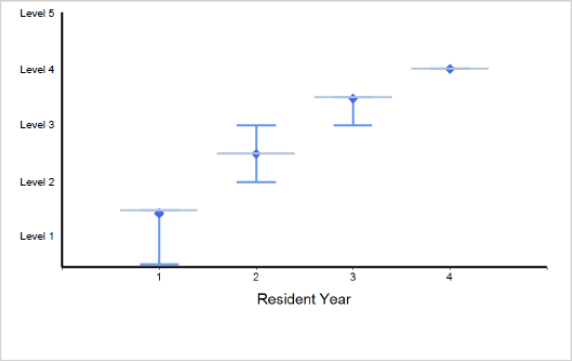
Program Report

Program Box Plot Report - Milestone Evaluation by Resident Year: Year-End 2018-2019

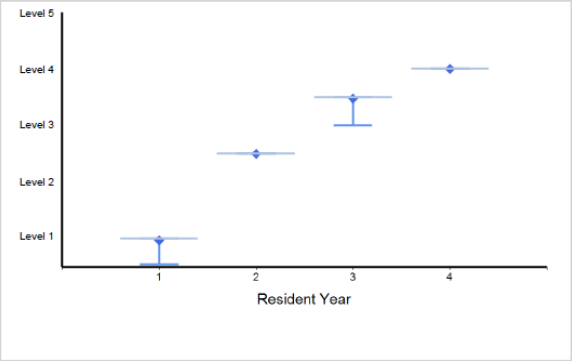
Program: **Program Name Program Code - Specialty**

Resident Year	1	2	3	4	Total Residents
# of Residents	XX	XX	XX	XX	XX

1. Patient Care - Patient Care 1: Pre-anesthetic Patient Evaluation, Assessment, and Preparation



2. Patient Care - Patient Care 2: Anesthetic Plan and Conduct



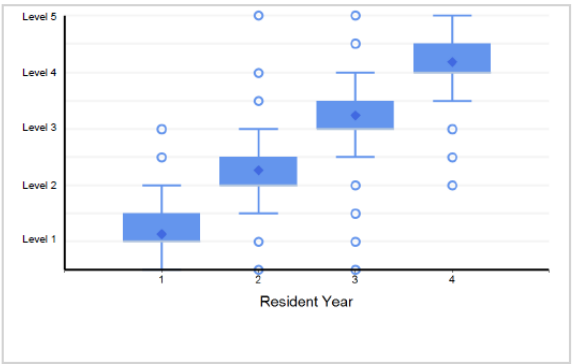
Specialty Report

Specialty Box Plot Report - Milestone Evaluation by Resident Year: Year-End 2018-2019

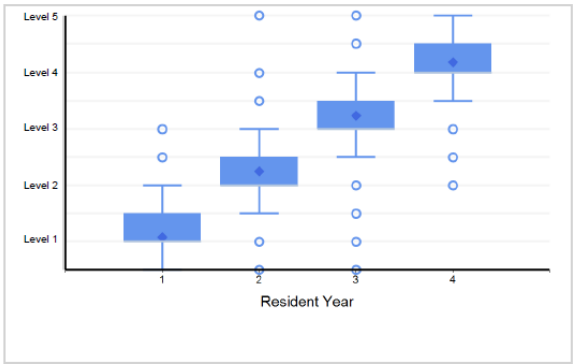
Specialty: **Specialty**

Resident Year	1	2	3	4	Total Residents
# of Residents	x,xxx	x,xxx	x,xxx	x,xxx	x,xxx

1. Patient Care - Patient Care 1: Pre-anesthetic Patient Evaluation, Assessment, and Preparation



2. Patient Care - Patient Care 2: Anesthetic Plan and Conduct



Predictive Probability Value (PPV) Tables

The ACGME began providing predictive probability value (PPV) tables with the *2019 Milestones National Report*. Program directors can now examine PPVs for program-level education and training, which are provided following the box plots in the report by specialty. PPVs are provided to help program directors identify residents/fellows who may be struggling to match normative national data during each six-month block of the educational program.

PPV tables provide the probability (in percentage terms) that a resident/fellow at or below a certain Milestone rating (Level) would not achieve Level 4 at time of graduation. In the example shown below, all PPVs for the Family Medicine Patient Care Subcompetency #03 that could be calculated as of June 2019 are included in the table. For example, a resident receiving a Milestone rating of 2.5 or lower at Milestones review occasion four, at the end of the PGY-2 has a 54.7 percent probability (based on national data) of not achieving Level 4 in this subcompetency by the end of the three-year family medicine residency.

Figure 3: PPV Matrix for the Patient Care Subcompetency #03 in Family Medicine: Partners with the patient, family, and community to improve health through disease prevention and health promotion

Subcompetency	Threshold	Yr1, Mid-Year	Yr1, Year-End	Yr2, Mid-Year	Yr2, Year-End	Yr3, Mid-Year
PC03						
Partners with the	≤ Level 5.0					
patient...	≤ Level 4.5					
	≤ Level 4.0					32.7
	≤ Level 3.5	31.9		32.4	33.4	40.6
	≤ Level 3.0	31.9	32.0	33.5	37.6	60.8
	≤ Level 2.5	32.1	32.8	37.5	54.7	67.4
	≤ Level 2.0	32.5	35.3	50.2	57.2	
	≤ Level 1.5	33.8	42.9	60.1	78.5	
	≤ Level 1.0	39.7	51.6			

The table in this example provides a matrix of all PPVs by Milestone rating threshold and Milestone review occasions for a single subcompetency for a single specialty. PPVs are provided to help identify residents/fellows within a program who may be struggling to match normative national data during each six-month block of the educational program. This can then be used to support decisions for remediation or individualized learning plans. The PPVs support the use of the Milestones as longitudinal assessment data to support professional development, feedback, coaching, and individualized learning plans.

Conclusions

The overarching goal of all GME programs is to produce graduates that can be entrusted to provide the highest quality of care for the benefit of the public they serve. It is important to remember that the principle driver for a shift to an outcomes-based educational model was the recognition both within and outside the medical education community that rapid changes in health care delivery and science necessitated concomitant changes in the medical education system. The Milestones, combined with CCCs, were developed to enable and accelerate the transformation to a competency-based system after a difficult early period of implementation. The success of the ACGME's current accreditation model and the Milestones will depend on an ongoing collaboration among the end users (i.e., programs, faculty members, and learners), regulators like the ACGME and the certification boards, Sponsoring Institutions and organizations, researchers, and policy makers.

