



How Residents Can Use Milestones Data

Readers' Guide

The abstracts in this document are organized by year of publication (in descending order). Readers with a particular interest in a subject may use the search function in the PDF to quickly find topics of their choosing. Please note that since the following content comes from a variety of sources, there may be variations in style and spelling. For official ACGME terminology and terminology usage, refer to the ACGME Glossary of Terms.

Background:

The use of Milestones is not limited to evaluating residents' progress, but it can also help identify trends and areas of improvement for residents. The data can also be used to spot trends and identify underperformers, not only within a program but also between institutions. Residents can use the Milestones to track their progress in each of the six Core Competencies and identify areas where they may need to improve. By comparing their own performance to the Milestones, residents can create targeted goals for self-improvement. Residents can use Milestones data to receive feedback from their faculty and program directors. By sharing their Milestones assessments with their mentors, residents can identify areas of strength and weakness and work with their mentors to create a plan for improvement.

Residents can also use Milestones data when selecting residency and fellowship programs. By reviewing Milestones data for programs they are considering, residents can get a sense of how well programs are preparing their trainees and whether they align with the resident's own career goals and interests. Residents can also use Milestones data to plan their future careers. By identifying areas of strength and interest, residents can begin to explore opportunities within their specialty and develop a plan for continuing education and professional development.

What's in the literature:

- Using predictive analytics (predictive probability values [PPVs]) and early Milestone ratings to identify the percentage of learners who will not achieve a target Milestone level by the time of graduation
- How to spot trends and identify underperformers
- How residents can track their progress and use the data to make informed decisions about their learning and professional development

Predicting Performance at Graduation From Early ACGME Milestone Ratings: Longitudinal Learning Analytics in Professionalism and Communication in Vascular Surgery

Smith BK, Yamazaki K, Luman A, Tekian A, Holmboe E, Mitchell EL, Park YS, Hamstra SJ. *J Surg Educ.* 2022 Sep 28:S1931-7204(22)00234-3. doi: 10.1016/j.jsurg.2022.09.002. Epub ahead of print. PMID: 36182635.

OBJECTIVE:

Program directors in surgical disciplines need more tools from the ACGME to help them use Milestone ratings to improve trainees' performance. This is especially true in competencies that are notoriously difficult to measure, such as professionalism (PROF) and interpersonal and communication skills (ICS). It is now widely understood that skills in these two areas have direct impact on patient care outcomes. This study investigated the potential for generating early predictors of final Milestone ratings within the PROF and ICS competency categories.

DESIGN:

This retrospective cohort study utilized Milestone ratings from all ACGME-accredited vascular surgery training programs, covering residents and fellows who completed training in June 2019. The outcome measure studied was the rate of achieving the recommended graduation target of Milestone Level 4 (possible range: 1-5), while the predictors were the Milestone ratings attained at earlier stages of training. Predictive probability values (PPVs) were calculated for each of the 3 PROF and two ICS sub-competencies to estimate the probability of trainees not reaching the recommended graduation target based on their previous Milestone ratings.

SETTING:

All ACGME-accredited vascular surgery training programs within the United States.

PARTICIPANTS:

All trainees completing a 2 year vascular surgery fellowship (VSF) in June 2019 (n = 119) or a 5 year integrated vascular surgery residency (IVSR) in June 2019 (n = 52) were included in the analyses.

RESULTS:

The overall rate of failing to achieve the recommended graduation target across all PROF and ICS sub-competencies ranged from 7.7% to 21.8% of all trainees. For trainees with a Milestone rating at ≤ 2.5 with 1 year remaining in their training program, the predictive probability of not achieving the recommended graduation target ranged from 37.0% to 71.5% across sub-competencies, with the highest risks observed under PROF for "Administrative Tasks" (71.5%) and under ICS for "Communication with the Healthcare Team" (56.7%).

CONCLUSIONS:

As many as 1 in 4 vascular surgery trainees did not achieve the ACGME vascular surgery Milestones targets for graduation in at least one of the PROF and ICS sub-competencies. Biannual ACGME Milestone assessment ratings of PROF and ICS during early training can be used to predict achievement of competency targets at time of graduation. Early clues to problems in PROF and ICS enable programs to address potential deficits early in training to ensure competency in these essential non-technical skills prior to entering unsupervised practice.

Do Pediatric Emergency Medicine Fellows Meet the Milestone Targets for Graduation?

Roskind, C. G., Leonard, K., Baghdassarian, A., Kou, M., Lévasseur, K., Rose, J., Shabanova, V., Vu, T., Zuckerbraun, N. S., & Langhan, M. L. *Academic Pediatrics*. 2020;20(7):e35-e36.

