



ACGME

The Milestones Guidebook

Eric S. Holmboe, MD
Laura Edgar, EdD, CAE
Stan Hamstra, PhD

Version 2016

TABLE OF CONTENTS

<u>Topic</u>	<u>Page(s)</u>
Preface.....	3
How to Use This Guidebook.....	4
Competency-based Education and the Rationale for the Educational Milestones	5 - 8
Milestones.....	9 - 30
• What are Milestones.....	9 - 11
• How Were Milestones Developed?.....	11 - 12
• Why Milestones.....	13- 15
• Implementing and Using Milestones Effectively.....	15 - 21
• Importance of Feedback.....	22 - 23
• Early Lessons Learned About Milestones.....	24 - 25
• Milestones and the NAS Assessment System.....	26 - 28
• How will the ACGME Evaluate the Milestones?.....	28 - 29
• Conclusions.....	30
Appendix 1: Additional Educational and Assessment Resources	31 - 32
Appendix 2: Annotated Bibliography of Selected Key Literature.....	33- 37
References.....	38 - 41

PREFACE

Milestones have become an important formative component of the accreditation system for graduate medical education (GME) in the United States. The Next Accreditation System (NAS) was part of the educational community's response to public and policy makers' concerns regarding the need to improve GME.¹ The NAS more fully embraces the outcomes-based principles that started with the release of the General Competencies in 1999, and the launch of the Outcomes Project in 2001.^{2,3} However the ACGME and programs struggled to operationalize the Competencies and create meaningful outcomes-based assessments. Recognizing these challenges, the NAS included two important new components to accreditation; Milestones and Clinical Competency Committees (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 guidebook will begin with a brief history of the ACGME, NAS, and competency-based medical education (CBME), followed by an overview of the philosophy and theory underlying the Milestones. Over the last two years, ACGME staff have interacted with hundreds of program directors and faculty members to begin to learn what is working and what is not. In addition, early research into the Milestones is beginning to provide some insights into the Milestones system. Building on these lessons "from the field" and the early research, the guidebook will provide some practical suggestions for the effective use of Milestones to help programs not only improve but also transform their assessment practices, curriculum, and overall residency and fellowship programs. A companion guidebook on the CCC can be found on the ACGME Milestones web page.⁴

HOW TO USE THIS GUIDEBOOK

This guidebook is designed to be both sufficiently comprehensive but also practical. For those interested in more of the history and rationale of competency-based medical education (CBME) and the Milestones, the early sections of this guidebook (pages 5 - 12) provide this information. In addition, Appendix 2 provides an annotated bibliography of key articles that can help readers deepen their understanding of the theory and research performed on CBME and the Milestones.

For those looking for practical solutions, the sections on *Implementing and Using Milestones Effectively* (pages 12 - 15), *Importance of Feedback* (pages 16 - 17), and *Early Lessons Learned About Milestones* (pages 18 - 20) will be the most helpful. “Practical Tips” boxes that offer quick summaries are provided in certain sections. Finally, Appendix 1 provides a list of potentially useful educational and assessment resources that may help programs with implementation of the Milestones.

We also welcome feedback on this first edition of the *Milestones Guidebook*. This guidebook will be reviewed yearly and updated. You can send your feedback 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 Next Accreditation System (NAS) (Table 1). Key dates include the approval of the Competencies in 1999, the launch of the Outcome Project in 2001, and the transition of the first phase of accredited specialties to the NAS in July 2013.^{1,2}

Table 1: Key Dates in Educational Milestones History

Dates	Structure
1999	The six General Competencies endorsed by ACGME and American Board of Medical Specialties (ABMS)
2001	The Outcomes Project formally launched
2009	ACGME approves structure of Next Accreditation System, including inclusion of Milestones
2013	First seven specialties fully enter NAS, including Milestones reporting
2014	Remaining accredited specialties and subspecialties enter NAS, including Milestone reporting
2015	All specialties and subspecialties begin to report Milestones data

Competency-based medical education (CBME) serves as the foundation for the NAS. The NAS is also grounded in a continuous quality improvement and innovation philosophy.^{1,5} Before we examine the role of the Milestones in assessment and programmatic improvement, it is useful to briefly review the history of CBME.

Overview: Competency-based Medical Education (CBME)

Competency-based educational models are not new. In other fields, is it often called competency-based education and training (CBET), a term transformed to CBME in medicine. What is CBET? As Sullivan notes (1995):⁶

“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 actually 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 competency-based education and training (CBET) had much of its genesis in the teacher education reform movement of the 1960s.⁷ This interest was spurred by a

US Office of Education National Center for Education research grant program in 1968 to 10 universities to develop and implement new teacher training models that focused on student achievement (outcomes). Elam laid down a series of principles and characteristics of CBET in 1971 (Table 2). From these beginnings, interest within medical education began to grow.⁷

Table 2: Principles and Characteristics of Competency-based Educational Models⁷

Principles	Characteristics
<ol style="list-style-type: none"> 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 also takes knowledge into account 4. Individual learners progress at rates dependent on demonstrated competency 5. The instructional program facilitates development and evaluation of the specific competencies 	<ol style="list-style-type: none"> 1. Learning is individualized 2. Feedback to the learner is critical 3. Emphasis is more on the exit criteria than on the admission criteria 4. CBET requires a systematic program (approach) 5. Training is modularized 6. Both the learner and the program have accountability

Competency-based models for medical education were first 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.”⁸

In a 2002 review, Carraccio and colleagues noted that some sectors in medical education explored competency-based models in the 1970s, but except for one study, no comparisons between competency-based and the traditional structure/process-based curricula were undertaken.⁹ Of the few studies within medical fields that have investigated competency-based models, there appear to be some benefits to learners in the CBME model.^{9,10}

In the context of medicine, Carraccio and colleagues compared the elements between the structure/process-based educational approach and the outcomes-based approach in 2002. (Table 3).⁹

Table 3: Comparison of Structure/Process-based vs. Competency-based Programs

Variable	Educational Program Approach	
	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 educ. encounter	Knowledge acquisition	Knowledge application
Typical assessment tool	Single subject measure	Multiple objective 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, 2002.

Finally, the authors also described a four-step process for implementing CBME: 1) Identification of the competencies (in the United States the six ACGME/ABMS General Competencies); 2) determination of competency components and performance levels (e.g., benchmarks and milestones); 3) competency evaluation; and 4) overall assessment of the process.⁹

More recently, 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:¹¹

“an outcomes-based approach to the design, implementation, assessment and evaluation of a medical education program using an organizing framework of competencies”

In addition, Elaine van Melle and colleagues outline five core components for CBME they are using as part of an institution-wide implementation of CBME at Queens University in Kingston, Ontario (van Melle, personal communication):

1. Competencies required for practice are clearly articulated.
2. Competencies are arranged progressively.
3. Learning experiences facilitate the progressive development of competencies.
4. Teaching practices promote the progressive development of competencies.
5. Assessment practices support and document the progressive development of competencies.

A key distinguishing feature of competency-based education and training is that learners could potentially progress through the educational process at *different* rates: the most capable and talented individuals should be able to make career

transitions *earlier*, while others will require more time (up to a point) 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 valuable resource that should be used wisely and effectively. No one would argue that a certain quantity of experience is not still important.¹² 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 will work in “hybrid models” of CBME – using competency-based educational principles in the context of fixed years of training. A second key feature is the increased emphasis on assessment, especially ongoing, longitudinal assessment that enables the faculty to more accurately determine the developmental progress of the learner, as well as to help the learner through frequent feedback, coaching, and adjustments to learning plans.^{13,14} This is consistent with Anders Ericsson’s work in expertise and deliberate practice demonstrating the need to tailor the educational experience to continually challenge the learner with experiences that are neither too easy nor overwhelming (too hard).¹⁵

While defining the Competencies was an important and necessary step, operationalizing and implementing them prior to the Milestones proved to be very 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 “look like” in practice. This lack of shared understanding (i.e., shared mental models) hampered curricular changes and development and evolution of better assessment methods. The challenges to operationalizing the Competencies was not restricted to the US, and over the last 10 years several notable concepts have emerged in an effort to enable more effective implementation of CBME, such as the Milestones and entrustable professional activities (EPAs). Both concepts approach competency as a developmental process and rely heavily on positivist behavioral theory. The Milestones have become an essential component of the NAS, and we hope this guidebook will provide 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 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 sub-competencies 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 clinical 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. The terminology used within the Milestones is included in Figures 1a and 1b.

