Milestones

Bibliography

January 2019

Sonia Sangha, MPH
Stanley J. Hamstra, PhD
Table of Contents

Introduction ........................................................................................................................................ 3

Use of Bibliography .......................................................................................................................... 4

Audience – Program Directors

Assessment Tools and Processes for Collecting Milestones Data ........................................... 5
CCCs – Structure and Function ....................................................................................................... 94
How to use Milestones Data to Improve your Program ............................................................... 109
Using Milestones to Guide Curriculum ......................................................................................... 148
Rationale for Milestones ............................................................................................................. 173
Trends in Milestones Data ........................................................................................................... 225
Content – Do the Milestones Represent my Specialty/Program? .............................................. 247
How Residents Can Use Milestones Data .................................................................................. 295

Appendix

Papers Describing the Initial Development of Milestones Language by Specialty
(JGME suppl. issues March 2013 and March 2014) ................................................................. 308
Introduction

This bibliography provides a list of peer-reviewed published articles, as well as commentaries, editorials, and opinion pieces that specifically address the topic of the Milestones.

This list is designed to be useful in tracking validity evidence for the Milestones, as well as in accessing links to educational theory and strategies for implementation of this framework in individual residency and fellowship programs.

Additionally, the list was envisioned to serve as a springboard for further research into the validity and use of the Milestones to help develop the best evidence for decisions around resident and fellows’ progression within accredited programs and beyond.

This list will be updated approximately every six months.

Stanley J. Hamstra, PhD
Vice President, Milestone Research and Evaluation
shamstra@acgme.org

Sonia Sangha, MPH
Research Associate for Milestones Evaluation
ssangha@acgme.org
Use of Bibliography

The bibliography is organized according to the categories below, and is presented by year of publication (in descending order) within each category. Note: some articles are coded in more than one category.

<table>
<thead>
<tr>
<th>Audience</th>
<th>Category</th>
<th>Number of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program Directors</td>
<td>Assessment Tools/Processes for Collecting Milestones Data</td>
<td>88</td>
</tr>
<tr>
<td>Program Directors</td>
<td>CCCs – Structure and Function</td>
<td>14</td>
</tr>
<tr>
<td>Program Directors</td>
<td>How to Use Milestones Data to Improve your Program</td>
<td>38</td>
</tr>
<tr>
<td>Program Directors</td>
<td>Using Milestones to Guide Curriculum</td>
<td>24</td>
</tr>
<tr>
<td>Program Directors</td>
<td>Rationale for Milestones</td>
<td>50</td>
</tr>
<tr>
<td>Program Directors</td>
<td>Trends in Milestones Data</td>
<td>21</td>
</tr>
<tr>
<td>Program Directors</td>
<td>Content - Do the Milestones Represent my Specialty/Program?</td>
<td>47</td>
</tr>
<tr>
<td>Program Directors</td>
<td>How Residents Can Use Milestones Data</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Total # of Articles</td>
<td>210</td>
</tr>
</tbody>
</table>

The breakdown by Year of Publication is as follows:

- 2018: 60
- 2017: 50
- 2016: 33
- 2015: 31
- 2014: 24
- 2013: 10
- 2012: 2

Keywords Used and Search Criteria:
- Date Range: January 1, 2012-December 31, 2018
- Source Type: Academic Journals
- Database(s) used: EBSCO, PubMed
- Keywords:
  - EBSCO: (TI milestone* OR AB milestone*) AND (AB (resident* OR fellow*) AND SU (graduate AND medical AND education))
  - PubMed: (milestone*[Title/Abstract]) AND (resident* OR fellow* OR (graduate AND medical AND education))

Note: since the content that follows comes from a variety of sources, there may be significant variations in style and spelling with respect to terminology. For official ACGME terminology and terminology usage, refer to the ACGME Glossary of Terms.
Assessment Tools and Processes for Collecting Milestones Data
Examining the Effects of Narrative Commentary on Evaluators’ Summative Assessments of Resident Performance