ABSTRACT:

The ACGME Milestone Project is a competency-based assessment tool. Subcompetencies (SC) are scored on a 5-point scale, and level 4 is recommended for graduation. The 2018 Milestones Report found that across subspecialties, not all graduates are attaining a level 4 for every SC. To describe the number of pediatric emergency medicine (PEM) fellows who achieve level 4 in all 23 SC at graduation and to identify SC and predictive factors where a level 4 is not achieved. This is a multicenter, retrospective cohort study of PEM fellows. Program directors provided de-identified milestone reports from 2015- 2018. Descriptive analysis of milestone scores at graduation was performed. Demographics were compared between fellows who did and did not meet level 4 at graduation for each SC. Sub- analyses assessed differences in residency and first year milestone scores and the rate of milestone attainment between fellows who did and did not attain level 4 at graduation. Data from 48 PEM fellowship programs yielded graduation scores for 392 fellows (62% of total). 87% completed pediatric residency and 60% were female. Residency scores were available for 45 fellows. There were no SC in which all fellows attained at least level 4 at graduation; the range of fellows scoring < level 4 per SC was 7-39%. (Table 1) 67% of fellows did not attain level 4 on one or more of the 23 SC at graduation. While some fellows failed to attain a level 4 on up to all 23 SC, 26% failed to meet level 4 on only 1 or 2 SC. In 19/23 SC, residency and/or first year milestones scores were significantly lower for those who did not attain level 4 at graduation compared to those who did (mean difference 0.74 points). Those who did not attain level 4 at graduation had a significantly faster rate of improvement in milestone scores for 10/23 SC compared to those who did attain level 4. In our sample, 67% of PEM fellows did not attain level 4 for at least 1 of the 23 SC at graduation. Low scores during residency or early in fellowship may predict difficulty in meeting level 4 by fellowship completion.

Cytopathology Milestones: Can You Get to Level 5?

Dyhdalo KS, Oshilaja O, Chute DJ, Booth CN, Suchy P, Smith K, Procop GW, Reynolds JP. *J Am Soc Cytopathol*. 2020 Mar 30. pii: S2213-2945(20)30046-6. doi: 10.1016/j.jasc.2020.03.002. [Epub ahead of print]

INTRODUCTION:

ACGME Milestones describe 6 areas of proficiency, indicating readiness for practice. Each is divided into 5 levels of mastery; Level 1 (new trainees) through Levels 4 (graduation) and 5 (aspirational). Milestones reporting began Spring 2016. We used Milestones to assess graduated fellows.

MATERIALS AND METHODS:

We conducted phone interviews with previous fellows and collected demographic information including practice setting. We asked graduates if they fulfilled each example of mastery and recorded their answers.

RESULTS:

A total of 22 fellows graduated from 2010 to 2017; 15 responded (10 academic, 5 private). Milestones in which nearly all respondents performed well (Level 4+) were: PC1, MK1, SBP2, SBP4, PROF1-4, ICS1-3. Some were more challenging (PC2, MK2, SBP1/3/5, PBL1). For PC2, 2 respondents achieved Level 1 (did not perform fine-needle aspirations). For MK2, 2 respondents achieved Level 1 (did not evaluate Papanicolaou). For SBP1, 80% in private practice achieved Level 5; 50% in academics achieved Level 3. For SBP3, 80% in private practice achieved Level 4+; 100% in academics achieved maximum Level 2. For SBP5, 60% of all respondents achieved maximum Level 3; only 1 achieved Level 5.

CONCLUSIONS:

Many Milestones are attainable. Eleven of 18 yielded Level 4+ from most respondents. Three (PC2, MK1, MK2) yielded rare Level 1 due to scope of practice. Others (SBP1, SBP3) reflect more of an all-or-nothing phenomenon. For SBP5, most respondents achieved Level 3; only 1 achieved Level 5. Some Milestones are highly dependent on practice setting, and others remain aspirational.

Using an Entrustable Professional Activity to Assess Consultation Requests Called on an Internal Medicine Teaching Service

Kang AJ, Gielissen K, Windish D. MedEdPORTAL. 2019 Nov 22;15:10854. doi: 10.15766/mep_2374- 8265.10854.

INTRODUCTION:

The Accreditation Council for Graduate Medical Education's milestones require internal medicine residents to have competency in calling consults. Based on a literature review, we developed an Entrustable Professional Activity (EPA) to delineate the knowledge, skills, and attitudes required for a consultation request and, building on the EPA, implemented an assessment instrument to provide feedback to interns calling consultation requests and assess the quality of their consult questions and the level of supervision required in performing this milestone.

METHODS:

Assessments were done on internal medicine inpatient teaching services. Consultation requests were performed by interns and observed by residents using the assessment instrument. Feedback was provided to the interns. Interns then completed a self-reflection instrument based on the feedback.

RESULTS:

Twenty-six paired observations were collected over three 1-month rotations. There was a moderate positive correlation ($r = .43$) comparing resident and intern responses to how they felt about the intern's ability to make a consultation request. There was a strong positive correlation ($r = .65$) comparing resident opinion of how strong the intern's ability in calling a consult to how well the consult question used the PICO (patient, intervention, comparators, outcomes of interest) framework. Twenty-five out of 28 interns (89%) said they would make a change during their next consultation request due to feedback from their resident.

DISCUSSION:

Our EPA-based assessment instrument provided an opportunity to give interns feedback and to assess the quality of the consultation requests they made.

Family Medicine Residency Graduates' Preparation for Quality Improvement Leadership

Lichkus J, Fang B, Peterson LE. J Grad Med Educ. 2019 Oct;11(5):558-564. doi: 10.4300/JGME-D- 18-01060.1.

BACKGROUND:

Training in quality improvement (QI) is a standard component of family medicine residency education. Graduating family medicine residents' ability to lead QI initiatives is unknown.