The Milestones describe the learning trajectory within a subcompetency that takes the resident or fellow from a beginner in the specialty or subspecialty, to a highly proficient resident or fellow or early practitioner. 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. 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 competency.

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, but programs will need to use good judgment to fill in the gaps in curriculum and assessment. It is essential that the Milestones are not thought of as curricula 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. Second, 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.

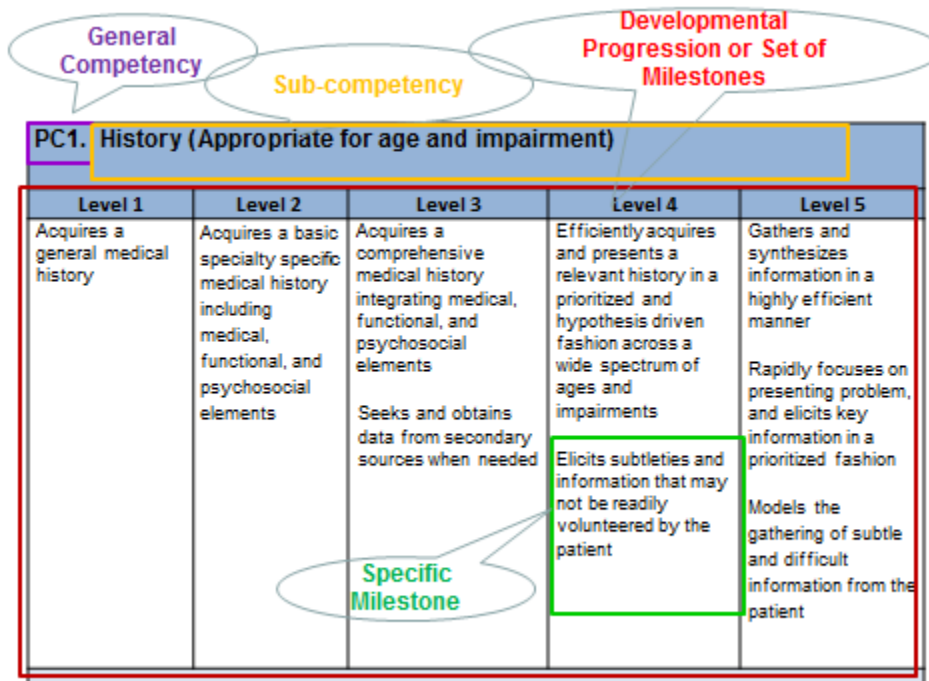
The *entire* Milestones document (set) used for NAS reporting was also never intended to serve as a regular assessment tool, especially for short rotations (e.g., 2-8 weeks). The Milestones, and even the more specific subcompetencies, do not

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

Milestone Description: Template				
Level 1	Level 2	Level 3	Level 4	Level 5
What are the expectations for a beginning resident?	What are the milestones for a resident who has advanced over entry, but is performing at a lower level than expected at mid-residency?	What are the key developmental milestones mid-residency? What should they be able to do well in the realm of the specialty at this point?	What does a graduating resident look like? What additional knowledge, skills & attitudes have they obtained? Are they ready for certification?	Stretch Goals – Exceeds expectations
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Comments:				

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



© 2013 Accreditation Council for Graduate Medical Education (ACGME)

How were Milestones Developed?

The process of Milestone 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 Working Group for each of the core specialties was fully developed. That same year, the decision to include five levels within the Milestones was made, guided by the Dreyfus Model of expertise development.² 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 developing (i.e., fewer Levels, levels with different descriptions, different graduation targets).

Each Working Group was composed of members of the relevant ACGME Review Committee, the American Board of Medical Specialties through the individual certification boards, the specialty colleges and associations, and program directors' groups, and representative residents and/or fellows. The core specialty Working Groups typically had 15 members representing the varying areas within the applicable specialty. For example, the Orthopaedic Surgery Working Group included

members who could represent the eight subspecialties within the practice of orthopaedic surgery.

Each group met three or four times to complete the process, which in each case started with reviews of published documents, including the Program Requirements, certification blue prints, competency statements, shared curricula, and other literature. The discussion of what knowledge, skills, and abilities 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, and the work of development started later.

The Milestones are written a few different ways across the specialties; some include general and broad categories, some are very specific, and others are a mix. Generally speaking, the specialties with more residents/fellows (e.g., internal medicine, family medicine, surgery) tended to be broader in their Milestones, and those specialties with fewer and more specialized residents/fellows (e.g., orthopaedic surgery, radiation oncology) were more specific. Another difference is that those sets of Milestones that were more specific tended to include more subcompetencies and individual milestones within them.

For most specialties, after the draft was completed, it was shared with an Advisory Committee to perform a review and comment on the work that had been completed. After several rounds of editing, the Milestones were shared with program directors for piloting and feedback. The outcomes of the pilot testing and the feedback received were utilized to edit and finalize the documents. Some specialties utilized Milestone sets that had been considered either duplicative or too elementary and were published as an appendix that could be used as a remediation or learning tool; these are sometimes referred to as “non-reportable Milestones.” Other specialties created more detailed documents to demonstrate the research behind the Milestones that were selected and/or potential assessment tools that could be used to better determine the Milestone level.

When the Milestones were published, the ACGME Milestones staff attended many program directors’ group meetings. An effort was made to educate the program directors on the background and purpose of the Milestones, as well as on how to interpret the language. Eventually, these presentations included workshops for Mock Clinical Competency Committee Meetings (more information on the requirements for the Clinical Competency Committee can be found in the *Clinical Competency Committee Guidebook*). Each core specialty also published an article in the *Journal of Graduate Medical Education* that detailed the specific process used.¹⁶⁻²⁴

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 general competencies since their introduction in 1999.² The Milestones, along with the related concept of 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 faculty members have found very difficult to use effectively. Third, the Milestones provide guidance on curriculum by defining the general, essential competencies within a discipline. Fourth, the Milestones can serve as a guide and “item bank” to create more meaningful assessments. Fifth, 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 Clinical Competency Committee (CCC) deliberations and judgments.

There are other Milestones, besides the Reporting Milestones, that should be mentioned. A few specialties have developed more granular, specialty-detailed Milestones that are often referred to as “curricular milestones.” As the name implies, these are designed to help programs attend to key elements of their curriculum in more detail, and to help create more meaningful assessments for specific curricular experiences. Not all specialties have created these more descriptive, granular milestones. Programs are not required to report to the ACGME on these milestones. Primarily, they are utilized by internal medicine and pediatrics and their related subspecialties, to guide curriculum development and specific assessments.²⁵

The Milestones play a number of important roles depending on the constituent or stakeholder. Table 4 provides an overview of the purposes and functions of the Milestones related to each key stakeholder.²⁶

Table 4: The Purpose and Function of Milestones

Constituency or Stakeholder	Purpose/Function
Residents and Fellows	<ul style="list-style-type: none"> • Provide a descriptive roadmap for training • Increased transparency of performance requirements • Encourage informed self-assessment and self-directed learning • Facilitate better feedback to trainee • Encourage self-directed feedback seeking behaviors
Residency and Fellowship Programs	<ul style="list-style-type: none"> • 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
ACGME	<ul style="list-style-type: none"> • Accreditation – enables continuous monitoring of programs and lengthening of site visit cycles • Public Accountability – report at a aggregated national level on competency outcomes • Community of practice for evaluation and research, with focus on continuous improvement
Certification Boards	<ul style="list-style-type: none"> • Enable research to improve certification processes

Several key aspects about the use of the Milestones deserve special attention. First, as noted above, the Milestones 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 Reporting 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 rotation and curricular experiences. Second, the Reporting Milestones can also be used for guided self-

assessment and reflection by the resident/fellow in preparation for feedback sessions and in creating individual learning plans. Residents and fellows should use the Milestones for self-assessment with input and feedback from 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 limiting assessments to the Milestones could leave many topics without proper and essential assessment and evaluation.