ABSTRACT

Anchor-based, end-of-shift ratings are commonly used to conduct performance assessments of resident physicians. These performance evaluations often include narrative assessments, such as solicited or "free-text" commentary. Although narrative commentary can help to create a more detailed and specific assessment of performance, there are limited data describing the effects of narrative commentary on the global assessment process. This single-group, observational study examined the effect of narrative comments on global performance assessments. A subgroup of the clinical competency committee, blinded to resident identity, assigned a single, consensus-based performance score (1-6) to each resident based solely on end-of-shift milestone scores. De-identified narrative comments from end-of-shift evaluations were then included and the process was repeated. We compared milestone-only scores to milestone plus narrative commentary scores using a nonparametric sign test. During the study period, 953 end-of-shift evaluations were submitted on 41 residents. Of these, 535 evaluations included free-text narrative comments. In 17 of the 41 observations, performance scores changed after the addition of narrative comments. In two cases, scores decreased with the addition of free-text commentary. In 15 cases, scores increased. The frequency of net positive change was significant (p = .0023). The addition of narrative commentary to anchor-based ratings significantly influenced the global performance assessment of Emergency Medicine residents by a committee of educators. Descriptive commentary collected at the end of shift may inform more meaningful appraisal of a resident's progress in a milestone-based paradigm. The authors recommend clinical training programs collect unstructured narrative impressions of residents' performance from supervising faculty.
Assessment Methods and Resource Requirements for Milestone Reporting by an Emergency Medicine Clinical Competency Committee


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) introduced milestones for Emergency Medicine (EM) in 2012. Clinical Competency Committees (CCC) are tasked with assessing residents on milestones and reporting them to the ACGME. Appropriate workflows for CCCs are not well defined.

OBJECTIVE:
Our objective was to compare different approaches to milestone assessment by a CCC, quantify resource requirements for each and to identify the most efficient workflow.

DESIGN:
Three distinct processes for rendering milestone assessments were compared: Full milestone assessments (FMA) utilizing all available resident assessment data, Ad-hoc milestone assessments (AMA) created by multiple expert educators using their personal assessment of resident performance, Self-assessments (SMA) completed by residents. FMA were selected as the theoretical gold standard. Intraclass correlation coefficients were used to analyze for agreement between different assessment methods. Kendall's coefficient was used to assess the inter-rater agreement for the AMA.

RESULTS:
All 13 second-year residents and 7 educational faculty of an urban EM Residency Program participated in the study in 2013. Substantial or better agreement between FMA and AMA was seen for 8 of the 23 total subcompetencies (PC4, PC8, PC9, PC11, MK, PROF2, ICS2, SBP2), and for 1 subcompetency (SBP1) between FMA and SMA. Multiple AMA for individual residents demonstrated substantial or better interobserver agreement in 3 subcompetencies (PC1, PC2, and PROF2). FMA took longer to complete compared to AMA (80.9 vs. 5.3 min, p < 0.001).

CONCLUSIONS:
Using AMA to evaluate residents on the milestones takes significantly less time than FMA. However, AMA and SMA agree with FMA on only 8 and 1 subcompetencies, respectively. An estimated 23.5 h of faculty time are required each month to fulfill the requirement for semiannual reporting for a residency with 42 trainees.
Assessment of Emergency Medicine Resident Performance in an Adult Simulation Using a Multisource Feedback Approach

Jong, Michael; Elliott, Nicole; Nguyen, Michael; Goyke, Terrence; Johnson, Steven; Cook, Matthew; Lindauer, Lisa; Best, Katie; Gernerd, Douglas; Morolla, Louis; Matuzsan, Zachary; Kane, Bryan. Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health. 2018 Dec 17. Vol. 20 Issue 1. DOI: 10.5811/westjem.2018.12.39844

INTRODUCTION:
The Accreditation Council for Graduate Medical Education (ACGME) specifically notes multisource feedback (MSF) as a recommended means of resident assessment in the emergency medicine (EM) Milestones. High-fidelity simulation is an environment wherein residents can receive MSF from various types of healthcare professionals. Previously, the Queen’s Simulation Assessment Tool (QSAT) has been validated for faculty to assess residents in five categories: assessment; diagnostic actions; therapeutic actions; interpersonal communication, and overall assessment. We sought to determine whether the QSAT could be used to provide MSF using a standardized simulation case.