OBJECTIVE:

We assessed the preparedness of graduating family medicine residents to lead QI projects and to identify factors that may increase such readiness.

METHODS:

Milestone data for all graduating family medicine residents were linked to a practice demographic questionnaire completed by the same residents who registered for the American Board of Family Medicine certification examination between 2014 and 2017. The change in self-assessed QI preparedness over time and its association with faculty-assigned milestone ratings were examined using descriptive and regression analyses.

RESULTS:

The questionnaire had a 100% response rate (12 208 responded). Between 2014 and 2017, the percentage of residents who self-reported being "extremely" or "moderately" prepared to lead QI projects increased from 72.7% (2208 of 3038) to 75.8% (2434 of 3210, $P = .009$). Self-reported QI team leadership was associated with 93% higher odds of feeling extremely prepared compared to moderately prepared (odds ratio 1.93, 95% CI 1.58-2.35). The average midyear faculty-assigned milestone rating for QI among residents who felt "extremely" prepared was 3.28 compared to 3.14 among those who felt "not at all" prepared.

CONCLUSIONS:

Over the past 4 years, family medicine residents' self-assessed preparedness to lead QI projects has barely increased. There was no correlation between self-assessed preparation and faculty-assigned milestone rating. However, we found a small association between self-reported QI leadership and self-assessed QI preparedness.

An Emergency Medicine Milestone-Based Simulation Curriculum: Acute Ischemic Stroke

Turner-Lawrence D, Hang BS, Shah P, Levasseur K. MedEdPORTAL. 2019 Jun 18;15:10829. doi: 10.15766/mep_2374-8265.10829.

INTRODUCTION:

The emergency medicine (EM) resident's ability to make independent decisions in the setting of acute ischemic stroke has been reduced as a result of the involvement of multidisciplinary teams. This simulation was created to give EM residents the opportunity to independently manage the early stages of ischemic stroke and its complications.

METHODS:

A solo learner was presented with a 55-year-old male with complaints consistent with an acute stroke. The resident had to calculate stroke severity; coordinate hospital resources; discuss risks, benefits, and alternatives to thrombolysis; and deal with subsequent complications. The learner had to keep a broad differential for sudden change in mental status and consider alternative interventions. Strategies to decrease intracranial pressure needed to be implemented while obtaining neurosurgical consultation. Debriefing included discussion of expected actions in the context of the Accreditation Council for Graduate Medical Education (ACGME) milestones. Residents' review of their video performance added additional self-reflection.

RESULTS:

A total of 69 PGY 3 EM residents independently participated in this simulation over a 5-year period. Thirty-two completed a postsimulation evaluation. Nearly all learners felt that this case reflected an actual patient encounter and increased their confidence in managing stroke. The milestone-based feedback tool was completed with all learners. Anticipated actions linked to Level 1 and 2 milestones were regularly achieved while acquisition of Level 3 and 4 actions varied.

DISCUSSION:

Case actions were uniquely characterized by the ACGME milestones, which helped to delineate learners' knowledge gaps and provided concrete areas for improvement.

The Implementation of an Introductory Surgical Pathology Didactic Series to Transition First Year Residents and Facilitate Upper Level Resident Teaching

Mehr CR, Montone KT, Schwartz LE. *Adv Anat Pathol*. 2019 May;26(3):210-214. doi: 10.1097/PAP.0000000000000229.

ABSTRACT:

The increasing complexity of the practice of pathology and health care in general requires that pathology residents acquire a vast number of skills during their training. This has been reflected by the broad range of skills addressed in the Accreditation Council for Graduate Medical Education (ACGME) milestones. In order to address some of these milestones, our residency program instituted an introductory didactic series in surgical pathology that focused on 2 objectives. First, the didactics provided basic grossing and histology training to first year residents transitioning from medical school.

Second, the sessions allowed upper level residents to refine their teaching and communication skills at the microscope and therefore served as an important career development tool. Surveys of both first year residents and the upper level residents that led these sessions confirm the utility of these didactics and the use of upper level residents to teach junior trainees. In addition, these sessions led to a dramatic increase in RISE scores among first year trainees. An introductory series with upper level residents leading slide sessions could easily be replicated at other institutions and provide similar benefits.

Resident Case Volume Correlates with Clinical Performance: Finding the Sweet Spot

Agarwal V, Bump GM, Heller MT, Chen LW, Branstetter BF 4th, Amesur NB, Hughes MA. Acad Radiol. 2019 Jan;26(1):136-140. doi: 10.1016/j.acra.2018.06.023. Epub 2018 Aug 4.

RATIONALE AND OBJECTIVES:

To determine whether the total number of studies interpreted during radiology residency correlates with clinical performance as measured by objective criteria.

MATERIALS AND METHODS:

We performed a retrospective cohort study of three graduating classes of radiology residents from a single residency program between the years 2015-2017. The total number of studies interpreted by each resident during residency was tracked. Clinical performance was determined by tracking an individual resident's major discordance rate. A major discordance was recorded when there was a difference between the preliminary resident interpretation and final attending interpretation that could immediately impact patient care. Accreditation council for graduate medical education milestones at the completion of residency, Diagnostic radiology in-training scores in the third year, and score from the American board of radiology core exam were also tabulated. Pearson correlation coefficients and polynomial regression analysis were used to identify correlations between the total number of interpreted films and clinical, test, and milestone performance.