Implementing and Using Milestones Effectively

While we still have 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 and fellows at the beginning of the program. This helps them to gain a shared understanding of the goals of the program and Milestones.*
2. *Have residents and fellows complete individualized learning plans, using the Milestones as an important guide.*
3. *Consider having residents and fellows complete a self-assessment of their Milestones that they can compare and contrast, with a trusted advisor, to the Milestone judgments of the CCC every six months.*
4. *Enable residents and fellows to seek out assessment (i.e., self-directed assessment seeking), especially direct observation, from faculty members.*

Residents and fellows are a primary stakeholder in the Milestones system. Education is always co-created and co-produced between teacher and learner.²⁷⁻³³ 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 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.³⁰⁻¹ Recently, Sabadossa and Batalden 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.”³⁰ Wagner, et. al. have described the importance of “activated patients” for the development of good care. Medical education-as-service is no different.³⁴

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,³⁵ 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. 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.;
3. 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;³⁶
4. develop personal learning plans that they revisit and revise at least twice a year;
5. actively seek out assessment and feedback on an ongoing basis;
6. provide systematic feedback to the program on their experience with the Milestones.

Summary - Practical Tips

1. *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.*
2. *Observe, observe, observe! Faculty observation of key competencies is essential to effective feedback, coaching, and professional development.*
3. *Embed observation in “what faculty 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, 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 training program. Our conception of faculty 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 training. As we’ll see below, this may mean that the program will need to revise the nature of the evaluation forms faculty members complete. Finally, 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:

1. Familiarize themselves with the overall Milestones
2. Hone in and 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
7. Provide ongoing feedback to learners, which is essential for good coaching and professional growth

Summary - Practical Tips

1. Share and discuss the pertinent Milestone set with faculty 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. Empower and facilitate direct observation by faculty members. Faculty observation of key competencies is essential to effective feedback, coaching, and professional development.
3. 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.
4. 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 the CCC meetings. Use the “practice makes perfect” principle through continued dialogue around the Milestones. This helps to deepen shared understanding.
5. Map your curriculum and assessment program against the pertinent milestones. This will help to identify curricular gaps and areas for opportunity, and ensure you have the most effective combination of assessments.

We recognize the NAS and the Milestones have substantially affected the role and nature of work for program directors and other program leaders. Program directors represent the critical and essential hub of the entire 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. Conduct a crosswalk of curriculum with the specialty Milestones 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.

Rotation or Curricular Experience	Goals and Objectives (Purpose: why this rotation?)	Essential Milestones (e.g., PCx; MK1x, etc.)	Assessment Method

2. Develop a program of assessment that aligns with the Milestones and functions as an integrated, holistic package. Assessment activities should tightly align with the actual training activity.
3. Identify and address gaps in assessment strategies to ensure meaningful and authentic Milestones judgments.
4. Conduct ongoing program evaluation to assess what is working, for who, 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 you have access to an education department or expertise, we'd encourage program leaders to sit down with these individuals to explore what the best program evaluation strategy would be for their programs.
5. Provide ongoing faculty development, especially around assessment. While workshops are clearly helpful, they are not enough. Think of ways the program can build “small aliquots” of faculty development into section or department meetings, grand rounds, CCC meetings, and so forth. 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. Build a team. Program directors cannot do this alone. 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. We have found on our travels that really good stuff is happening in at least some of the specialties within institutions of which others in the same institutions are unaware. Check with your designated institutional official (DIO) and graduate medical education committee (GMEC) to learn what is happening in your own institution.
7. Explore the functionality of your 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*.³⁷ 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:

Table 5: Common Assessment Methods for the Six General Competencies

General Competency	Common Assessment Methods
Patient Care	<ul style="list-style-type: none"> ▪ 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	<ul style="list-style-type: none"> ▪ In-training examinations ▪ Oral questioning methods (e.g., SNAPPS) ▪ Direct observation (live or video)
Professionalism	<ul style="list-style-type: none"> ▪ Multi-source feedback (MSF) ▪ Patient surveys (can be part of MSF) ▪ Direct observation
Interpersonal and Communication Skills	<ul style="list-style-type: none"> ▪ Multi-source feedback (MSF) ▪ Patient surveys (can be part of MSF) ▪ Direct observation (live or video) ▪ Simulation (including standardized patients)
Practice-based Learning and Improvement	<ul style="list-style-type: none"> ▪ Audit of clinical practice (e.g., quality performance measures) ▪ Evidence-based medicine logs ▪ Case Logs ▪ Rating scales/evaluation forms
Systems-based practice	<ul style="list-style-type: none"> ▪ Audit of clinical practice (e.g., quality performance measures) ▪ Multi-source feedback (MSF) ▪ Rating scales/evaluation forms

However, some key principles in building an assessment program can be helpful:

1. First and foremost, recognize that it is the individual completing an assessment tool (e.g., a faculty member) who serves as the measurement instrument, not the tool itself. The majority of variance in ratings and judgments is always with the rater. Therefore, faculty development is absolutely essential.
2. Where possible, use existing tools and modify as necessary. It is very difficult to develop new tools de novo, so taking existing tools that have been evaluated for validity is a good place to start, and modify if needed. Understand that if as a program chooses to modify an assessment instrument, it would be good to assess some of its basic performance, such as reliability and validity if feasible.

3. Ensure the assessments chosen possess a high degree of utility, as exemplified by this hybrid equation from Norcini and van der Vleuten.³⁸

$$\text{Utility} = \text{Validity} \times \text{Reliability} \times \text{Educational impact} \times \text{Acceptability} \times \text{Feasibility}$$

As you can see, if any one of these variables is zero, by definition utility is zero. For most purposes, the tools programs will use do not need to have high levels of reliability (i.e., reproducibility) as required for high stakes testing. For educational impact, the primary focus should be on the tool's "catalytic effects;" in other words the assessment should help to drive future learning forward. Tools have to be feasible, including in terms of time and financial cost, and must also be acceptable to both faculty members and learners.

4. Avoid tools that ask faculty members or others to judge too many items, competencies, or questions (too much "cognitive load"). Think carefully about what a faculty member can realistically observe and judge in the allotted time period. Many faculty members now spend only one to two weeks attending during a rotation in many specialties. Furthermore, the forms should align and comport with the purpose and main objectives of the rotation or curricular experience.
5. Ensure "construct alignment" of rating scales on evaluation forms.³⁹ Construct alignment refers to how the descriptors used in a rating scale align with the judgments being asked of the rater. For example, for years a typical scale included the descriptors "unsatisfactory," "satisfactory," and "superior." However, these descriptors are "misaligned" with the task of making a developmental or criterion-based judgment. Recent research has shown that scales and descriptors better aligned with the rating task and judgment have better reliability. An example of a construct-aligned scale is one that asks faculty members to judge entrustment or level of supervision. While much work remains to be done, it is worth taking a look at the program's current scales and descriptors to see if they are effectively guiding faculty members' judgment and ratings.
6. Encourage the use and capture of narrative assessment. While rating scales and evaluation forms can certainly be useful, it is the narrative that is often most useful, especially for feedback. Offer faculty development to structure the outline of narrative evaluations to facilitate synthesis of narrative comments to Milestones achievement.

The CCC is also a vital component of the assessment program and overall program system. We strongly encourage you to, in conjunction with this Milestones Guidebook, review the CCC Guidebook available on the ACGME website (www.acgme.org).