References

ACGME Common Program Requirements. July 1, 2019. Accessed May 7, 2020: https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRResidency2019.pdf

Achike Francis I., Shaheen E. Lakhan, and Mohsin Yakub. 2019. "Competency-based Medical Education: Philosophy, What, How, Why, and the Challenges Therein." *Journal of Medical Education* 23 (1): 1-13.

Batalden, Paul, David Leach, Susan Swing, Hubert Dreyfus, and Stuart Dreyfus. 2002. "General Competencies And Accreditation In Graduate Medical Education." *Health Affairs* 21 (5): 103–11. https://doi.org/10.1377/hlthaff.21.5.103.

Bate, P., and G. Robert. 2006. "Experience-Based Design: from Redesigning the System around the Patient to Co-Designing Services with the Patient." *Quality and Safety in Health Care* 15 (5): 307–10. https://doi.org/10.1136/qshc.2005.016527.

Bonnema Rachel A., Abby L. Spencer. 2012. "Remediating residents: Determining when enough is enough." *Academic Internal Medicine Insight* 10(4): 6-7.

Campbell, Neil C., Elizabeth Murray, Janet Darbyshire, Jon Emery, Andrew Farmer, Frances Griffiths, Bruce Guthrie, Helen Lester, Phil Wilson, and Ann Louise Kinmonth. 2007. "Designing and Evaluating Complex Interventions to Improve Health Care." *BMJ* 334 (7591): 455–59. https://doi.org/10.1136/bmj.39108.379965.be.

Carraccio, Carol, Susan D. Wolfsthal, Robert Englander, Kevin Ferentz, and Christine Martin. 2002. "Shifting Paradigms: From Flexner to competencies." *Academic Medicine* 77 (5): 361–67. https://doi.org/10.1097/00001888-200205000-00003.

Caverzagie, Kelly J., Markku T. Nousiainen, Peter C. Ferguson, Olle Ten Cate, Shelley Ross, Kenneth A. Harris, Jamiu Busari, et al. 2017. "Overarching Challenges to the Implementation of Competency-Based Medical Education." *Medical Teacher* 39 (6): 588–93. https://doi.org/10.1080/0142159x.2017.1315075.

Cook, David A., and Thomas J. Beckman. 2006. "Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application." *The American Journal of Medicine* 119 (2). https://doi.org/10.1016/j.amjmed.2005.10.036.

Crawford, Lindsay, Nicholas Cofie, Laura Mcewen, Damon Dagnone, and Sean W. Taylor. 2020. "Perceptions and Barriers to Competency-Based Education in Canadian Postgraduate Medical Education." *Journal of Evaluation in Clinical Practice*. https://doi.org/10.1111/jep.13371.

Edgar, Laura, Sydney Roberts, and Eric S. Holmboe. 2018. "Milestones 2.0: A Step Forward." *Journal of Graduate Medical Education* 10 (3): 367–69. https://doi.org/10.4300/jgme-d-18-00372.1.

Edgar, Laura, Sydney Roberts, Nicholas A. Yaghmour, Andrea Leep Hunderfund, Stanley J. Hamstra, Lisa Conforti, and Eric S. Holmboe. 2018. "Competency Crosswalk: a Multispecialty Review of the Accreditation Council for Graduate Medical Education Milestones Across Four Competency Domains." *Academic Medicine* 93 (7): 1035–41. https://doi.org/10.1097/acm.0000000000002059.

Elam, Stanley. *Performance-Based Teacher Education: What is the State of the Art?* 1971. Washington: American Association of Colleges for Teacher Education. 1–36.

Englander, Robert, Jason R. Frank, Carol Carraccio, Jonathan Sherbino, Shelley Ross, and Linda Snell. 2017. "Toward a Shared Language for Competency-Based Medical Education." *Medical Teacher* 39 (6): 582–87. https://doi.org/10.1080/0142159x.2017.1315066.

Ericsson, K. Anders. 2007. "An Expert-Performance Perspective of Research on Medical Expertise: The Study of Clinical Performance." *Medical Education* 41 (12): 1124–30. https://doi.org/10.1111/j.1365-2923.2007.02946.x.

Ferguson, Peter C., Kelly J. Caverzagie, Markku T. Nousiainen, and Linda Snell. 2017. "Changing the Culture of Medical Training: An Important Step toward the Implementation of Competency-Based Medical Education." *Medical Teacher* 39 (6): 599–602. https://doi.org/10.1080/0142159x.2017.1315079.

Frank, Jason R., Linda S. Snell, Olle Ten Cate, Eric S. Holmboe, Carol Carraccio, Susan R. Swing, Peter Harris, et al. 2010. "Competency-Based Medical Education: Theory to Practice." *Medical Teacher* 32 (8): 638–45. https://doi.org/10.3109/0142159x.2010.501190.

Freire, Karine and Daniela Sangiorgi. 2010. "Service Design and Healthcare Innovation: From Consumption to Co-Production and Co-Creation." Paper Nordic Service Design Conference, Linkoping, Sweden. Accessed on November 30, 2014 at http://www.servdes.org/pdf/freire-sangiorgi.pdf.

Friedman, Karen A., Sandy Balwan, Frank Cacace, Kyle Katona, Suzanne Sunday, and Saima Chaudhry. 2014. "Impact on House Staff Evaluation Scores When Changing from a Dreyfus- to a Milestone-Based Evaluation Model: One Internal Medicine Residency Programs Findings." *Medical Education Online* 19 (1): 25185. https://doi.org/10.3402/meo.v19.25185.

Fuchs, Victor R. 1968. "Summary of Findings." In *The Service Economy*, 87:12. Cambridge, MA: National Bureau of Economic Research. Accessed November 30, 2014 at http://www.nber.org/books/fuch68-1

Garn Harvey A., M.J. Flax, M. Springer, and J.B. Taylor. 1976. "Models for Indicator Development: A Framework for Policy Analysis." *Urban Institute Paper,* April:1206–17. Washington, DC: Urban Institute.