METHODS:
Prospectively after institutional review board approval, residents from a dual ACGME/osteopathic-approved postgraduate years (PGY) 1-4 EM residency were consented for participation. We developed a standardized resuscitation after overdose case with specific 1-5 Likert anchors used by the QSAT. A PGY 2-4 resident participated in the role of team leader, who completed a QSAT as self-assessment. The team consisted of a PGY-1 peer, an emergency medical services (EMS) provider, and a nurse. Two core faculty were present to administer the simulation case and assess. Demographics were gathered from all participants completing QSATs. We analyzed QSATs by each category and on cumulative score. Hypothesis testing was performed using intra-class correlation coefficients (ICC), with 95% confidence intervals. Interpretation of ICC results was based on previously published definitions.

RESULTS:
We enrolled 34 team leader residents along with 34 nurses. A single PGY-1, a single EMS provider and two faculty were also enrolled. Faculty provided higher cumulative QSAT scores than the other sources of MSF. QSAT scores did not increase with team leader PGY level. ICC for inter-rater reliability for all sources of MSF was 0.754 (0.572-0.867). Removing the self-evaluation scores increased inter-rater reliability to 0.838 (0.733-0.910). There was lesser agreement between faculty and nurse evaluations than from the EMS or peer evaluation. Conclusion: In this single-site cohort using an internally developed simulation case, the QSAT provided MSF with excellent reliability. Self-assessment decreases the reliability of the MSF, and our data suggest self-assessment should not be a component of MSF. Use of the QSAT for MSF may be considered as a source of data for clinical competency committees.
Using Operating Room Turnover Time by Anesthesia Trainee Level to Assess Improving Systems-Based Practice Milestones


BACKGROUND:
Operating room (OR) metrics are frequently cited when optimizing cost efficacy and quality of care (Weiss et al, Characteristics of operating room procedures in U.S. hospitals, 2011: Statistical brief #170, 2013; Macario A, Anesthesiology 105:237-240, 2006; Childers et al, JAMA Surg 153:e176233, 2018). Little has been reported to evaluate how anesthesia trainees change anesthesia-related efficiencies in the OR. Statistical correlation may demonstrate awareness and implementation of efficient systems-based practice.

METHODS:
Utilizing computerized OR information systems, specific data regarding anesthesia controlled turnover times were collected (546 data points) over the course of 4 months. The type of surgery performed, patient's American Society of Anesthesiologists (ASA) physical status and OR turnover times were compared for clinical anesthesia (CA) trainee levels CA1, CA2, CA3 and CRNAs. Standard descriptive statistics were computed. Analysis of variance (ANOVA) was performed to compare the average turnover time.

RESULTS:
Average OR turnover time was 31 min ranging from 8 to 60 min. There was a significant difference between the OR turnover time of CA-1 (32 min) compared to CA-3 (29 min) (p = 0.017) and CA-1 compared to CRNA (30 min) (p = 0.016). OR turnover time was significantly shorter in CA-3 and CRNA. The analysis showed no differences between OR turnover time of ASA categories.

CONCLUSIONS:
These findings posit that trainees improve efficiency over time, but that education may for a time come at the expense of productivity. This trend may demonstrate a more profound understanding and mastery of a learner progressing in the graduate medical education system. This interplay plays a key role in clinical and academic shared success.
Milestones in Plastic Surgery: Attending Assessment versus Resident Assessment


BACKGROUND:
The Plastic Surgery Milestones Project was jointly conceived by the ACGME and American Board of Plastic Surgery as a tool to improve granularity in resident feedback. Resident self-evaluations were compared to attending clinical competency committee evaluations to gauge resident self-perceptions and understanding of the milestones framework.

METHODS:
Semi-annual evaluations from June 2014-2017 were analyzed and compared with corresponding resident self-evaluations from the 2015-2017 academic year at the University of Pittsburgh Medical Center. Evaluations were analyzed for overall trends in performance. The Wilcoxon Rank-Sign test was used to identify any systematic differences between evaluations. Subgroup analysis using the chi-square test was performed to determine factors that may contribute to major assessment disparity (≥1).

RESULTS:
6,207 milestones across 187 faculty evaluations and 3,139 milestones across 106 resident self-evaluations were available for review. With the exception of PGY-2 residents, residents rated themselves at a significantly lower level in the competencies of medical education and patient care. Post-graduate year, academic year timing, and ACGME competency were associated with major assessment discrepancies.