IMPORTANCE OF FEEDBACK

Feedback to the resident or fellow 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 and programs have to help residents and fellows learn and improve. The Milestones should be used to help residents and fellows develop action 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 performance data, as represented by the Milestones, should be facilitated and guided by a trusted advisor.

Five basic features of high quality feedback are:⁴⁰

1. Timeliness. Faculty members should always try to provide feedback in a timely fashion. The results of the CCC deliberations and Milestone determinations should also be shared 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 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.⁴⁶

Two feedback models might also be helpful.

ADAPT

Ask-Discuss-Ask-Plan Together. This model is built upon the work of Lyuba Konopasek and her earlier Ask-Tell-Ask (ATA) model. Since feedback should be a dialogue and not a one way conversation, the ATA model was revised to recognize this important aspect of feedback. Start by *asking* how things are going and encourage a self-assessment. *Discuss* with the learner your observations of their

self-assessment and how it relates to your feedback, looking for areas of concordance and discordance. Discordances are especially good opportunities for professional growth and helping the learner with the skill of self-reflection and calibration. Ask the learner to then reflect on the feedback session and gain their further input. Finally, *plan together* to decide action plans and needs for ongoing coaching.

R2C2 Model

This model was developed by Joan Sargeant and colleagues, who specifically included feedback sessions that involved the review of multi-source performance data, such as multi-source feedback and clinical performance measures, in their research.³⁶ The model builds on robust educational theory. The steps of the model are:

Rapport Building: In this initial stage, the faculty member should build rapport and establish the relationship. The goal of this stage is to explain the purpose of the assessment, engage the resident/fellow, and establish the credibility of the assessment. At this stage you want to outline and negotiate the agenda with the learner to ensure issues he/she wishes to discuss are surfaced during the review of the Milestones data, discuss what the process means to him/her, and confirm that the session should lead to an action plan.

Explore **R**eaction: The next stage is to explore reactions, emotions, and perceptions of the feedback. If the resident/fellow has completed a self-assessment (e.g., his/her own Milestones judgments), emotion and reaction are likely around areas of concordance and especially discordance between his/her impressions of his/her performance. These concordances, and especially the discordances, should be explored. The goal of this stage is to ensure the resident/fellow feels heard and that his/her views are respected, even if there is disagreement.

Explore **C**ontent: In this stage, explore how and what the resident/fellow understands about the feedback data. In this stage you want to ensure the resident/fellow fully understands the meaning of the feedback data and how he/she can use it for action plans and professional development. Helping the resident/fellow also understand how the various assessments are used to inform the Milestones may also be helpful.

Coach for Performance Change: In this last stage, the faculty member facilitates and engages the resident/fellow in “change talk” and the creation of an action plan.

One more observation of the R2C2 model – emotion, reaction, or misinterpretation can arise at any time during a session, so you may need to “loop back” to explore reactions or content.³⁶

EARLY LESSONS LEARNED ABOUT THE MILESTONES

Milestone department staff (ESH, SJH, LE) attended or visited multiple society meetings and institutions between January 2014 and December 2015. These encounters enabled high-level conversations on the benefits and challenges of the Milestones that have produced some early themes. While more systematic and rigorous research is underway, these conversations do provide an early signal and can help guide future evaluation activities, and in that spirit Table 5 provides a topline summary.

While it is simply too early to perform a systematic review, several studies on the early experience with the Milestones are also worth noting. One study investigating implementation of the first set of Internal Medicine Milestones found the Milestones improved faculty evaluations and feedback.⁴¹ One of the first national studies to find evidence of validity involved the first-year experience with the Emergency Medicine Milestones.⁴² Another emergency medicine study examined reliability, performed exploratory factor analysis and Milestones judgment distributions by training year across all emergency medicine residencies, showing encouraging results.⁴³ An earlier mixed-methods study involving program directors from 17 internal medicine programs found the Milestones to be useful for formative assessment, but faculty development was recognized as an important need to operationalize the Milestones.⁴⁴ However, a group of internal medicine programs found only modest differences in perceived quality of feedback by residents after implementation of the Milestones system.⁴⁵

One study in a large internal medicine program found that transitioning to a Milestones-based model produced a larger separation in the scores between post-graduate years (PGY)-1-3 classes and a wider use of a five-point scale on an end-of-rotation evaluation form.⁴⁶ Two studies found use of the Milestones was more effective than use of previous evaluation forms, finding better discrimination in ratings and a reduction in common rater errors.⁴⁷⁻⁸ On the other hand, a study of a Milestones “passport” intervention in an emergency medicine program found only modest increases in resident satisfaction with feedback.⁴⁹ Still another study reported that Milestones-based assessments for end-of-shift evaluations led to grade inflation in another emergency medicine program.⁵⁰ Using information technology is another growing theme of research. For example, a surgery program is using a smart phone application to complete a Zwisch Scale immediately after a procedure and link to the Milestones.⁵¹ All of these studies highlight the critical importance of ongoing, iterative, and rigorous research on the Milestones initiative.

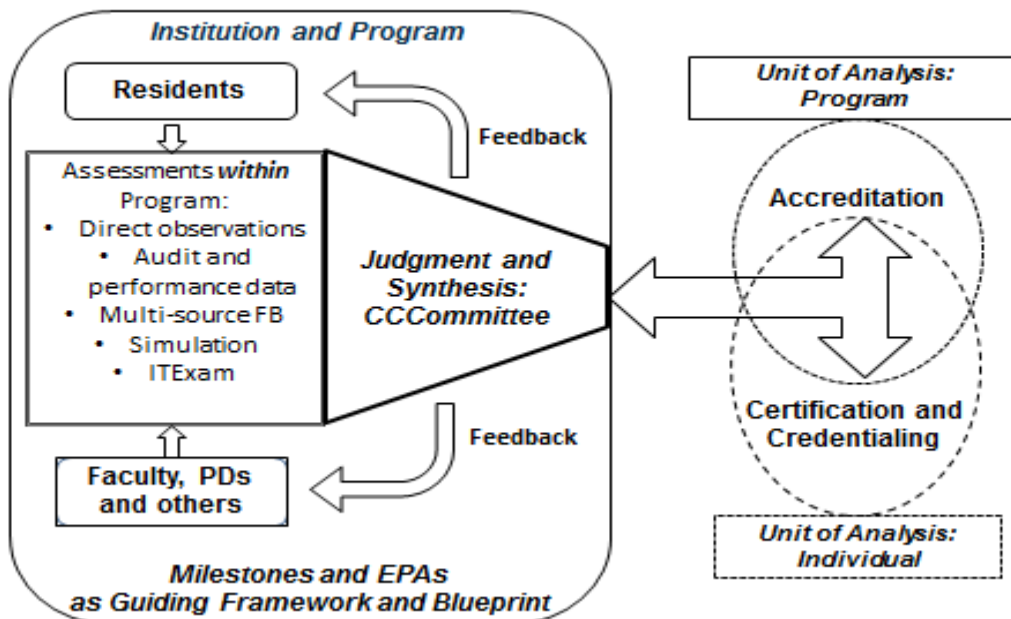
Table 6: Perceived Benefits and Challenges of Milestones Implementation

Benefits	Challenges
<ul style="list-style-type: none"> • Milestones and CCC process provides better feedback for residents and fellows • Milestones system catalyzing feedback for residents and fellows (e.g. for many first time formal feedback given) • Milestones provide a useful language for assessment and feedback • Milestones helping faculty develop shared mental model of competence. • Milestones have helped to identify curricular gaps • Milestone mapping onto curricular activities has facilitated better assessment • Milestones facilitating earlier identification of residents and fellows in difficulty • CCCs a useful mechanism to facilitate working with residents and fellows in difficulty • Milestones facilitate faculty development • Continuous quality improvement philosophy of system • Common framework of Milestones allows for more generalizability of medical education research on assessment in GME 	<ul style="list-style-type: none"> • Time and resources (“RVUs always win”) <ul style="list-style-type: none"> ○ 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 reporting Milestones as rotation evaluation form (e.g., “cognitive load”) • Need for faculty development • Assessment burden on faculty • Increasingly short faculty attending periods (e.g., 1-2 weeks) in a number of specialties <ul style="list-style-type: none"> ○ Insufficient faculty exposure to properly perform assessment • Challenging using a five level Milestone rubric for one year fellowships • Educational jargon and framing of language (select Milestones sets) • Lack of harmonization across the non-Patient Care and Medical Knowledge Milestones