Gordon, Morris, Jeanne Farnan, Ciaran Grafton-Clarke, Ridwaan Ahmed, Dawne Gurbutt, John Mclachlan, and Michelle Daniel. 2019. "Non-Technical Skills Assessments in Undergraduate Medical Education: A Focused BEME Systematic Review: BEME Guide No. 54." *Medical Teacher* 41 (7): 732–45. https://doi.org/10.1080/0142159x.2018.1562166.

Hall, Andrew K., Jessica Rich, J. Damon Dagnone, Kristen Weersink, Jaelyn Caudle, Jonathan Sherbino, Jason R. Frank, Glen Bandiera, and Elaine Van Melle. 2020. "It's a Marathon, Not a Sprint: Rapid Evaluation of CBME Program Implementation.." *Academic Medicine* 95 (5): 786–93. https://doi.org/10.1097/acm.0000000000003040.

Hamza, Deena M., Shelley Ross, and Ivy Oandasan. 2020. "Process and Outcome Evaluation of a CBME Intervention Guided by Program Theory." Journal of Evaluation in Clinical Practice. https://doi.org/10.1111/jep.13344.

Holmboe, Eric S., Jonathan Sherbino, Donlin M. Long, Susan R. Swing, and Jason R Frank. 2010. "The Role of Assessment in Competency-based Medical Education." *Medical Teacher* 32 (8): 676–82.

Holmboe, Eric S., Kenji Yamazaki, Laura Edgar, Lisa Conforti, Nicholas Yaghmour, Rebecca S. Miller, and Stanley J. Hamstra. 2015. "Reflections on the First 2 Years of Milestone Implementation." *Journal of Graduate Medical Education* 7 (3): 506–11. https://doi.org/10.4300/jgme-07-03-43.

Institute of Medicine (IOM). 2014. *Graduate Medical Education that Meets the Nation's Health Needs.* Washington, DC: The National Academies Press.

Kogan, Jennifer R., and Eric Holmboe. 2013. "Realizing the Promise and Importance of Performance-Based Assessment." *Teaching and Learning in Medicine* 25 (sup1). https://doi.org/10.1080/10401334.2013.842912.

McGaghie, William C. and Laurette Lipson. 1978. "Competency-Based Curriculum Development in Medical Education: An Introduction." *Public Health Papers* 68 Geneva: World Health Organization.

Medical Research Council (United Kingdom). 2014. “Developing and Evaluating Complex Interventions: New Guidance.” Accessed at www.mrc.ac.uk on January 5, 2014.
Molloy E. and D. Boud. 2013. “Changing Conceptions of Feedback.” In *Feedback in Higher and Professional Education*. Edited by D. Boud and E. Molly. New York: Routledge.
Nasca, Thomas J., Ingrid Philibert, Timothy Brigham, and Timothy C. Flynn. 2012. “The Next GME Accreditation System — Rationale and Benefits.” *New England Journal of Medicine* 366 (11): 1051–56. https://doi.org/10.1056/nejmsr1200117.
Normann, Richard. 2001. *Reframing Business: When the Map Changes the Landscape.* London: Wiley Publishing.
Oandasan, Ivy, Liz Martin, Melissa McGuire, and Rochelle Zorzi. 2020. “Twelve Tips for Improvement-Oriented Evaluation of Competency-Based Medical Education. *Medical Teacher* 42(3): 272–277.
Ostrom, Elinor. 1996. “Crossing the Great Divide: Coproduction, Synergy, and Development.” *World Development* 24 (6): 1073–87. https://doi.org/10.1016/0305-750x(96)00023-x.
Pawson, Ray. 2013. *The Science of Evaluation: A Realist Manifesto*. London: Sage Publications.
Pawson, Ray and N. Tilley. 1997. *Realistic Evaluation.* London: Sage Publications.
Rogers P.J. 2011. “Implications of complicated and complex characteristics for key tasks in evaluation.” In *Evaluating the Complex: Attribution, Contribution and Beyond*, edited by Kim Forss, Mita Marra, and Robert Schwartz. 33-53. New Brunswick, New Jersey: Transaction Publishers.
Ross, Shelley, Natalia M. Binczyk, Deena M. Hamza, Shirley Schipper, Paul Humphries, Darren Nichols, and Michel G. Donoff. 2018. “Association of a Competency-Based Assessment System with Identification of and Support for Medical Residents in Difficulty.” *JAMA Network Open* 1 (7). https://doi.org/10.1001/jamanetworkopen.2018.4581.
Sabadossa K.A. and P.B. Batalden. 2014. “The Interdepedent Roles of Patients, Families and Professionals in Cystic Fibrosis: A System for the Coproduction of Healthcare and its Improvement. *BMJ Quality & Safety* 23: i90-i94 https://doi.org/10.1136/bmjqs-2013-002782.
Sangha, Sonia and Stanley J. Hamstra. 2019. “Milestones Bibliography: December 2019.” Accessed April 3, 2020: https://www.acgme.org/Portals/0/PDFs/Milestones/MilestonesBibliography_Dec%202019%20Final%2003302020.pdf?ver=2020-04-03-130735-153
Sargeant, Joan, Jocelyn Lockyer, Karen Mann, Eric Holmboe, Ivan Silver, Heather Armson, Erik Driessen, et al. 2015. “Facilitated Reflective Performance Feedback.” *Academic Medicine* 90 (12): 1698–1706. https://doi.org/10.1097/acm.0000000000000809.
Schuwirth, Lambert W.T. and Cees P.M. Van der Vleuten. 2011. “Programmatic Assessment: From Assessment of Learning to Assessment for Learning.” *Medical Teache*r 33: 478–85.
Skeff, K. and G. Stratos. 2015. “Feedback.” Stanford Clinical Teaching Program. Accessed at http://sfdc.stanford.edu/clinical_teaching.html January 24, 2015.
Sullivan, Rick L. 1995. “The Competency-Based Approach to Training.” *Strategy Paper No 1*. Baltimore, Maryland: JHPIEGO Corporation.
Tannenbaum, Evan, Hossai Furmli, Nancy Kent, Sharon Dore, Margaret Sagle, and Nicolette Caccia. 2020. “Exploring Faculty Perceptions of Competency-Based Medical Education and Assessing Needs for Implementation in Obstetrics and Gynaecology Residency.” *Journal of Obstetrics and Gynaecology Canada* 42 (6): 707–17. doi:10.1016/j.jogc.2019.10.034.