CONCLUSIONS:
Overall, resident and faculty evaluations at our program were concordant which demonstrates that residents are capable of accurately assessing their own abilities and understanding the milestones framework. Areas of discordance between resident and faculty evaluations fostered discussion between residents and faculty and have led to multiple changes in our program. The introduction of self-evaluation tools at other programs may provide them with similar benefits.
Standardized Patients to Assess Resident Interpersonal Communication Skills and Professional Values Milestones


ABSTRACT
It has been a challenge to assess communication and professional values Milestones in emergency medicine (EM) residents using standardized methods, as mandated by the Accreditation Council for Graduate Medical Education (ACGME). This paper outlines an innovative method of assessing these Milestones using an established instructional method. EM faculty mapped the communication and professional values Milestones to an existing communication and interpersonal skills scale. We identified six communication-focused scenarios: death notification; informed consent; medical non-compliance; medical error; treatment refusal; and advanced directives. In a pilot, 18 EM residents completed these six standardized patient (SP) encounters. Our experience suggests SP encounters can support standardized direct observation of residents’ achievement of ACGME Milestones. Further effort can be made to create a tailored, behaviorally-anchored tool that uses the Milestones as the conceptual framework.
BACKGROUND: Recently the American Board of Orthopaedic Surgery and the Accreditation Council of Graduate Medical Education have identified the treatment of septic arthritis of the hip in children as a milestone skill for all US orthopaedic residents. The purpose of this study was to test correlation between clinical knowledge and examination score on a mobile app-based training module for the treatment of pediatric septic hip arthritis.

METHODS: A 4-part simulation model on surgical decision-making associated with the treatment of pediatric septic arthritis was developed through expert consensus. Orthopaedic trainees participating in the "Top Gun" program of the 2015 and 2016 International Pediatric Orthopaedic Symposia were recruited to participate in this pilot study. Trainees completed a presimulation quiz on their knowledge of diagnosis, arthrocentesis, and surgical irrigation and debridement on a pediatric patient presenting with septic arthritis of the hip. Trainees then completed the 4-part simulation on the mobile app. Pearson correlation analysis was used to assess the relationship between the quiz and the simulation.

RESULTS: A total of 53 orthopaedic residents and fellows participated in the simulation. Median quiz score was 87 points [interquartile range (IQR), 81 to 94] before the intervention and 100 points (IQR, 94 to 100) postintervention. The median simulation test score was 89 (IQR, 81 to 92) which demonstrated a positive correlation with the postintervention quiz (r=0.44, P<0.001). The preintervention metrics demonstrated a positive correlation with postintervention metrics (r=0.53, P<0.001).

CONCLUSIONS: This study revealed a statistically significant positive correlation between the mobile app simulation and the clinical knowledge of the participants, as well as the ability to improve knowledge about a procedure during the testing period. These findings support the ability for the mobile app to test clinical knowledge. In the current environment of decreased work hours and patient exposure for orthopaedic trainees, mobile app-based simulation has the potential to safely aid in assessment of orthopaedic residents and fellows.
Four-Year Analysis of a Novel Milestone-Based Assessment of Faculty by General Surgical Residents


OBJECTIVE:
In response to our faculty's concerns about the quality and reliability of feedback from general surgery residents, we developed a novel faculty assessment tool. This study was designed as an interim analysis of the tool's effectiveness and discriminatory ability.

METHODS:
Our department's educational leadership developed milestones in 7 domains that were scored from 1 to 4, with each level representing an educational approach that ranged from ineffective (1) to ideal (4). Each postgraduate year (PGY) class meets annually to develop a consensus regarding each faculty member's effectiveness in each of the 7 domains: (1) operative supervision, (2) operative teaching, (3) clinic and/or hospital supervision, (4) clinic and/or hospital teaching, (5) conference participation, (6) availability, and (7) overall contribution to the training program. We reviewed the results from the initial 4 years of this project. We also analyzed the annual national faculty survey administered by the Accreditation Council for Graduate Medical Education (ACGME) to evaluate faculty satisfaction regarding feedback during the same study period. Data were assessed using the Levene test for homogeneity, analysis of variance, and Wilcoxon-Mann-Whitney tests.