MILESTONES AND THE NAS ASSESSMENT SYSTEM

Figure 2 provides an overview of how the Milestones inform the GME system. At the program level, residents/fellows are assessed routinely through a combination of assessment tools. These include: direct observations; global evaluation; audits and review of clinical performance data; Case Logs; multisource feedback from team members, including peers, nurses, patients, and families; simulation; in-service training examinations (ITEs); self-assessment; and others. Increasingly, the Milestones should be used as a guiding framework and “blueprint” for the curriculum and assessment of individual learner performance. Assessment tools should be selected intentionally to allow routine, frequent, formative feedback to the resident or fellow to affirm areas of successful performance and to highlight competencies on which they need to improve. The CCC should help to analyze and synthesize the assessment data, such as “quantitative” information from in-service exams and clinical performance audits, and “qualitative” information from observers and co-workers through surveys and direct observation.

Figure 2: Overview of Professional Self-Regulatory Assessment System in the U.S.



The figure also highlights the critical importance of active resident engagement in the assessment system. Effective group process via the Clinical Competence Committee (CCC) leads to better decisions and judgments about learner development. Professional self-regulation, represented in the U.S. by the ACGME (accreditation) and the American Board of Medical Specialties (certification) are the public facing entities of the system, but depend substantially on the programs for execution of standards. The bi-directional arrows signify the co-dependent relationships of all actors in the system.

Using the Milestones, the CCC should try to reach a consensus judgment regarding each resident's or fellow's performance. However, do not suppress minority opinions as research demonstrates these can help improve group decisions. The CCC provides those conclusions to the program director, who possesses the ultimate authority for determining the residents' or fellows' Milestones developmental level at least twice yearly. The unit of analysis for the ACGME is the program, and ACGME uses the national data as a mechanism to help improve training nationwide. The unit of analysis is the "individual" for certification and credentialing entities. Collectively, the goal of this system is help the entire medical education enterprise be accountable to the public for honest assessments of resident and fellow performance and truthful verification of their readiness to progress to unsupervised practice. As you can also see from the figure, the ACGME is involved with the certification boards around research on the effectiveness of the Milestones. Milestones data is not used to determine eligibility for certification by the boards. Together, the ACGME and the certification boards constitute the foundation of professional self-regulation in the United States.

USE OF MILESTONES BY THE ACGME

Residents' and fellows' performance on the Milestones will become a source of specialty-specific data for the specialty Review Committees to use in their continuous quality improvement efforts in assessing programs and for facilitating improvements to program curricula and resident/fellow assessment. The critical concept here is that the Milestones' primary purpose is to drive improvement in training programs and enhance the resident and fellow educational experience. The Milestones data, especially in this early phase, will be used as formative assessment of the quality of residency and fellowship programs. The Milestones will also be used by the ACGME to demonstrate accountability of the effectiveness of graduate medical education within ACGME-accredited programs in meeting the needs of the public over time.

The ACGME also fully expects that the Milestones will need to be revised. The Milestones are currently truly in Version 1.0. The ACGME will collect feedback through several mechanisms, including through its own research and evaluation activities, as well as collaborative research and evaluation with other stakeholders, and 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 when "Version 2.0" of the Milestones might roll out is yet to be determined, but it will be at least several years of learning and planning before the next versions of the specialty and subspecialty Milestones will be implemented.

Data Security and Milestones

The ACGME is dedicated to protecting the data collected from programs and residents. 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 CQI 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.

The ACGME fully understands the concerns of all stakeholders in the data security issue and will update the community as the system evolves. Stakeholders are encouraged to periodically check the ACGME website and review the ACGME's regular weekly *e-Communication* newsletter. The Frequently Asked Questions (FAQs) document on the Milestones section of the ACGME website is also a good resource, as the FAQs are updated, at a minimum, twice a year and any time a change occurs.

After a program submits Milestones data through ADS, a report is prepared (PDF) for each individual resident and fellow. The report includes all of the milestones reported for each resident during the previous reporting cycle. The program director can choose to print this report and use it as part of the semiannual evaluation with the resident or fellow. 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 will be available 10-14 days after the close of the reporting window.

How will the ACGME Evaluate the Milestones?

Evaluation of the Milestones iteratively and longitudinally will be essential in achieving the desired goals of the NAS. The ACGME fully understands the legitimate concerns raised by some in the educational community regarding the effects of the Milestones.⁵²⁻⁴ 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.⁵⁵⁻⁵⁹ Several specialties are already conducting pilot studies to gather information about the clarity, feasibility, acceptability, and performance characteristics of the Milestones. One advantage of the Milestones, compared to some of the evaluation tools currently used by individual programs, is that assessment data will be collected on thousands of residents and fellows, producing a sample that, over time, will make it possible to establish their reliability and validity on a national scale.

Validity frameworks will also guide the research work. The Messick framework is a useful framework in understanding validity:⁶⁰

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 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.

CONCLUSIONS

The overarching goal of all residency and fellowship 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 Clinical Competency Committees, were developed to enable and accelerate the transformation to a competency-based system after a difficult early period of implementation. The success of the Next Accreditation System 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 certification boards, Sponsoring Institutions and organizations, researchers, and policy makers.

Appendix 1: Additional Educational and Assessment Resources

Note: The listing of these resources does not constitute endorsement by the ACGME. Readers are encouraged to carefully consider the pros and cons of each resource.

1. Helpful Websites: Practice-based Learning and Improvement/Systems-based Practice
 - Healthcare Improvement Skills Center
 - Six web-based modules with assessment (<https://www.improvementskills.org/index.cfm>)
 - Cost for 6 modules and 10 CME credits is \$75/person
 - Institute for Healthcare Improvement Open School
 - Online courses in quality and patient safety (<http://www.ihl.org/IHI/Programs/IHIOpenSchool/>)
 - HRSA Quality Improvement Module
 - This website has a number of useful tools and is free. (<http://www.hrsa.gov/quality/toolbox/methodology/qualityimprovement/>)
 - Mayo Clinic Quality Academy
 - This is a modular program that covers all the basics and also includes several module on evidence-based medicine (<http://qiresources.mayo.edu/>)
 - Association of American Medical Colleges
 - Teaching for Quality Initiative (Te4Q)
 - Available onsite faculty development
 - Website: <https://www.aamc.org/initiatives/cei/te4q/>
 - Requires an institutional fee
 - Johns Hopkins University
 - Guided Care (PCMH): <http://www.guidedcare.org/module-listing.asp>
 - Multiple online training modules
 - Fee required (\$15 per module)
 - MD Content
 - Covers management, finance and liability (www.mdcontent.com)
 - Small fee for use
 - Choosing Wisely (ABIM Foundation)
 - Provides five diagnostic and/or therapeutic interventions that have little to no benefit to most patients across multiple specialties.
 - Website: <http://www.choosingwisely.org/>

- High Value Cost Conscious Care
 - Initiative of ACP: <https://hvc.acponline.org/>
 - Helpful toolkit for programs
- Costs of Care
 - Helpful resources: <http://www.costsofcare.org/>
- National Patient Safety Foundation
 - Patient safety curriculum
 - Full 10 module curriculum \$399/person
 - Certificate program also available
 - <http://www.npsf.org/?page=pscurriculum>
- World Health Organization (WHO) Patient Safety Curriculum Guide
 - Provides a number of free resources, including guide and slides
 - Website: <http://www.who.int/patientsafety/education/curriculum/en/>

2. Patient and Multisource Feedback Surveys:

- Physician Achievement Review (PAR) Program (Canada):
 - <http://par-program.org/par/for-physicians/how-par-works/how-par-is-scored/par-questionnaires/>
 - This website contains a number of assessment tools that have been studied for validity.
- CAHPS Surveys (Patient surveys for multiple settings):
 - <https://cahps.ahrq.com>
 - This site contains a number of CAHPS patient surveys depending on site of care delivery. Note –CAHPS or a CAHPS-like instrument cannot be used in the hospital because of Medicare. Institutions' QI department can assist if a program wants to use CAHPS data in the inpatient setting.
- Assessment of Professional Behaviors
 - <https://www.mededportal.org/publication/9902>
 - MSF tool developed by the National Board of Medical Examiners. The program is no longer available through the NBME but the assessment instrument plus some supporting material are on the MedEdPortal site. Users will need to create an account on MedEdPortal, and there is no charge.