Ten Cate, Olle. 2014. “The False Dichotomy of Quality and Quantity in the Discourse around Assessment in Competency-Based Education.” *Advances in Health Sciences Education* 20 (3): 835–38. doi:10.1007/s10459-014-9527-3.
VanMelle, Elaine Van, Jason R. Frank, Eric S. Holmboe, Damon Dagnone, Denise Stockley, and Jonathan Sherbino. 2019. “A Core Components Framework for Evaluating Implementation of Competency-Based Medical Education Programs.” *Academic Medicine* 94 (7): 1002–9. doi:10.1097/acm.0000000000002743.
Wagner, Edward H., Brian T. Austin, and Michael Von Korff. 1996. “Organizing Care for Patients with Chronic Illness.” *The Managed Care Quarterly* (4): 12–25.
Weiss, Kevin B., James P. Bagian, and Thomas J. Nasca. 2013. “The Clinical Learning Environment.” *JAMA.* 309 (16): 1687. doi:10.1001/jama.2013.1931.
Whitcomb, Michael E. 2016. “Transforming Medical Education.” *Academic Medicine* 91 (5): 618–20. doi:10.1097/acm.0000000000001049.

Appendix A: Additional CBME References

Amin, Zubair. 2011. “Purposeful Assessment.” *Medical Education* 46 (1): 4–7. https://doi.org/10.1111/j.1365-2923.2011.04170.x.

Aagaard, Eva, Gregory C. Kane, Lisa Conforti, Sarah Hood, Kelly J. Caverzagie, Cynthia Smith, Davoren A. Chick, Eric S. Holmboe, and William F. Iobst. 2013. “Early Feedback on the Use of the Internal Medicine Reporting Milestones in Assessment of Resident Performance.” *Journal of Graduate Medical Education* 5 (3): 433–38. https://doi.org/10.4300/jgme-d-13-00001.1.

Achike Francis I., Shaheen E. Lakhan, and Mohsin Yakub. 2019. “Competency-based Medical Education: Philosophy, What, How, Why, and the Challenges Therein.” *Journal of Medical Education* 23 (1): 1-13.

Albanese, Mark A. 2001. “Challenges in Using Rater Judgements in Medical Education.” *Journal of Evaluation in Clinical Practice* 6 (3): 305–19. https://doi.org/10.1046/j.1365-2753.2000.00253.x.

Baker, Keith. 2011. “Determining Resident Clinical Performance.” *Anesthesiology* 115 (4): 862–78. https://doi.org/10.1097/aln.0b013e318229a27d.

Banerjee, Yajnavalka, Christopher Tuffnell, and Rania Alkhadragy. 2019. “Mento’s Change Model in Teaching Competency-Based Medical Education.” *BMC Medical Education* 19 (1). https://doi.org/10.1186/s12909-019-1896-0.

Berger-Estilita, Joana, Sabine Nabecker, and Robert Greif. 2019. “A Delphi Consensus Study for Teaching ‘Basic Trauma Management’ to Third-Year Medical Students.” *Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine* 27 (1). https://doi.org/10.1186/s13049-019-0675-6.

Black, David. 2013. “Revalidation for Trainees and the Annual Review of Competency Progression (ARCP).” *Clinical Medicine* 13 (6): 570–72. https://doi.org/10.7861/clinmedicine.13-6-570.

Bonnema, Rachel A. and Abby L. Spencer. 2012. “Remediating Residents: Determining When Enough is Enough.” *Academic Internal Medicine Insight* 10 (4): 6–7.

Boyd, Victoria A., Cynthia R. Whitehead, Patricia Thille, Shiphra Ginsburg, Ryan Brydges, and Ayelet Kuper. 2017. “Competency-Based Medical Education: the Discourse of Infallibility.” *Medical Education* 52 (1): 45–57. https://doi.org/10.1111/medu.13467.

Branzetti, Jeremy, Michael A. Gisondi, Laura R. Hopson, and Linda Regan. 2019. “Aiming Beyond Competent: The Application of the Taxonomy of Significant Learning to Medical Education.” *Teaching and Learning in Medicine* 31 (4): 466–78. doi:10.1080/10401334.2018.1561368.

Carr, S.J. 2004. “Assessing Clinical Competency in Medical Senior House Officers: How and Why Should We Do It?” *Postgraduate Medical Journal* 80 (940): 63–66. https://doi.org/10.1136/pmj.2003.011718.

Carraccio, Carol, Susan D. Wolfsthal, Robert Englander, Kevin Ferentz, and Christine Martin. 2002. “Shifting Paradigms.” *Academic Medicine* 77 (5): 361–67. https://doi.org/10.1097/00001888-200205000-00003.

Cohen, Gerald S., Nancy L. Henry, and Pearl E. Dodd. 1990. “A Self-Study of Clinical Evaluation in the McMaster Clerkship.” *Medical Teacher* 12 (3-4): 265–72. https://doi.org/10.3109/01421599009006630.

Cohen, Gerald S., Phyllis Blumberg, Nancy C. Ryan, and Patricia L. Sullivan. 1993. “Do Final Grades Reflect Written Qualitative Evaluations of Student Performance?” *Teaching and Learning in Medicine* 5 (1): 10–15. https://doi.org/10.1080/10401339309539580.

Davis, David A., Paul E. Mazmanian, Michael Fordis, R. Van Harrison, Kevin E. Thorpe, and Laure Perrier. 2006. “Accuracy of Physician Self-Assessment Compared with Observed Measures of Competence.” *JAMA* 296 (9): 1094. https://doi.org/10.1001/jama.296.9.1094.

Downing, Steven M. 2005. “Threats to the Validity of Clinical Teaching Assessments: What about Rater Error?” *Medical Education* 39 (4): 353–55. https://doi.org/10.1111/j.1365-2929.2005.02138.x.

Dudek, Nancy L., Meridith B. Marks, and Glenn Regehr. 2005. “Failure to Fail: The Perspectives of Clinical Supervisors.” *Academic Medicine* 80 (Supplement). https://doi.org/10.1097/00001888-200510001-00023.