RESULTS:
Forty-two faculty members were annually evaluated by 29 to 32 residents. Each resident PGY class assigned faculty milestone scores that varied across the 7 domains, demonstrating that faculty scores reflected variable opinions about each specific domain, while avoiding labeling an effective faculty member with all high scores and a less effective member with all poor scores. (p<0.0001). Milestone scores for a given faculty member differed across PGY classes, indicating that junior residents might evaluate a specific faculty member differently than senior residents (p<0.0001). Eleven faculty members received low scores of 1 or 2 on the overall contribution to training domain and 8/11 (73%) improved to 3 or 4, the following year. Twenty core faculty members were included on the annual ACGME survey. The results from the study period on the ACGME anonymous faculty survey reflected enhanced satisfaction with resident feedback during the study period, improving from 68% to 88% compliance with ACGME standards and our mean program score improved from 4.1 to 4.4 compared to the national mean of 4.3 (p = 0.02).

CONCLUSIONS:
This milestone-based faculty assessment tool improves the quality of the feedback from surgical residents when evaluating faculty. When residents assign a negative statement to describe faculty educational effectiveness in a specific domain, performance improves. A milestone-based faculty assessment strategy should be explored on a national level.
Predicting Performance of First-Year Residents: Correlations between Structured Interview, Licensure Exam, and Competency Scores in a Multi-Institutional Study


**PURPOSE:**
To determine whether scores on structured interview (SI) questions designed to measure noncognitive competencies in physicians (1) predict subsequent first-year resident performance on Accreditation Council for Graduate Medical Education (ACGME) milestones and (2) add incremental validity over United States Medical Licensing Examination (USMLE) Step 1 and Step 2 Clinical Knowledge (CK) scores in predicting performance.

**METHOD:**
The authors developed 18 behavioral description questions to measure key noncognitive competencies (e.g., teamwork). In 2013-2015, 14 programs (13 residency, 1 fellowship) from 6 institutions used subsets of these questions in their selection processes. The authors conducted analyses to determine the validity of SIs and USMLE scores in predicting first-year resident milestone performance in the ACGME’s core competency domains and overall.

**RESULTS:**
SI scores predicted mid-year and year-end overall performance ($r = .18$ and $.19$, respectively, $P < .05$) and year-end performance on patient care, interpersonal and communication skills, and professionalism competencies ($r = .23$, $r = .22$, and $r = .20$, respectively, $p < .05$). SI scores contributed incremental validity over USMLE scores in predicting year-end performance on patient care ($\Delta R = .05$), interpersonal and communication skills ($\Delta R = .09$), and professionalism ($\Delta R = .09$; all $P < .05$). USMLE scores contributed incremental validity over SI scores in predicting year-end performance overall and on patient care and medical knowledge.

**CONCLUSIONS:**
SI scores predict first-year resident year-end performance in the interpersonal and communication skills, patient care, and professionalism competency domains. Future research should investigate whether SIs predict a range of clinically relevant outcomes.
Can Deficiencies in Performance Be Identified Earlier in Surgical Residency? An Initial Report of a Surgical Trainee Assessment of Readiness Exam


OBJECTIVE:
Identifying gaps in medical knowledge, patient management, and procedural competence is difficult early in surgical residency. We designed and implemented an end-of-year examination for our postgraduate year 1 residents, entitled Surgical Trainee Assessment of Readiness (STAR). Our objective in this study was to determine whether STAR scores correlated with other available indicators of resident performance, such as the American Board of Surgery in-training exam (ABSITE) and Milestone scores, and if they provided evidence of additional discriminatory value.

STUDY DESIGN:
Overall and component scores of the STAR exam were compared to the ABSITE and Milestone assessment scores for the 17 categorical residents that took the exam in 2016 and 2017.

SETTING:
Harbor-UCLA Medical Center, a university-affiliated academic medical center.

PARTICIPANTS:
Seventeen categorical general surgery residents.

RESULTS:
The STAR Total Test Score ($\beta = 2.77, p = 0.006$) was an independent predictor of the ABSITE taken the same year, and components of the STAR were independent predictors of ABSITE taken the following year. The STAR Total Test Score was lowest in the 3 residents who had at least 1 low Milestone score assessed in the same year; and 2 of these 3 residents had at least 1 low Milestone score assigned the next year after STAR. Lastly, the Patient Care 1 and 2 Milestones assessed in the same year as STAR were uniformly scored as appropriate for level of training, yet the corresponding STAR component for those milestones demonstrated 3 residents as having deficiencies.