3. Teaching Physical Diagnosis:

- The Stanford Medicine 25: This website provides helpful teaching materials and videos on 25 evidence-based physical exam maneuvers.
- Website: <http://stanfordmedicine25.stanford.edu/index.html>

Appendix 2: Annotated Bibliography of Selected Articles

Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system – rationale and benefits. *NEJM*. 2012; 366(11): 1051-1056.

The article discusses the Accreditation Council for Graduate Medical Education's (ACGME) Next Accreditation System (NAS) which is scheduled for phased implementation in July 2013. The aims of the NAS include to enhance the peer-review system and accelerate the movement toward accreditation on the basis of educational outcomes. An overview of the major problems with the graduate medical education (GME) when ACGME was established in 1981 is given.

Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, Horsley T. Toward a definition of competency-based education in medicine: a systematic review of published definitions. *Med Teach*. 2010; 32(8): 631-637.

Background: Competency-based education (CBE) has emerged in the health professions to address criticisms of contemporary approaches to training. However, the literature has no clear, widely accepted definition of CBE that furthers innovation, debate, and scholarship in this area.

Aim: To systematically review CBE-related literature in order to identify key terms and constructs to inform the development of a useful working definition of CBE for medical education.

Methods: We searched electronic databases and supplemented searches by using authors' files, checking reference lists, contacting relevant organizations and conducting Internet searches. Screening was carried out by duplicate assessment, and disagreements were resolved by consensus. We included any English- or French-language sources that defined competency-based education. Data were analyzed qualitatively and summarized descriptively.

Results: We identified 15,956 records for initial relevancy screening by title and abstract. The full text of 1,826 records was then retrieved and assessed further for relevance. A total of 173 records were analyzed. We identified 4 major themes (organizing framework, rationale, contrast with time, and implementing CBE) and 6 sub-themes (outcomes defined, curriculum of competencies, demonstrable, assessment, learner-centered and societal needs). From these themes, a new definition of CBE was synthesized.

Conclusion: This is the first comprehensive systematic review of the medical education literature related to CBE definitions. The themes and definition identified should be considered by educators to advance the field.

Frank JR, Snell LS, Ten Cate O, Holmboe ES, Carraccio C, Swing SR, Harris P, Glasgow NJ, Campbell C, Dath D, Harden RM, Iobst W, Long DM, Mungroo R, Richardson DL, Sherbino J, Silver I, Taber S, Talbot M, Harris KA. Competency-based medical education: theory to practice. *Med Teach*. 2010; 32(8): 638-645.

Although competency-based medical education (CBME) has attracted renewed interest in recent years among educators and policy-makers in the health care professions, there is little agreement on many aspects of this paradigm. We convened a unique partnership – the International CBME Collaborators – to examine conceptual issues and current debates in CBME. We engaged in a multi-stage group process and held a consensus conference with the aim of reviewing the scholarly literature of competency-based medical education, identifying controversies in need of clarification, proposing definitions and concepts that could be useful to educators across many jurisdictions, and exploring future directions for this approach to preparing health professionals. In this paper, we describe the evolution of CBME from the outcomes movement in the 20th century to a renewed approach that, focused on accountability and curricular outcomes and organized around competencies, promotes greater learner-centeredness and de-emphasizes time-based curricular design. In this paradigm, competence and related terms are redefined to emphasize their multi-dimensional, dynamic, developmental, and contextual nature. CBME therefore has significant implications for the planning of medical curricula and will have an important impact in reshaping the enterprise of medical education. We elaborate on this emerging CBME approach and its related concepts, and invite medical educators everywhere to enter into further dialogue about the promise and the potential perils of competency-based medical curricula for the 21st century.

Irby DM, Cooke M, O'Brien BC. Calls for reform of medical education by the Carnegie Foundation 2010. Acad Med. 2010 Feb; 85(2): 220-7.

The Carnegie Foundation for the Advancement of Teaching, which in 1910 helped stimulate the transformation of North American medical education with the publication of the Flexner Report, has a venerated place in the history of American medical education. Within a decade following Flexner's report, a strong scientifically oriented and rigorous form of medical education became well established; its structures and processes have changed relatively little since. However, the forces of change are again challenging medical education, and new calls for reform are emerging. In 2010, the Carnegie Foundation will issue another report, *Educating Physicians: A Call for Reform of Medical School and Residency* that calls for (1) standardizing learning outcomes and individualizing the learning process, (2) promoting multiple forms of integration, (3) incorporating habits of inquiry and improvement, and (4) focusing on the progressive formation of the physician's professional identity. The authors, who wrote the 2010 Carnegie report, trace the seeds of these themes in Flexner's work and describe their own conceptions of them, addressing the prior and current challenges to medical education as well as recommendations for achieving excellence. The authors hope that the new report will generate the same excitement about educational innovation and reform of undergraduate and graduate medical education as the Flexner Report did a century ago.

Ten Cate O, Scheele F. Competency-based postgraduate training: Can we bridge the gap between theory and clinical practice? Acad Med. 2007 June; 82(6): 542–547.

The introduction of competency-based postgraduate medical training, as recently stimulated by national governing bodies in Canada, the United States, the United Kingdom, The Netherlands, and other countries, is a major advancement, but at the same time it evokes critical issues of curricular implementation. A source of concern is the translation of general competencies into the practice of clinical teaching. The authors observe confusion around the term *competency*, which may have adverse effects when a teaching and assessment program is to be designed. This article aims to clarify the competency terminology. To connect the ideas behind a competency framework with the work environment of patient care, the authors propose to analyze the critical activities of professional practice and relate these to predetermined competencies. The use of entrustable professional activities (EPAs) and statements of awarded responsibility (STARs) may bridge a potential gap between the theory of competency-based education and clinical practice. EPAs reflect those activities that together constitute the profession. Carrying out most of these EPAs requires the possession of several competencies. The authors propose not to go to great lengths to assess competencies as such, in the way they are abstractly defined in competency frameworks but, instead, to focus on the observation of concrete critical clinical activities and to infer the presence of multiple competencies from several observed activities. Residents may then be awarded responsibility for EPAs. This can serve to move toward competency-based training, in which a flexible length of training is possible and the outcome of training becomes more important than its length.

Ten Cate O, Snell L, Carraccio C. Medical competence: The interplay between individual ability and the health care environment. Med Teach. 2010; 32: 669-675.

Competency-based education in the health care professions has become a prominent approach to postgraduate training in Canada, the Netherlands, the United Kingdom, the United States, and many other countries. Competency frameworks devised at national and international levels have been well received, and in many cases mandated, by governing bodies. However, the teaching and assessment of competencies pose questions of practicality, validity, and reliability. In this article we propose that competence and competencies be approached in the context of the particular clinical environment, such that the assessment of competence is tied to a trainee's performance of essential clinical activities that define the profession. Competence is implicit in the eventual entrustment of trainees to perform these professional activities. Competencies and "entrustable professional activities" (EPAs) relate to each other as two dimensions of a grid in which each EPA can be mapped back to a number of competencies. This backward visioning from EPAs to competencies is proposed as a guide to curriculum planning and assessment. The authors discuss experiences with this conceptual model in research, curriculum development and learner assessment.