Dudek, Nancy L., Meridith B. Marks, Timothy J. Wood, et al. 2012. “Quality Evaluation Reports: Can a Faculty Development Program Make a Difference?” *Medical Teacher* 34: e725–e731.

Englander, Robert, Jason R. Frank, Carol Carraccio, et al. 2017. “Toward a Shared Language for Competency-Based Medical Education.” *Medical Teacher* 39 (6): 582–587.

Frank, Jason R., Linda Snell, Robert Englander, Eric S. Holmboe, and ICBME Collaborators. 2017. “Implementing Competency-Based Medical Education: Moving Forward.” *Medical Teacher* 39 (6): 568–573.

Friedlander, R.B., Victoria Green, Jamie S. Padmore, and Kerry M. Richard. 2006. “Legal Issues in Residency Training.” In *The Life Curriculum Teachers Guide II*, edited by Karen Andolsek, 8–35. Durham, North Carolina: Duke University School of Medicine. http://wellness.som.jhu.edu/dl/LIFE%20Teachers%20Guide%202.pdf.

Gaglione, Margaret Mackrell, Lisa Moores, Louis Pangaro, and Paul A. Hemmer. 2005. “Does Group Discussion of Student Clerkship Performance at an Education Committee Affect an Individual Committee Member???s Decisions?” *Academic Medicine* 80 (Supplement). https://doi.org/10.1097/00001888-200510001-00016.

Gifford, Kimberly A., and Leslie H. Fall. 2014. “Doctor Coach: A Deliberate Practice Approach to Teaching and Learning Clinical Skills.” *Academic Medicine* 89 (2): 272–76. https://doi.org/10.1097/acm.0000000000000097.

Ginsburg, Shiphra, Jodi Mcilroy, Olga Oulanova, Kevin Eva, and Glenn Regehr. 2010. “Toward Authentic Clinical Evaluation: Pitfalls in the Pursuit of Competency.” *Academic Medicine* 85 (5): 780–86. https://doi.org/10.1097/acm.0b013e3181d73fb6.

Ginsburg, Shiphra, Kevin Eva, and Glenn Regehr. 2013. “Do In-Training Evaluation Reports Deserve Their Bad Reputations? A Study of the Reliability and Predictive Ability of ITER Scores and Narrative Comments.” *Academic Medicine* 88 (10): 1539–44. https://doi.org/10.1097/acm.0b013e3182a36c3d.

Ginsburg, Shiphra, Jodi Mcilroy, Olga Oulanova, Kevin Eva, and Glenn Regehr. 2010. “Toward Authentic Clinical Evaluation: Pitfalls in the Pursuit of Competency.” *Academic Medicine* 85 (5): 780–86. https://doi.org/10.1097/acm.0b013e3181d73fb6.

Greaves, J.D., and J. Grant. 2000. “Watching Anaesthetists Work: Using the Professional Judgement of Consultants to Assess the Developing Clinical Competence of Trainees.” *British Journal of Anaesthesia* 84 (4): 525–33. https://doi.org/10.1093/oxfordjournals.bja.a013485.

Goebel, Emily A., Matthew J. Cecchini, and Michele M. Weir. 2017. “Resident and Supervisor Evaluation Outcomes of a CBME Pathology Curriculum.” *Canadian Journal of Pathology* 9 (1): 7.

Govaerts, Marjan J.B., Lambert W.T. Schuwirth, Arno M.M. Muijtjens, and Cees P.M. Van der Vleuten. 2006. “Broadening Perspectives on Clinical Performance Assessment: Rethinking the Nature of In-Training Assessment.” *Advances in Health Sciences Education* 12 (2): 239–60. https://doi.org/10.1007/s10459-006-9043-1.

Griewatz, Jan, Amir Yousef, Miriam Rothdiener, and Maria Lammerding-Koeppel. “Are We Preparing for Collaboration, Advocacy and Leadership? Targeted Multi-Site Analysis of Collaborative Intrinsic Roles Implementation in Medical Undergraduate Curricula.” *BMC Medical Education* 20, no. 1 (2020). https://doi.org/10.1186/s12909-020-1940-0.

Hamdy, Hossam, Kameshwar Prasad, M. Brownell Anderson, Albert Scherpbier, Reed Williams, Rein Zwierstra, and Helen Cuddihy. 2006. “BEME systematic review: Predictive values of measurements obtained in medical schools and future performance in medical practice.” *Medical Teacher* 28: 103–16.

Hamby, Hossam, Kameshwar Prasad, Reed Williams, and Fathi A. Salih. 2003. “Reliability and validity of the direct observation clinical encounter validation (DOCEE).” *Medical Education* 37: 205–212.

Hattie, John, and Helen Timperley. 2007. “The Power of Feedback.” *Review of Educational Research* 77, no. 1: 81–112. https://doi.org/10.3102/003465430298487.

Hatala, R., and G.R. Norman. 1999. “In-Training Evaluation during an Internal Medicine Clerkship.” *Academic Medicine* 74, no. 10. https://doi.org/10.1097/00001888-199910000-00059.

Hauer, Karen E., Lindsay Mazotti, Bridget O’Brien, Paul A. Hemmer, and Lowell Tong. 2011. “Faculty Verbal Evaluations Reveal Strategies Used to Promote Medical Student Performance.” *Medical Education Online* 16, no. 1: 6354. https://doi.org/10.3402/meo.v16i0.6354.

Hemmer, Paul A., Richard Hawkins, Jeffrey L. Jackson, and Louis N. Pangaro. 2000. “Assessing How Well Three Evaluation Methods Detect Deficiencies in Medical Students’ Professionalism in Two Settings of an Internal Medicine Clerkship.” *Academic Medicine* 75, no. 2: 167–73. https://doi.org/10.1097/00001888-200002000-00016.

Herbers, Jerome E., Gordon L. Noel, Glinda S. Cooper, Joan Harvey, Louis N. Pangaro, and Michael J. Weaver. 1989. “How Accurate Are Faculty Evaluations of Clinical Competence?” *Journal of General Internal Medicine* 4 (3): 202–8. https://doi.org/10.1007/bf02599524.

Hodges B. 2013. “Assessment in the Post-Psychometric Ear: Learning to Love the Subjective and Collective.” *Medical Teacher* 35 (7): 564–8.