RESULTS:

Thirty-seven residents interpreted a mean of 12,709 studies (range 8898-19,818; standard deviation [SD] 2351.9) in residency with a mean major discordance rate of 1.1% (range 0.34%- 2.54%; stand dev 0.49%). There was a nonlinear correlation between total number of interpreted films and performance. As the number of interpreted films increased to approximately 16,000, clinical performance ($p = 0.004$) and test performance ($p = 0.01$) improved, but volumes over 16,000 correlated with worse performance.

CONCLUSION:

The total number of studies interpreted during radiology training correlates with performance. Residencies should endeavor to find the "sweet spot": the amount of work that maximizes clinical exposure and knowledge without overburdening trainees.

Intern Self-Reported Preparedness for Residency: An ACGME Milestone-Based Study

McLean ME, Huls TA, Park JC, Anana MC, Chen AS, Chien GK, Cygan L, Gupta SJ, Husain A, Mishra DN, Ng KM, Russell JT, Surlles RT, Kulkarni ML. *Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health*. 2019;20:S10.

OBJECTIVE:

Transitioning from medical school to internship is challenging. While several curricula for medical students and interns have been proposed during this transition period, there has not been a large-scale self-assessment of incoming emergency medicine (EM) interns' preparedness for EM milestones. While many medical schools and EM residencies host "boot camps" or other intensive orientation programs for EM-bound students, having knowledge of incoming EM residents' self-perceived strengths and weaknesses will help clerkship directors and EM residency leadership better serve this group of learners. Milestones in EM are used in the United States to measure residents' progress and determine competence at residency completion.⁴ Incoming interns are expected to have achieved level 1 milestones by the time they enter residency, to have achieved level 2 milestones between the first and second year, and to have achieved level 4 milestones before completion of residency. We reached out to 151 newly-matched interns at 11 different sites to ask about their self-perceived "preparedness" for levels 1, 2, and 3 of our eight selected EM milestones (numbers 1, 3, 4, 5, 7, 9, 10, and 12).

DESIGN AND METHOD:

This was a prospective, cross-sectional study of 151 newly-matched pre-interns at 11 EM residency programs. We included all newly-matched interns at each program. Interns were invited via email from their programs to complete a voluntary, anonymous survey prior to the start of residency. The survey used a Likert scale (1 = very unprepared to 5 = extremely prepared) to assess self-reported preparedness to perform levels 1 and 2 of milestones 1, 3, 4, 5, 7, 9, 10, and 12. Milestones were chosen based on ease of teaching in an EM case curriculum that was later implemented.

RESULTS:

A total of 126 pre-interns completed the survey (response rate 83.4%).

CONCLUSION:

Subjects reported highest level of preparedness for emergency stabilization (PC1), and lowest levels of preparedness for airway management (PC10) and pharmacological management (PC5). The data suggest that teachers of fourth-year medical students and new EM interns may want to emphasize milestones 5 and 10 early in internship or late in medical school.

Radiology Resident Assessment and Feedback Dashboard

Durojaiye AB, Snyder E, Cohen M, Nagy P, Hong K, Johnson PT. Radiographics. 2018 Sep- Oct;38 (5):1443-1453. doi: 10.1148/rg.2018170117. Epub 2018 Aug 10.

ABSTRACT:

Assessment of residents is optimally performed through processes and platforms that provide daily feedback, which can be immediately acted on. Given the documentation required by the Accreditation Council for Graduate Medical Education (ACGME), effective data management, integration, and presentation are crucial to ease the burden of manual documentation and increase the timeliness of actionable information. To this end, the authors modeled the learning activities of residents using the Experience Application Programming Interface (xAPI) framework, which is a standard framework for the learning community. On the basis of the xAPI framework and using open-source software to extend their existing infrastructure, the authors developed a Web-based dashboard that provides residents with a more holistic view of their educational experience. The dashboard was designed around the ACGME radiology milestones and provides real-time feedback to residents using various assessment metrics derived from multiple data sources. The purpose of this article is to describe the dashboard's architecture and components, the design and technical considerations, and the lessons learned in implementing the dashboard.

A Cross-Specialty Examination of Resident Error Disclosure and Communication Skills Using Simulation

Gardner AK, Lim G, Minard CG, Guffey D, Pillow MT. J Grad Med Educ. 2018 Aug;10(4):438-441. doi: 10.4300/JGME-D-17-00603.1.

BACKGROUND:

Disclosure of medical errors is important to patients and physicians, but formal disclosure training during the graduate medical education curriculum is limited.

OBJECTIVE:

We examined resident competence related to error disclosure, using standardized patient (SP) ratings of resident communication skills.

METHODS:

All first-year residents from medicine, radiology, emergency medicine, orthopedic surgery, and neurological surgery completed a 20-minute simulated session in which they were provided background information on a medical error they had made and were asked to disclose the error to an SP acting as a family member. Residents were then debriefed and completed a post-scenario questionnaire. The SPs completed an 11-item communication assessment and 3 milestone rating tools on professionalism (PROF-1, PROF-3) and interpersonal and communication skills (ICS-1).