Holmboe ES, Sherbino J, Long DM, Swing SR, & Frank JR. The role of assessment in competency-based medical education. *Med Teach.* 2010; 32(8): 676-682.

Competency-based medical education (CBME), by definition, necessitates a robust and multifaceted assessment system. Assessment and the judgments or evaluations that arise from it are important at the level of the trainee, the program, and the public. When designing an assessment system for CBME, medical education leaders must attend to the context of the multiple settings where clinical training occurs. CBME further requires assessment processes that are more continuous and frequent, criterion-based, developmental, work-based where possible, use assessment methods and tools that meet minimum requirements for quality, use both quantitative and qualitative measures and methods, and involve the wisdom of group process in making judgments about trainee progress. Like all changes in medical education, CBME is a work in progress. Given the importance of assessment and evaluation for CBME, the medical education community will need more collaborative research to address several major challenges in assessment, including “best practices” in the context of systems and institutional culture and how to best to train faculty to be better evaluators. Finally, we must remember that expertise, not competence, is the ultimate goal. CBME does not end with graduation from a training program, but should represent a career that includes ongoing assessment.

Holmboe ES, Ward DS, Reznick RK, Katsufraakis PJ, Leslie KM, Patel VL, Ray DD, Nelson EA. Faculty development in assessment: the missing link in competency-based medical education. *Acad Med.* 2011 Apr; 86(4): 460-7.

As the medical education community celebrates the 100th anniversary of the seminal Flexner Report, medical education is once again experiencing significant pressure to transform. Multiple reports from many of medicine’s specialties and external stakeholders highlight the inadequacies of current training models to prepare a physician workforce to meet the needs of an increasingly diverse and aging population. This transformation, driven by competency based medical education (CBME) principles that emphasize the outcomes, will require more effective evaluation and feedback by faculty. Substantial evidence suggests, however, that current faculty are insufficiently prepared for this task across both the traditional competencies of medical knowledge, clinical skills, and professionalism and the newer competencies of evidence-based practice, quality improvement, interdisciplinary teamwork, and systems. The implication of these observations is that the medical education enterprise urgently needs an international initiative of faculty development around CBME and assessment. In this article, the authors outline the current challenges and provide suggestions on where faculty development efforts should be focused and how such an initiative might be accomplished. The public, patients, and trainees need the medical education enterprise to improve training and outcomes now.

Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011; 33: 478-85.

In assessment a considerable shift in thinking has occurred from assessment of learning to assessment for learning. This has important implications for the conceptual framework from which to approach the issue of assessment, but also with respect to the research agenda. The main conceptual changes pertain to programmes of assessment. This has led to a broadened perspective on the types of construct assessment tries to capture, the way information from various sources is collected and collated, the role of human judgement and the variety of psychometric methods to determine the quality of the assessment. Research into the quality of assessment programmes, how assessment influences learning and teaching, new psychometric models and the role of human judgement is much needed.

References:

1. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system--rationale and benefits. *N Engl J Med*. 2012 Mar 15;366(11):1051-6.
2. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. An antidote to over specification in the education of medical specialists. *Health Affairs*. 2002; 21: 103-111.
3. IOM (Institute of Medicine). 2014. *Graduate medical education that meets the nation's health needs*. Washington, DC: The National Academies Press.
4. Clinical Competency Committee Guidebook. Accessed at <http://www.acgme.org/acgmeweb/Portals/0/ACGMEClinicalCompetencyCommitteeGuidebook.pdf>.
5. Weiss KB, Bagian JP, Nasca TJ. The clinical learning environment: the foundation of graduate medical education. *JAMA*. 2013 Apr 24;309(16):1687-8.
6. Sullivan RL. 1995. *The Competency-Based Approach to Training*. Strategy Paper No 1. JHPIEGO Corporation: Baltimore, Maryland.
7. Elam S. *Performance-based teacher education: What is the state of the art?* Washington: American Association of Colleges for Teacher Education. 1971.
8. McGaghie WC, Lipson L. Competency-based curriculum development in medical education: An introduction: World Health Organization Geneva; 1978.
9. Carraccio C, Wolfstahl SD, Englander R, Ferentz K, Martin C. Shifting paradigms: from Flexner to competencies. *Acad Med*. 2002; 77: 361-67.
10. Thurman GK, Sanders MK. Competency-based education versus traditional education: a comparison of effectiveness. *Radiol Technol*. 1987; 59: 164-9
11. Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, Harris P, Glasgow NJ, Campbell C, Dath D, Harden RM, Iobst W, Long DM, Mungroo R, Richardson DL, Sherbino J, Silver I, Taber S, Talbot M, Harris KA. Competency-based medical education: theory to practice. *Med Teach*. 2010;32(8):638-45
12. Ten Cate O. The false dichotomy of quality and quantity in the discourse around assessment in competency-based education. *Adv in Health Sci Educ*. 2014 (online): DOI 10.1007/s10459-014-9527-3
13. Holmboe ES, Sherbino J, Long DM, Swing SR, Frank JR. The Role of Assessment in Competency-based Medical Education. *Med Teach*. 2010;32(8):676-82.
14. Kogan JR and Holmboe ES. Realizing the Promise and Importance of Performance-based Assessment. *Teach Learn Med* 2013;25 Suppl 1:S68-74.
15. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ*. 2007; 41: 1124-30.
16. Swing SR, Beeson MS, Carraccio C, Coburn M, Iobst W, Selden NR, Stern PJ, Vydareny K. Educational milestone development in the first 7 specialties to enter the next accreditation system. *J Grad Med Educ*. 2013 Mar;5(1):98-106.
17. Coburn M, Amling C, Bahnson RR, Dahm P, Kerfoot BP, King L, Lane B, Ritchey ML, Scales CD Jr, Sundaram CP, Swing SR. Urology milestones. *J Grad Med Educ*. 2013 Mar;5(1 Suppl 1):79-98. doi: 10.4300/JGME-05-01s1-07. Similar articles