Holmboe, Eric S. 2004. “Faculty and the Observation of Trainees’ Clinical Skills: Problems and Opportunities.” *Academic Medicine* 79 (1): 16–22. https://doi.org/10.1097/00001888-200401000-00006.

Holmboe, Eric S., R.E. Hawkins. 1998. “Methods for Evaluating the Clinical Competence of Residents in Internal Medicine: A Review.” *Annals of Internal Medicine* 129 (1): 42. https://doi.org/10.7326/0003-4819-129-1-199807010-00011.

Holmboe Eric S., Jonathan Sherbino, Donlin M. Long, Susan R. Swing, Jason R. Frank. 2010. The role of assessing in competency-based medical education. *Medical Teacher*. 32: 676–682.

Holmboe, Eric S., Denham S. Ward, Richard K. Reznick, Peter J. Katsufrakis, Karen M. Leslie, Vimla L. Patel, Donna D. Ray, and Elizabeth A. Nelson. 2011. “Faculty Development in Assessment: The Missing Link in Competency-Based Medical Education.” *Academic Medicine* 86 (4): 460–467.

Iobst, William F., and Kelly J. Caverzagie. 2013. “Milestones and Competency-Based Medical Education.” Gastroenterology 145 (5): 921–24. https://doi.org/10.1053/j.gastro.2013.09.029.

Issenberg, S. Barry, William C. McGaghie, and Robert A. Waugh. 1999. “Computers and Evaluation of Clinical Competence.” *Annals of Internal Medicine* 130 (3): 244. https://doi.org/10.7326/0003-4819-130-3-199902020-00020.

Jones, M. Douglas, and Tai M. Lockspeiser. 2018. “Proceed with Caution: Implementing Competency-Based Graduate Medical Education.” *Journal of Graduate Medical Education* 10 (3): 276–78. https://doi.org/10.4300/jgme-d-18-00311.1.

Ketteler, Erika R., Edward D. Auyang, Kathy E. Beard, Erica L. Mcbride, Rohini Mckee, John C. Russell, Nova L. Szoka, and M. Timothy Nelson. 2014. “Competency Champions in the Clinical Competency Committee: A Successful Strategy to Implement Milestone Evaluations and Competency Coaching.” *Journal of Surgical Education* 71 (1): 36–38. https://doi.org/10.1016/j.jsurg.2013.09.012.

Kogan, Jennifer R., Eric S. Holmboe, and Karen E. Hauer. 2009. “Tools for Direct Observation and Assessment of Clinical Skills of Medical Trainees.” *JAMA* 302 (12): 1316. https://doi.org/10.1001/jama.2009.1365.

Langsley, Donald G. 1991. “Medical Competence and Performance Assessment.” *JAMA* 266 (7): 977. https://doi.org/10.1001/jama.1991.03470070107013.

Lavin, B., and L. Pangaro. 1998. “Internship Ratings as a Validity Outcome Measure for an Evaluation System to Identify Inadequate Clerkship Performance.” *Academic Medicine* 73 (9): 998–1002. https://doi.org/10.1097/00001888-199809000-00021.

Littlefield, J.H., D.A. DaRosa, K.D. Anderson, R.M. Bell, G.G. Nicholas, and P.J. Wolfson. 1991. “Accuracy of Surgery Clerkship Performance Raters.” *Academic Medicine* 66: S16–S18.

Lockyer, Jocelyn, Carol Carraccio, Ming-Ka Chan, Danielle Hart, Sydney Smee, Claire Touchie, Eric S. Holmboe, and Jason R. Frank. 2017. “Core Principles of Assessment in Competency-Based Medical Education. *Medical Teacher* 39 (6): 609–616.

Lurie, Stephen J., Christopher J. Mooney, and Jeffrey M. Lyness. 2009. “Measurement of the General Competencies of the Accreditation Council for Graduate Medical Education: A Systematic Review.” *Academic Medicine* 84 (3): 301–9. https://doi.org/10.1097/acm.0b013e3181971f08.

Melvin, Lindsay, and Rodrigo B. Cavalcanti. 2016. “The Oral Case Presentation.” JAMA 316 (21): 2187. https://doi.org/10.1001/jama.2016.16415.

Miller, A., and J. Archer. 2010. “Impact of Workplace Based Assessment on Doctors' Education and Performance: a Systematic Review.” *BMJ* 341 (sep24 1): c5064–c5064. https://doi.org/10.1136/bmj.c5064.

Moideen, Nikitha, Catherine De Metz, Maria Kalyvas, Eleftherios Soleas, Rylan Egan, and Nancy Dalgarno. 2020. “Aligning Requirements of Training and Assessment in Radiation Treatment Planning in the Era of Competency-Based Medical Education.” *International Journal of Radiation Oncology*Biology*Physics* 106 (1): 32–36. https://doi.org/10.1016/j.ijrobp.2019.10.005.

Nasca, Thomas J., Ingrid Philibert, Timothy Brigham, and Timothy C. Flynn. 2012. “The Next GME Accreditation System — Rationale and Benefits.” *New England Journal of Medicine* 366 (11): 1051–56. https://doi.org/10.1056/nejmsr1200117.

Noel, Gordon L., Jerome E. Herbert Jr., Madlen P. Caplow, Glinda S. Cooper, Louis N. Pangaro, Joan Harvey. 1992. “How Well Do Internal Medicine Faculty Members Evaluate the Clinical Skills of Residents?” *Annals of Internal Medicine* 117: 757–65.

Nousiainen Markku T., Kelly J. Caverzagie, Peter C. Ferguson, and Jason R. Frank. 2017. “Implementing Competency-Based Medical Education: What Changes in Curricular Structure and Processes are Needed?” *Medical Teacher* 39(6): 594–598.

O'Dowd, Emily, Sinéad Lydon, Paul O'Connor, Caoimhe Madden, and Dara Byrne. 2019. “A Systematic Review of 7 Years of Research on Entrustable Professional Activities in

Graduate Medical Education, 2011-2018." *Medical Education* 53 (3): 234–49. https://doi.org/10.1111/medu.13792.

Orr, Christine, and Ranil Sonnadara. 2019. "Coaching by Design: Exploring a New Approach to Faculty Development in a Competency-Based Medical Education Curriculum." *Advances in Medical Education and Practice Volume* 10: 229–44. https://doi.org/10.2147/amep.s191470.