RESULTS:

Ninety-six residents from a single institution participated toward the end of the intern year. Communication assessment scores ranged from 23% to 100% (mean [SD], 80.6 [17.0]). Mean (SD) milestone ratings across specialties were 2.80 ± 0.92 for PROF-1, 2.48 ± 0.92 for PROF-3, and 2.45 ± 0.92 for ICS-1. One-way analysis of variance revealed no significant differences among specialties on milestone or communication ratings. Residents who accepted personal responsibility for the error (84.55 [14.06]) received significantly higher communication ratings from SPs compared with residents who did not (66.67 [19.52], $P < .001$).

CONCLUSIONS:

This SP assessment of error disclosure by first-year residents from multiple specialties was feasible and acceptable. It revealed areas of improvement as well as considerable variation in communication skills and professionalism among residents.

A Survey on Recent Medical School Graduate Comfort With the Level 1 Milestones

Petravick ME, Marsh JL, Karam MD, Dirschl DR. J Surg Educ. 2018 Jul - Aug;75(4):911-917. doi: 10.1016/j.jsurg.2017.10.004. Epub 2017 Nov 7.

OBJECTIVE:

The Next Accreditation System implemented 5 levels of milestones for orthopedic surgery residents in 2013. The Level 1 milestones were noted as those "expected of an incoming resident." While the milestones were intended for assessing resident progression and readiness for independent practice, this designation can also be used to assess how well prepared graduating medical students are for beginning an orthopedic surgery residency. The primary objective of this paper is to measure recent medical school graduate comfort with the Level 1 milestones.

DESIGN, SETTING, AND PARTICIPANTS:

In June 2015, the program directors for the Midwest Orthopaedic Surgical Skills (MOSS) Consortium affiliated residency programs were sent an online survey for distribution to the recent medical school graduates who matched at their respective programs. The survey was about recent graduate comfort with the Level 1 milestone handles associated with 16 orthopedic milestones spanning multiple subspecialties. Responses were grouped based on comfort with individual milestone handles with orthopedic conditions (e.g., carpal tunnel) or with broader categories spanning orthopedic milestones (e.g., imaging).

RESULTS:

In all, 66 of 112 graduates (58.9%) responded. Of 60 milestone handles surveyed, respondents were "Comfortable" with an average of 31.6 ± 14.2 handles with some conditions performing much better than others. The median "Comfortable" response rate was 31 handles. The 8 broader categories had "Comfortable" response rates between 35% and 70%. All 8 orthopedic conditions had significantly higher "Comfortable" response rates for "Evaluation & Knowledge" handles than for "Decision Making & Treatment" handles.

CONCLUSIONS:

Most recent medical student graduates who matched into an orthopedic surgery residencies are only comfortable with about half of the Level 1 milestone handles even though they are expected to meet the Level 1 milestones upon beginning residency. This finding suggests the development of an assessment based on the Level 1 milestones would be appropriate to better inform both graduate and undergraduate medical education in orthopedic surgery.

Resident and Program Director's Perceptions of Milestone-Based Feedback in Obstetrics and Gynecology

Hariton E, Bortoletto P, Barnes KL, Kaimal AJ, Stagg AR. J Med Educ Curric Dev. 2018 May 20;5:2382120518774794. doi: 10.1177/2382120518774794. eCollection 2018 Jan-Dec.

INTRODUCTION:

In July 2014, US residency programs fully implemented the Next Accreditation System including the use of milestone evaluation and reporting. Currently, there has been little investigation into the result of implementation of this new system. Therefore, this study sought to evaluate perceptions of Obstetrics and Gynecology residents and program directors regarding the use of milestone-based feedback and identify areas of deficiency.

METHODS:

A Web-based survey was sent to US-based Obstetrics and Gynecology residents and program directors regarding milestone-based assessment implementation.

RESULTS:

Out of 245 program directors, 84 responded to our survey (34.3% response rate). Of responding program directors, most reported that milestone-based feedback was useful (74.7%), fair (83.0%), and accurate (76.5%); however, they found it administratively burdensome (78.1%). Residents felt that milestone-based feedback was useful (62.7%) and fair (70.0%). About 64.3% of residents and 74.7% of program directors stated that milestone-based feedback is an effective tool to track resident progression; however, a sizable minority of both groups believe that it does not capture surgical aptitude. Qualitative analysis of free response comments was largely negative and highlighted the administrative burden and lack of accuracy of milestone-based feedback.

CONCLUSION:

Overall, both Obstetrics and Gynecology program directors and residents report that milestone-based feedback is useful and fair. Issues of administrative burden, timeliness, evaluation of surgical aptitude, and ability to act on assigned milestone levels were identified. Although this study is limited to one specialty, such issues are likely important to all residents, faculty, and program directors who have implemented the Next Accreditation System requirements.

Shadowing Emergency Medicine Residents by Medical Education Specialists to Provide Feedback on Non-Medical Knowledge-Based ACGME Subcompetencies

Waterbrook AL, Spear Ellinwood KC, Pritchard TG, Bertels K, Johnson AC, Min A, Stoneking LR. *Adv Med Educ Pract.* 2018 May 4;9:307-315. doi: 10.2147/AMEP.S151216. eCollection 2018.