18. Vydareny KH, Amis ES Jr, Becker GJ, Borgstede JP, Bulas DI, Collins J, Davis LP, Gould JE, Itri J, Laberge JM, Meyer L, Mezwa DG, Morin RL, Nestler SP, Zimmerman R. Diagnostic radiology milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):74-8. doi: 10.4300/JGME-05-01s1-01.
19. Carraccio C, Benson B, Burke A, Englander R, Guralnick S, Hicks P, Ludwig S, Schumacher D, Vasiliadis J. milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):59-73. doi: 10.4300/JGME-05-01s1-06.
20. Stern PJ, Albanese S, Bostrom M, Day CS, Frick SL, Hopkinson W, Hurwitz S, Kenter K, Kirkpatrick JS, Marsh JL, Murthi AM, Taitsman LA, Toolan BC, Weber K, Wright RW, Derstine PL, Edgar L. Orthopaedic surgery milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):36-58. doi: 10.4300/JGME-05-01s1-05. S
21. Selden NR, Abosch A, Byrne RW, Harbaugh RE, Krauss WE, Mapstone TB, Sagher O, Zipfel GJ, Derstine PL, Edgar L. Neurological surgery milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):24-35. doi: 10.4300/JGME-05-01s1-04.
22. Iobst W, Aagaard E, Bazari H, Brigham T, Bush RW, Caverzagie K, Chick D, Green M, Hinchey K, Holmboe E, Hood S, Kane G, Kirk L, Meade L, Smith C, Swing S. Internal medicine milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):14-23. doi: 10.4300/JGME-05-01s1-03.
23. Beeson MS, Carter WA, Christopher TA, Heidt JW, Jones JH, Meyer LE, Promes SB, Rodgers KG, Shayne PH, Wagner MJ, Swing SR. Emergency medicine milestones. *J Grad Med Educ.* 2013 Mar;5(1 Suppl 1):5-13. doi: 10.4300/JGME-05-01s1-02.
24. Hicks PJ, Schumacher DJ, Benson BJ, Burke AE, Englander R, Guralnick S, Ludwig S, Carraccio C. The pediatrics milestones: conceptual framework, guiding principles, and approach to development. *J Grad Med Educ.* 2010 Sep;2(3):410-8. doi: 10.4300/JGME-D-10-00126.1.
25. Green ML, Aagaard EM, Caverzagie KJ, Chick DA, Holmboe ES, Kane G, Smith CD, Iobst W. Charting the Road to Competence: Developmental Milestones for Internal Medicine Residency Training. *J Graduate Med Educ.* 2009; 1: 5-20.
26. Holmboe ES, Yamazaki K, Edgar L, Conforti L, Yaghmour N, Miller R, Hamstra SJ. Reflections on the first 2 years of milestone implementation. *J Grad Med Educ.* 2015; Published online. DOI: <http://dx.doi.org/10.4300/JGME-07-03-43>.
27. Fuchs V. The Service Economy. Pp. 12. National Bureau of Economic Research. 1968. Accessed November 30, 2014 at <http://www.nber.org/books/fuch68-1>
28. Bate P and Robert G. Experience-based design: from redesigning the system around the patient to co-designing services with the patient. *BMJ Qual Saf.* 2006; 15: 307-310.
29. Freire K and Sangiorgi D. Service design and Healthcare innovation: from consumption to co-production and co-creation. Paper Nordic Service Design Conference, Linkoping, Sweden 2010. Accessed on November 30, 2014 at <http://www.servdes.org/pdf/freire-sangiorgi.pdf>.
30. Sabadossa KA and Batalden PB. The interdependent roles of patients, families and professionals in cystic fibrosis: a system for the coproduction of healthcare and its improvement. *BMJ Qual Saf* 2014; 23:i90-i94 doi: 10.1136/bmjqs-2013-002782.

31. Normann, R. Reframing Business: When the Map Changes the Landscape. Wiley Publishing. 2001. London.
32. Ostrom E. Crossing the Great Divide: Coproduction, Synergy, and Development. World Development, Vol. 24, No. 6, pp. 1073-1087, 1996.
33. Garn HA, Flax MJ, Springer M, Taylor JB. Models for Indicator Development: A Framework for Policy Analysis. An Urban Institute Paper (1206-17). Urban Institute. April, 1976. Washington, DC
34. Wagner EH, Austin BT, Von Korff M. Improving outcomes in chronic illness. Manag Care Q. 1996; 4:12–25.
35. Molloy E and Boud D. Changing conceptions of feedback. In Feedback in Higher and Professional Education. Boud D and Molly E, editors. Routledge. 2013. New York.
36. Sargeant J, Lockyer J, Mann K, Holmboe E, Silver I, Armson H, Driessen E, MacLeod T, Yen W, Ross K, Power M. Facilitated Reflective Performance Feedback: Developing an Evidence- and Theory-Based Model That Builds Relationship, Explores Reactions and Content, and Coaches for Performance Change (R2C2). Acad Med. 2015 Dec; 90(12):1698-706.
37. Schuwirth LWT, Van der Vleuten CPM. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011; 33: 478-85.
38. Norcini J, Anderson B, Bollela V, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. Med Teach. 2011; 33:206-14.
39. Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. Med Educ. 2012; 46: 28-37.
40. Skeff K and Stratos G. Feedback. Stanford Clinical Teaching Program. Accessed at http://sfdc.stanford.edu/clinical_teaching.html January 24, 2015.
41. Nabors C, Peterson SJ, Aronow W, Sute S, Mumtaz A, Delorenzo L, Chandy D, Lehman S, Frishman WH, Holmboe ES. Operationalizing the Internal Medicine Milestones - An Early Status Report. J Grad Medical Educ. 2013; 5: 130-137.
42. Korte RC, Beeson MS, Russ CM, Carter WA; Emergency Medicine Milestones Working Group, Reisdorff EJ. The emergency medicine milestones: a validation study. Acad Emerg Med. 2013 Jul;20(7):730-5.
43. Beeson MS, Holmboe ES, Korte RC, Nasca TJ, Brigham T, Russ CM, Whitley CT, Reisdorf EJ. Initial Validity Analysis of the Emergency Medicine Milestones. Acad Emerg Med. 2015 Jul; 22(7):838-44.
44. Aagaard E, Kane GC, Conforti L, Hood S, Caverzagie KJ, Smith C, Chick DA, Holmboe ES, Iobst WF. Early feedback on the use of the internal medicine reporting milestones in assessment of resident performance. J Grad Med Educ. 2013;5(3):433-8
45. Angus S, Moriarty J, Nardino RJ, Chmielewski A, Rosenblum MJ. Internal Medicine Residents' Perspectives on Receiving Feedback in Milestone Format. J Graduate Med Educ. 2015; 7:220-224.
46. Friedman KA, Balwan S, Cacace F, Katona K, Sunday S, Chaudhry S. Impact on house staff evaluation scores when changing from a Dreyfus- to a Milestone-

- based evaluation model: one internal medicine residency program's findings. *Med Educ Online*. 2014; 19:251-85.
47. Bartlett KW, Whicker SA, Bookman J, Narayan AP, Staples BB, Hering H, McGann KA. Milestones-based ratings are superior to Likert-type assessments in illustrating trainee progression. *J Grad Med Educ*. 2015; 7: 75-80.
 48. Raj, JM and Thorn PM. A faculty development program to reduce rater error o milestones-based assessments. *J Grad Med Educ*. 2014; 6: 680-85.
 49. Yarris LM, Jones D, Kornegay JG, Hansen M. The milestones passport: a learner centered application of the Milestone framework to prompt real-time feedback in the emergency department. *J Grad Med Educ*. 2014; 6: 555-60.
 50. Dehon E, Jones J, Puskarich M, Sandifer JP, Sikes K. Use of Emergency Medicine Milestones as Items on End-of-Shift Evaluations Results in Overestimates of Residents' Proficiency Level. *J Graduate Med Educ*. 2015; 7: 192-196.
 51. George BC, Teitelbaum EN, Meyerson SL, Schuller MC, DaRosa DA, Petrusa ER, Petito LC, Fryer JP. Reliability, validity, and feasibility of the Zwisch scale for the assessment of intraoperative performance. *J Surg Educ*. 2014; 71(6):e90-6.
 52. Norman G, Norcini J, Bordage G. Competency-Based Education: Milestones or Millstones? *J Graduate Medical Education*. 2014; DOI: <http://dx.doi.org/10.4300/JGME-D-13-00445.1> [epub ahead of print]
 53. Pangaro LN. Two cheers for the milestones. *J Grad Med Educ*. 2015; 7: 4-6.
 54. Dewan M, Manning J, Satish U. The new milestones: do we need to take a step back to go a mile forward? *Acad Psychiatry*. 2015; 39(2):147-50.
 55. Rogers PJ. Implications of complicated and complex characteristics for key tasks in evaluation. Pps. 33-53. In *Evaluating the Complex: Attribution, Contribution and Beyond*. Forss K, Marra M, Schwartz R, eds. Transaction Publishers. New Brunswick, NJ. 2011.
 56. Pawson R. *The Science of Evaluation: A Realist Manifesto*. Sage Pub. London. 2013.
 57. Campbell NC, Murray E, Darbyshire J, et al. Designing and evaluating complex interventions to improve care. *BMJ*. 2007; 334: 455-9.
 58. Medical Research Council (United Kingdom). *Developing and evaluating complex interventions: new guidance*. Accessed at www.mrc.ac.uk on Jan. 5, 2014.
 59. Pawson R and Tilley N. *Realistic Evaluation*. Sage Publications. London. 1997.
 60. Cook DA, Beckman TJ. Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *Am J Med*. 2006; 119: 166.e7-166.e16.