Pandit, Subhendu, Merlin R. Thomas, A. Banerjee, Mohan Angadi, Sushil Kumar, Aseem Tandon, Tripti Shrivastava, Debasis Bandopadhyay, V.D.S. Jamwal, and D.R. Basannar. 2019. "A Crossover Comparative Study to Assess Efficacy of Competency Based Medical Education (CBME) and the Traditional Structured (TS) Method in Selected Competencies of Living Anatomy of First Year MBBS Curriculum: A Pilot Study." *Medical Journal Armed Forces India* 75 (3): 259–65. https://doi.org/10.1016/j.mjafi.2018.01.010.

Pangaro, L. 1999. "A New Vocabulary and Other Innovations for Improving Descriptive in-Training Evaluations." Academic Medicine 74 (11): 1203–7. https://doi.org/10.1097/00001888-199911000-00012.

Regehr, Glenn, Shiphra Ginsburg, Jodi Herold, Rose Hatala, Kevin Eva, and Olga Oulanova. 2012. "Using 'Standardized Narratives' to Explore New Ways to Represent Faculty Opinions of Resident Performance." *Academic Medicine* 87 (4): 419–27. https://doi.org/10.1097/acm.0b013e31824858a9.

Rosenberger, Kyle, Daniel Skinner, and Jody Monk. 2017. "Ready for Residency: A Bloomian Analysis of Competency-Based Osteopathic Medical Education." *The Journal of the American Osteopathic Association* 117 (8): 529. https://doi.org/10.7556/jaoa.2017.103.

Sanfey, Hilary, Janet Ketchum, Jennifer Bartlett, Stephen Markwell, Andreas H. Meier, Reed Williams, and Gary Dunnington. 2010. "Verification of Proficiency in Basic Skills for Postgraduate Year 1 Residents." *Surgery* 148 (4): 759–67. https://doi.org/10.1016/j.surg.2010.07.018.

Scavone, B.M., M.T. Sproviero, R.J. McCarthy, C.A. Wong, J.T. Sullivan, V.J. Siddall, and L.D. Wade. 2006. "Development of an Objective Scoring System for Measurement of Resident Performance on the Human Patient Simulator." *Anesthesiology*. 105: 260–6.

Schwind, Cathy J., Reed G. Williams, Margaret L. Boehler, and Gary L. Dunnington. 2004. "Do Individual Attendings' Post-Rotation Performance Ratings Detect Residents' Clinical Performance Deficiencies?" *Academic Medicine* 79 (5): 453–57. https://doi.org/10.1097/00001888-200405000-00016.

Stillman, Paula L., D.B. Swanson, S. Smee, A.E. Stillman et al. 1986. "Assessing Clinical Skills of Residents with Standardized Patients." *Annals of Internal Medicine* 105 (5): 762. https://doi.org/10.7326/0003-4819-105-5-762.

Storrar, Neill, David Hope, and Helen Cameron. 2018. "Student Perspective on Outcomes and Process – Recommendations for Implementing Competency-Based Medical Education." *Medical Teacher* 41 (2): 161–66. https://doi.org/10.1080/0142159x.2018.1450496.

Swing, Susan R., Stephen G. Clyman, Eric S. Holmboe, and Reed G. Williams. 2009. "Advancing Resident Assessment in Graduate Medical Education." *Journal of Graduate Medical Education* 1 (2): 278–86. https://doi.org/10.4300/jgme-d-09-00010.1.

Swing, Susan R. and International CBME Collaborators. 2010. “Perspectives on Competency-Based Medical Education from the Learning Sciences.” *Medical Teacher* 32 (8): 663–68. https://doi.org/10.3109/0142159x.2010.500705.
Tannenbaum, Evan, Hossai Furmli, Nancy Kent, Sharon Dore, Margaret Sagle, and Nicolette Caccia. 2020. “Exploring Faculty Perceptions of Competency-Based Medical Education and Assessing Needs for Implementation in Obstetrics and Gynaecology Residency.” *Journal of Obstetrics and Gynaecology Canada* 42 (6): 707–17. https://doi.org/10.1016/j.jogc.2019.10.034.
Ten Cate, Olle, and Stephen Billett. 2014. “Competency-Based Medical Education: Origins, Perspectives and Potentialities.” *Medical Education* 48 (3): 325–32. https://doi.org/10.1111/medu.12355.
Tesser, Abraham, and Sidney Rosen. 1975. “The Reluctance to Transmit Bad News.” *Advances in Experimental Social Psychology Advances in Experimental Social Psychology Volume 8*, 193–232. https://doi.org/10.1016/s0065-2601(08)60251-8.
Tonesk, X., and R.G. Buchanan. 1987. “An AAMC Pilot Study by 10 Medical Schools of Clinical Evaluation of Students.” *Academic Medicine* 62 (9): 707–18. https://doi.org/10.1097/00001888-198709000-00001.
Touchie, Claire, and Olle Ten Cate. 2015. “The Promise, Perils, Problems and Progress of Competency-Based Medical Education.” *Medical Education* 50 (1): 93–100. https://doi.org/10.1111/medu.12839.
Visram, Kash. 2018. “The Rationale for CBME and Early Impressions.” *Canadian Urological Association Journal* 12 (6): 155. https://doi.org/10.5489/cuaj.5422.
Visram, Kashif. 2019. “The Role of Mobile Technology for Resident Assessment of Surgical Skills in the CBME Era.” *Canadian Urological Association Journal* 13 (2). https://doi.org/10.5489/cuaj.5849.
Walsh, Allyn, Sudha Koppula, Viola Antao, Cheri Bethune, Stewart Cameron, Teresa Cavett, Diane Clavet, and Marion Dove. 2017. “Preparing Teachers for Competency-Based Medical Education: Fundamental Teaching Activities.” *Medical Teacher* 40 (1): 80–85. https://doi.org/10.1080/0142159x.2017.1394998.
Whitcomb, Michael E. 2016. “Transforming Medical Education: Is Competency-Based Medical Education the Right Approach?” *Academic Medicine* 91 (5): 618–20. https://doi.org/10.1097/acm.0000000000001049.
Whitehead, Cynthia R., and Ayelet Kuper. 2014. “Competency-Based Training for Physicians: Are We Doing No Harm?” *Canadian Medical Association Journal* 187 (4). https://doi.org/10.1503/cmaj.140873.
Wilkinson, James R., James G.M. Crossley, Andrew Wragg, Peter Mills, George Cowan, and Winnie Wade. 2008. “Implementing Workplace-Based Assessment across the Medical Specialties in the United Kingdom.” *Medical Education* 42 (4): 364–73. https://doi.org/10.1111/j.1365-2923.2008.03010.x.
Williams, Reed G., Debra A. Klamen, and William C. McGaghie. 2003. “Cognitive, Social and Environmental Sources of Bias in Clinical Performance Ratings.” *Teaching and Learning in Medicine* 15 (4): 270–92. https://doi.org/10.1207/s15328015tlm1504_11.
Williams, Reed G., Hilary Sanfey, Xiaodong (Phoenix) Chen, and Gary L. Dunnington. 2012. “A Controlled Study to Determine Measurement Conditions Necessary for a Reliable and Valid Operative Performance Assessment.” *Annals of Surgery* 256 (1): 177–87. https://doi.org/10.1097/sla.0b013e31825b6de4.
Williams, Reed G., Steven Verhulst, Jerry A. Colliver, and Gary L. Dunnington. 2005. “Assuring the Reliability of Resident Performance Appraisals: More Items or More Observations?” Surgery 137 (2): 141–47. https://doi.org/10.1016/j.surg.2004.06.011.