OBJECTIVE:

Non-medical knowledge-based subcompetencies (multitasking, professionalism, accountability, patient-centered communication, and team management) are challenging for a supervising emergency medicine (EM) physician to evaluate in real-time on shift while also managing a busy emergency department (ED). This study examines residents' perceptions of having a medical education specialist shadow and evaluate their nonmedical knowledge skills.

METHODS:

Medical education specialists shadowed postgraduate year 1 and postgraduate year 2 EM residents during an ED shift once per academic year. In an attempt to increase meaningful feedback to the residents, these specialists evaluated resident performance in selected non-medical knowledge-based Accreditation Council of Graduate Medical Education (ACGME) sub-competencies and provided residents with direct, real-time feedback, followed by a written evaluation sent via email. Evaluations provided specific references to examples of behaviors observed during the shift and connected these back to ACGME competencies and milestones.

RESULTS:

Twelve residents participated in this shadow experience (six postgraduate year 1 and six postgraduate year 2). Two residents emailed the medical education specialists ahead of the scheduled shadow shift requesting specific feedback. When queried, five residents voluntarily requested their feedback to be included in their formal biannual review. Residents received milestone scores and narrative feedback on the non-medical knowledge-based ACGME subcompetencies and indicated the shadow experience and subsequent feedback were valuable.

CONCLUSION:

Medical education specialists who observe residents over the course of an entire shift and evaluate non-medical knowledge-based skills are perceived by EM residents to provide meaningful feedback and add valuable information for the biannual review process.

Interim Analysis of a Prospective Multi-Institutional Study of Surgery Resident Experience with Flexibility in Surgical Training

Cullinan DR, Wise PE, Delman KA, Potts JR, Awad MM, Eberlein TJ, Klingensmith ME. *J Am Coll Surg*. 2018 Apr;226(4):425-431. doi: 10.1016/j.jamcollsurg.2017.12.024. Epub 2018 Jan 5.

BACKGROUND:

The Flexibility in Surgical Training (FIST) consortium project was designed to evaluate the feasibility and resident outcomes of optional subspecialty-focused training within general surgery residency training.

STUDY DESIGN:

After approval by the American Board of Surgery, R4 and R5 residents were permitted to customize up to 12 of the final 24 months of residency for early tracking into 1 of 9 subspecialty tracks. A prospective IRB-approved study was designed across 7 institutions to evaluate the impact of this option on operative experience, in-service exam (American Board of Surgery In-Training Examination [ABSITE]) and ACGME milestone performance, and resident and program director (PD) perceptions. The FIST residents were compared with chief residents before FIST initiation (controls) as well as residents during the study period who did not participate in FIST (no specialization track, NonS).

RESULTS:

From 2013 to 2017, 122 of 214 chief residents (57%) completed a FIST subspecialty track. There were no differences in median ABSITE scores between FIST, NonS residents, and controls. The ACGME milestones at the end of the R5 year favored the FIST residents in 13 of 16 milestones compared with NonS. Case logs demonstrated an increase in track-specific cases compared with NonS residents. Resident and PD surveys reported a generally favorable experience with FIST.

CONCLUSIONS:

In this prospective study, FIST is a feasible option in participating institutions. All FIST residents, regardless of track, met requirements for ABS Board eligibility, despite modifications to rotations and case experience. Future studies will assess the impact of FIST on ABS exam results and fellowship success.

The Effect and Use of Milestones in the Assessment of Neurological Surgery Residents and Residency Programs

Conforti LN, Yaghmour NA, Hamstra SJ, Holmboe ES, Kennedy B, Liu JJ, Waldo H, Selden NR. *J Surg Educ*. 2018 Jan - Feb;75(1):147-155. doi: 10.1016/j.jsurg.2017.06.001. Epub 2017 Jun 22.

OBJECTIVES:

The purpose of this study was to determine the effect of the Accreditation Council for Graduate Medical Education Milestones on the assessment of neurological surgery residents. The authors sought to determine the feasibility, acceptability, and utility of this new framework in making judgments of progressive competence, its implementation within programs, and the influence on curricula. Residents were also surveyed to elicit the effect of Milestones on their educational experience and professional development.

DESIGN, SETTING, AND PARTICIPANTS:

In 2015, program leadership and residents from 21 neurological surgery residency programs participated in an online survey and telephone interview in which they reflected on their experiences with the Milestones. Survey data were analyzed using descriptive statistics. Interview transcripts were analyzed using grounded theory.

RESULTS:

Response themes were categorized into 2 groups: outcomes of the Milestones implementation process, and facilitators and barriers. Because of Milestones implementation, participants reported changes to the quality of the assessment process, including the ability to identify struggling residents earlier and design individualized improvement plans. Some programs revised their curricula based on training gaps identified using the Milestones. Barriers to implementation included limitations to the adoption of a developmental progression model in the context of rotation block schedules and misalignment between progression targets and clinical experience. The shift from time-based to competency-based evaluation presented an ongoing adjustment for many programs. Organized preparation before clinical competency committee meetings and diverse clinical competency committee composition led to more productive meetings and perceived improvement in promotion decisions.