CONCLUSIONS:
We have created a multifaceted standardized STAR exam, which correlates with performance on the ABSITE and early milestone scores. It also appears to discriminate resident performance where milestone assessments do not. Further evaluation of the STAR exam with longer term follow-up is needed to confirm these initial findings.
Resident Case Volume Correlates with Clinical Performance: Finding the Sweet Spot


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.
Individualized Learning Plan (ILP) Is an Effective Tool in Assessing Achievement of Otology-related Subcompetency Milestones


OBJECTIVE:
To investigate the individualized learning plan (ILP) as a tool in assessment of residents’ milestone achievements as they pertain to Otology subcompetencies: Chronic Ear Disease, Pediatric Otitis Media, and Hearing Loss.

STUDY DESIGN:
Prospective study.

METHODS:
Twenty otolaryngology residents were instructed to use an ILP and identify six milestones from three otology-related subcompetencies to focus on during the course of a 3-month rotation. They were also asked to plan out specific activities which would help them achieve these milestones, to specify whether or not they successfully achieved them, by what instructional or learning methods and to identify any barriers. The completed ILPs were reviewed by a faculty member.

MAIN OUTCOME MEASURES:
The effectiveness of the ILP was assessed by response compliance rate, corroboration of self-reported milestone achievement with faculty evaluations and the ability to set attainable milestones.

RESULTS:
There was 95% compliance in using an ILP to achieve milestones. Self-reported milestone scores corresponded to the faculty evaluations in a large majority (89.6%) of patients, and tended to be underestimated by the residents. Out of 114 total milestones identified, 44 (38.6%) were not achieved, with particular overestimation in the use of independent study as a learning method.

CONCLUSION:
The ILP is an effective tool in measuring residents' achievement of otology-related milestones, and could possibly be used to supplement or replace faculty assessment. The ILP provides valuable information on barriers to achieving milestones and informs trainees on how to set attainable goals as they pertain to patient care and medical knowledge in otology.
Validity of Simulation-Based Assessment for Accreditation Council for Graduate Medical Education Milestone Achievement


INTRODUCTION:
The Accreditation Council for Graduate Medical Education requires biannual evaluation of anesthesiology residents on 25 subcompetency milestones. Some milestone domains are particularly challenging to repeatedly and reliably observe during clinical care. Simulation-Based Milestones Assessment (SBMA) may help overcome these challenges. However, few studies have examined the external validation of simulation assessment scores (ie, the relationships between simulation-based assessment scores and other standard measures of ability) for milestones. This study analyzed whether SBMA scores (1) discriminate by postgraduate year, (2) improve over time, and (3) correlate with traditional measures of performance.

METHODS:
This is a retrospective analysis of 55 residents' SBMA data from 30 scenarios for two academic years. Each scenario was evaluated for time-in-training discrimination. Scenarios were then analyzed for SBMA scoring trends over time, and SBMA scores were compared with residents' clinical evaluations.

RESULTS:
Twenty-four SBMA scenarios discriminated by postgraduate year. Repeated measure analysis of variance showed statistically significant between-session score improvements (F (3, 54) = 17.79, P < 0.001). Pearson correlation coefficients demonstrated moderate to strong correlation between SBMA and clinical evaluations: January 2015 r = 0.67, P < 0.01 (n = 27); May 2015 r = 0.43, P = 0.09 (n = 17); November 2015 r = 0.70, P < 0.01 (n = 24); and April 2016 r = 0.70, P <0.01 (n = 27).

CONCLUSIONS:
The associations between SBMA scores and experience level, time-in-training, and clinical performance evaluations provide evidence that SBMA may be used as metrics of residents' Accreditation Council for Graduate Medical Education milestone competencies.
Preliminary Validity Evidence for a Milestones-Based Rating Scale for Chart-Stimulated Recall


BACKGROUND:
Minimally anchored Standard Rating Scales (SRSs), which are widely used in medical education, are hampered by suboptimal interrater reliability. Expert-derived frameworks, such as the Accreditation Council for Graduate Medical Education (ACGME) Milestones, may be helpful in defining level-specific anchors to use on rating scales.