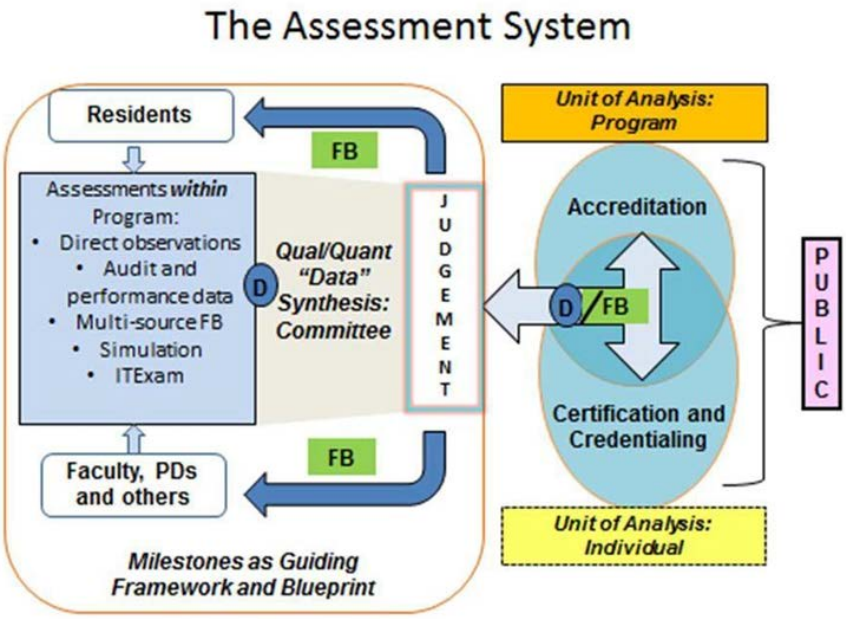
Williams, Reed G., Gary L. Dunnington, and Debra L. Klamen. 2005. “Forecasting Residents’ Performance—Partly Cloudy.” *Academic Medicine* 80 (5): 415–22. https://doi.org/10.1097/00001888-200505000-00002.

Williams, Reed G., Debra A. Klamen, and William C. McGaghie. 2003. “Cognitive, Social and Environmental Sources of Bias in Clinical Performance Ratings.” *Teaching and Learning in Medicine* 15 (4): 270–92. https://doi.org/10.1207/s15328015tlm1504_11.

Williams, Reed G., Hilary Sanfey, Xiaodong (Phoenix) Chen, and Gary L. Dunnington. 2012. “A Controlled Study to Determine Measurement Conditions Necessary for a Reliable and Valid Operative Performance Assessment.” *Annals of Surgery* 256 (1): 177–87. https://doi.org/10.1097/sla.0b013e31825b6de4.

Williams, Reed G., Cathy J. Schwind, Gary L. Dunnington, John Fortune, David Rogers, and Margaret Boehler. 2005. “The Effects of Group Dynamics on Resident Progress Committee Deliberations.” *Teaching and Learning in Medicine* 17 (2): 96–100. https://doi.org/10.1207/s15328015tlm1702_1.

Appendix B: The High Performing Residency/Fellowship Assessment System



At the program level, residents/fellows are assessed routinely through a combination of many assessment tools, including direct observations; global evaluation; audits and review of clinical performance data; multisource feedback from team members, including peers, nurses, patients, and family members; simulation; in-service training examinations (ITE); self-assessment; and others. Increasingly, the Milestones and EPAs are used as a guiding framework and "blueprint" for expected performance. Assessment tools are selected intentionally to allow routine, frequent, formative feedback to a resident/fellow to affirm areas of successful performance and to highlight those aspects that need to be improved. The CCC is the committee that synthesizes data—quantitative from in-service exams and clinical performance audits, and qualitative from observers and co-workers. Using the Milestones, the CCC forms a consensus decision, or a judgment, regarding each resident's/fellow's performance. The CCC provides those conclusions to the program director, who makes the final determination on each resident's/fellow's Milestone "level" at least twice yearly. These are provided to the applicable ACGME Review Committee and, in some cases, the pertinent specialty boards. The ACGME's unit of analysis is the program, and the Review Committees use aggregate Milestones information comparing a program with all accredited programs in the given specialty.

The comparison against these benchmarks serves as one source of input into the ACGME's determination of program quality and accreditation decisions. The unit of analysis is the "individual" for certification and credentialing entities. Collectively, all residents/fellows, faculty members/program directors/programs, the ACGME, and certification and credentialing entities are accountable to the public for honest assessments of residents'/fellows' performance and truthful verification of their readiness to progress to independent practice. Data (D) is essential for the entire system to engage in continuous quality improvement, especially to create meaningful feedback (FB) loops within the program and back to programs from the ACGME. Programs, residents, and fellows can currently download their Milestones report after each reporting period.