CONCLUSIONS:

The results of this study can be used by program leadership to help guide further implementation of the Milestones and program improvement. These results also help to guide the evolution of Milestones language and their implementation across specialties.

How Effective are New Milestones Assessments at Demonstrating Resident Growth? 1 Year of Data

Goldman RH, Tuomala RE, Bengtson JM, Stagg AR. J Surg Educ. 2017 Jan-Feb;74(1):68-73. doi: 10.1016/j.jsurg.2016.06.009. Epub 2016 Jul 6.

OBJECTIVE:

Assessment tools that accrue data for the Accreditation Council for Graduate Medical Education Milestones must evaluate residents across multiple dimensions, including medical knowledge, procedural skills, teaching, and professionalism. Our objectives were to: (1) develop an assessment tool to evaluate resident performance in accordance with the Milestones and (2) review trends in resident achievements during the inaugural year of Milestone implementation.

DESIGN:

A novel venue and postgraduate year (PGY) specific assessment tool was built, tested, and implemented for both operating room and labor and delivery "venues." Resident development of competence and independence was captured over time. To account for variable rotation schedules, the year was divided into thirds and compared using two-tailed Fisher's exact test.

SETTING:

Brigham and Women's and Massachusetts General Hospitals, Boston MA.

PARTICIPANTS:

Faculty evaluators and obstetrics and gynecology residents.

RESULTS:

A total of 822 assessments of 44 residents were completed between 9/2014 and 6/2015. The percentage of labor and delivery tasks completed "independently" increased monotonically across the start of all years: 8.4% for PGY-1, 60.3% for PGY-2, 73.7% for PGY-3, and 87.5% for PGY-4. Assessments of PGY-1 residents demonstrated a significant shift toward "with minimal supervision" and "independent" for the management of normal labor ($p = 0.03$). PGY-3 residents demonstrated an increase in "able to be primary surgeon" in the operating room, from 36% of the time in the first 2/3 of the year, to 62.3% in the last 1/3 ($p < 0.01$).

CONCLUSION:

Assessment tools developed to assist with Milestone assignments capture the growth of residents over time and demonstrate quantifiable differences in achievements between PGY classes. These tools will allow for targeted teaching opportunities for both individual residents and residency programs.

Using the ACGME Milestones for Resident Self-Evaluation and Faculty Engagement

Meier AH, Gruessner A, Cooney RN. J Surg Educ. 2016 Nov-Dec;73(6):e150-e157. doi: 10.1016/j.jsurg.2016.09.001.

BACKGROUND:

Since July 2014 General Surgery residency programs have been required to use the Accreditation Council for Graduate Medical Education milestones twice annually to assess the progress of their trainees. We felt this change was a great opportunity to use this new evaluation tool for resident self- assessment and to furthermore engage the faculty in the educational efforts of the program.

METHODS:

We piloted the milestones with postgraduate year (PGY) II and IV residents during the 2013/2014 academic year to get faculty and residents acquainted with the instrument. In July 2014, we implemented the same protocol for all residents. Residents meet with their advisers quarterly. Two of these meetings are used for milestones assessment. The residents perform an independent self- evaluation and the adviser grades them independently. They discuss the evaluations focusing mainly on areas of greatest disagreement. The faculty member then presents the resident to the clinical competency committee (CCC) and the committee decides on the final scores and submits them to the Accreditation Council for Graduate Medical Education website. We stored all records anonymously in a MySQL database. We used Anova with Tukey post hoc analysis to evaluate differences between groups. We used intraclass correlation coefficients and Krippendorff's α to assess interrater reliability.

RESULTS:

We analyzed evaluations for 44 residents. We created scale scores across all Likert items for each evaluation. We compared score differences by PGY level and raters (self, adviser, and CCC). We found highly significant increases of scores between most PGY levels ($p < 0.05$). There were no significant score differences per PGY level between the raters. The interrater reliability for the total score and 6 competency domains was very high (ICC: 0.87-0.98 and α : 0.84-0.97). Even though this milestone evaluation process added additional work for residents and faculty we had very good participation (93.9% by residents and 92.9% by faculty) and feedback was generally positive.

CONCLUSION:

Even though implementation of the milestones has added additional work for general surgery residency programs, it has also opened opportunities to furthermore engage the residents in reflection and self- evaluation and to create additional venues for faculty to get involved with the educational process within the residency program. Using the adviser as the initial rater seems to correlate closely with the final CCC assessment. Self-evaluation by the resident is a requirement by the RRC and the milestones seem to be a good instrument to use for this purpose. Our early assessment suggests the milestones provide a useful instrument to track trainee progression through their residency.

The Internal Medicine Reporting Milestones: Cross-Sectional Description of Initial Implementation in US Residency Programs

Hauer KE, Clauser J, Lipner RS, Holmboe ES, Caverzagie K, Hamstra SJ, Hood S, Iobst W, Warm E, McDonald FS. *Ann Intern Med.* 2016 Sep 6;165(5):356-62. doi: 10.7326/M15-2411. Epub 2016 May 10.

BACKGROUND:

High-quality assessment of resident performance is needed to guide individual residents' development and ensure their preparedness to provide patient care. To facilitate this aim, reporting milestones are now required across all internal medicine (IM) residency programs.