OBJECTIVE:
We examined validity evidence for a Milestones-Based Rating Scale (MBRS) for scoring chart-stimulated recall (CSR).

METHODS:
Two 11-item scoring forms with either an MBRS or SRS were developed. Items and anchors for the MBRS were adapted from the ACGME Internal Medicine Milestones. Six CSR standardized videos were developed. Clinical faculty scored videos using either the MBRS or SRS and following a randomized crossover design. Reliability of the MBRS versus the SRS was compared using intraclass correlation.

RESULTS:
Twenty-two faculty were recruited for instrument testing. Some participants did not complete scoring, leaving a response rate of 15 faculty (7 in the MBRS group and 8 in the SRS group). A total of 529 ratings (number of items × number of scores) using SRSs and 540 using MBRSs were available. Percent agreement was higher for MBRSs for only 2 of 11 items—use of consultants (92 versus 75, \( P = .019 \)) and unique characteristics of patients (96 versus 79, \( P = .011 \))—and the overall score (89 versus 82, \( P < .001 \)). Interrater agreement was 0.61 for MBRSs and 0.51 for SRSs.

CONCLUSIONS:
Adding milestones to our rating form resulted in significant, but not substantial, improvement in intraclass correlation coefficient. Improvement was inconsistent across items.
Psychotherapy Competency Milestones: an Exploratory Pilot of CBT and Psychodynamic Psychotherapy Skills Acquisition in Junior Psychiatry Residents


OBJECTIVE:
Psychiatry residents train in Psychodynamic Psychotherapy and Cognitive Behavioral Therapy (CBT), evidence-supported treatments used in mental health care that can facilitate clinical reasoning, foster therapeutic alliances, and improve clinical outcomes. However, empirically derived milestones are needed to evaluate competency. This exploratory pilot examined changes over 1 year of training in junior psychiatry residents' competency milestone elements in Psychodynamic Psychotherapy and CBT.

METHODS:
Seventy-nine randomly selected audio-recorded sessions from differing phases of Psychodynamic Psychotherapy and CBT with five junior residents and ten patients were rated using the Psychotherapy Process Q-sort (PQS).

RESULTS:
In both treatments, patient engagement with attention to in-session emotions improved. In CBT, residents were directive, supported patients' self-efficacy, emphasized patients' accepting responsibility for their problems, discussed homework such as thought records, and focused on termination in the concluding sessions. In Psychodynamic Psychotherapy, residents attended to emotional arousal and linked patients' feelings or perceptions to past situations or behavior. Growth and hierarchical linear modeling differentiated these treatments, with CBT v. Psychodynamic adherence to PQS modality-specific ideal elements being 52% v. 19%.

CONCLUSION:
Teaching and observation using empirically derived observable psychotherapy practice behaviors is feasible and can be used to assess milestone elements for competency-based education of psychiatry trainees.
Use of a Clinical Pathologic Conference to Demonstrate Residents’ ACGME Emergency Medicine Milestones, Aid in Faculty Development, and Increase Academic Output


ABSTRACT
The Emergency Medicine Milestones Project, developed by the Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Emergency Medicine, includes competence targets for residents to attain and, ultimately, to exceed American Osteopathic Association and ACGME expectations for residents. The authors sought to use the clinical pathologic conference (CPC) format in their institutions’ Emergency Medicine Milestones Project to provide measurable residency academic and faculty development outcomes. The CPC is an event in which a resident presents an unknown case to a discussant in advance of a didactic session to demonstrate an organized approach and decision-making rationale to a differential diagnosis. Feedback forms included the assessment of resident discussants from the perspective of level-5 Milestone achievements in particular. Developing an internal CPC competition with a dedicated core faculty coordinator who provides skill development for both resident and faculty presentation has proven successful. Such a competition can document the level-5 achievements for senior residents, be a source of faculty development, and increase peer-reviewed academic output.
Evaluating Surgical Residents’ Patient-Centered Communication Skills: Practical Alternatives to the "Apprenticeship Model"


OBJECTIVES:
The Accreditation Council for Graduate Medical Education (ACGME) requires residency programs to assess communication skills and provide feedback to residents. We aimed to develop a feasible data collection process that generates objective clinical performance information to guide training activities, inform ACGME milestone evaluations, and validate assessment instruments.