OBJECTIVE:

To describe initial milestone ratings for the population of IM residents by IM residency programs.

DESIGN:

Cross-sectional study.

SETTING:

IM residency programs.

PARTICIPANTS:

All IM residents whose residency program directors submitted milestone data at the end of the 2013-2014 academic year.

MEASUREMENTS:

Ratings addressed 6 competencies and 22 subcompetencies. A rating of "not assessable" indicated insufficient information to evaluate the given subcompetency. Descriptive statistics were calculated to describe ratings across competencies and training years.

RESULTS:

Data were available for all 21 774 U.S. IM residents from all 383 programs. Overall, 2889 residents (1621 in postgraduate year 1 [PGY-1], 902 in PGY-2, and 366 in PGY-3) had at least 1 subcompetency rated as not assessable. Summaries of average ratings by competency and training year showed higher ratings for PGY-3 residents in all competencies. Overall ratings for each of the 6 individual competencies showed that fewer than 1% of third-year residents were rated as "unsatisfactory" or "conditional on improvement." However, when subcompetency milestone ratings were used, 861 residents (12.8%) who successfully completed training had at least 1 competency with all corresponding subcompetencies graded below the threshold of "readiness for unsupervised practice."

LIMITATION:

Data were derived from a point in time in the first reporting period in which milestones were used.

CONCLUSION:

The initial milestone-based evaluations of IM residents nationally suggest that documenting developmental progression of competency is possible over training years. Subcompetencies may identify areas in which residents might benefit from additional feedback and experience. Future work is needed to explore how milestones are used to support residents' development and enhance residency curricula.

A Pilot Study of Orthopaedic Resident Self-Assessment Using a Milestones Survey just Prior to Milestones Implementation

Bradley KE, Andolsek KM. Int J Med Educ. 2016 Jan 11;7:11-8. doi: 10.5116/ijme.5682.6dfd.

OBJECTIVE:

To pilot test if Orthopaedic Surgery residents could self-assess their performance using newly created milestones, as defined by the Accreditation Council on Graduate Medical Education.

METHODS:

In June 2012, an email was sent to Program Directors and administrative coordinators of the 154 accredited Orthopaedic Surgery Programs, asking them to send their residents a link to an online survey. The survey was adapted from the Orthopaedic Surgery Milestone Project. Completed surveys were aggregated in an anonymous, confidential database. SAS 9.3 was used to perform the analyses.

RESULTS:

Responses from 71 residents were analyzed. First and second year residents indicated through self-assessment that they had substantially achieved Level 1 and Level 2 milestones. Third year residents reported they had substantially achieved 30/41, and fourth year residents, all Level 3 milestones. Fifth year, graduating residents, reported they had substantially achieved 17 Level 4 milestones, and were extremely close on another 15. No milestone was rated at Level 5, the maximum possible. Earlier in training, Patient Care and Medical Knowledge milestones were rated lower than the milestones reflecting the other four competencies of Practice Based Learning and Improvement, Systems Based Practice, Professionalism, and Interpersonal Communication. The gap was closed by the fourth year.

CONCLUSIONS:

Residents were able to successfully self-assess using the 41 Orthopaedic Surgery milestones. Respondents' rate improved proficiency over time. Graduating residents report they have substantially, or close to substantially, achieved all Level 4 milestones. Milestone self-assessment may be a useful tool as one component of a program's overall performance assessment strategy.

Emergency Medicine Residents Consistently Rate Themselves Higher than Attending Assessments on ACGME Milestones

Goldflam K, Bod J, Della-Giustina D, Tsyrluk A. West J Emerg Med. 2015 Nov;16(6):931-5. doi: 10.5811/westjem.2015.8.27247. Epub 2015 Nov 12.

INTRODUCTION:

In 2012, the Accreditation Council for Graduate Medical Education (ACGME) introduced the Next Accreditation System (NAS), which implemented milestones to assess the competency of residents and fellows. While attending evaluation and feedback is crucial for resident development, perhaps equally important is a resident's self-assessment. If a resident does not accurately self-assess, clinical and professional progress may be compromised. The objective of our study was to compare emergency medicine (EM) resident milestone evaluation by EM faculty with the same resident's self-assessment.

METHODS:

This is an observational, cross-sectional study that was performed at an academic, four-year EM residency program. Twenty-five randomly chosen residents completed milestone self-assessment using eight ACGME sub-competencies deemed by residency leadership as representative of core EM principles. These residents were also evaluated by 20 faculty members. The milestone levels were evaluated on a nine-point scale. We calculated the average difference between resident self-ratings and faculty ratings, and used sample t-tests to determine statistical significance of the difference in scores.

RESULTS:

Eighteen residents evaluated themselves. Each resident was assessed by an average of 16 attendings (min=10, max=20). Residents gave themselves statistically significant higher milestone ratings than attendings did for each sub-competency examined ($p < 0.0001$).

CONCLUSION:

Residents over-estimated their abilities in every sub-competency assessed. This underscores the importance of feedback and assessment transparency. More attention needs to be paid to methods by which residency leadership can make residents' self-perception of their clinical ability more congruent with that of their teachers and evaluators. The major limitation of our study is small sample size of both residents and attendings.