DESIGN:
Residents care for patients in the surgical clinic and in the hospital, and participate in a communication curriculum providing practice with standardized patients (SPs). We measured perception of resident communication using the 14-item Communication Assessment Tool (CAT), collecting data from patients at the surgery clinic and surgical wards in the hospital, and from SP encounters during simulated training scenarios. We developed a handout of CAT example behaviors to guide patients completing the communication assessment.

SETTING: Independent academic medical center.

PARTICIPANTS: General surgery residents.

RESULTS:
The primary outcome is the percentage of total items patients rated "excellent;" we collected data on 24 of 25 residents. Outpatient evaluations resulted in significantly higher scores (mean 84.5% vs. 68.6%, p < 0.001), and female patients provided nearly statistically significantly higher ratings (mean 85.2% vs. 76.7%, p = 0.084). In multivariate analysis, after controlling for patient gender, visit reason, and race, (1) residents' CAT scores from SPs in simulation were independently associated with communication assessments in their concurrent patient population (p = 0.017), and (2) receiving CAT example instructions was associated with a lower percentage of excellent ratings by 9.3% (p = 0.047).

CONCLUSIONS:
Our data collection process provides a model for obtaining meaningful information about resident communication proficiency. CAT evaluations of surgical residents by the inpatient population had not previously been described in the literature; our results provide important insight into relationships between the evaluations provided by inpatients, clinic patients, and SPs in simulation. Our example behaviors guide shows promise for addressing a common concern, minimizing ceiling effects when measuring physician-patient communication.
Teaching and Assessing Professionalism in Radiology: Resources and Scholarly Opportunities to Contribute to Required Expectations


ABSTRACT
Teaching and assessing trainees' professionalism now represents an explicit expectation for Accreditation Council Graduate Medical Education-accredited radiology programs. Challenges to meeting this expectation include variability in defining the construct of professionalism; limits of traditional teaching and assessment methods, used for competencies historically more prominent in medical education, for professionalism; and emerging expectations for credible and feasible professionalism teaching and assessment practices in the current context of health care training and practice. This article identifies promising teaching resources and methods that can be used strategically to augment traditional teaching of the cognitive basis for professionalism, including role modeling, case-based scenarios, debriefing, simulations, narrative medicine (storytelling), guided discussions, peer-assisted learning, and reflective practice. This article also summarizes assessment practices intended to promote learning, as well as to inform how and when to assess trainees as their professional identities develop over time, settings, and autonomous practice, particularly in terms of measurable behaviors. This includes assessment tools (including mini observations, critical incident reports, and appreciative inquiry) for authentic assessment in the workplace; engaging multiple sources (self-peer, other health professionals, and patients) in assessment; and intentional practices for trainees to take responsibility for seeking our actionable feedback and reflection. This article examines the emerging evidence of the feasibility and value added of assessment of medical competency milestones, including professionalism, coordinated by the Accreditation Council Graduate Medical Education in radiology and other medical specialties. Radiology has a strategic opportunity to contribute to scholarship and inform policies in professionalism teaching and assessment practices.
Resident and Program Director’s Perceptions of Milestone-Based Feedback in Obstetrics and Gynecology


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 Sub-Competencies


OBJECTIVE:
Non-medical knowledge-based sub-competencies (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 post graduate 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 sub-competencies 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.
Patient-Centered Medical Home Status and Preparedness to Assess Resident Milestones: A CERA Study


PURPOSE:
The patient-centered medical home (PCMH) model has been proposed as the ideal model for delivering primary care and is focused on improving patient safety and quality, reducing costs, and enhancing patient satisfaction. The mandated Accreditation Council for Graduate Medical Education educational milestones for evaluation of resident competency represent the skills graduates will utilize after graduation. Many of these skills are reflected in the PCMH model. We sought to determine if residency programs whose main family medicine (FM) practice sites have achieved PCMH recognition are therefore more prepared to evaluate milestones.

METHOD:
A national Council of Academic Family Medicine Educational Research Alliance (CERA) survey of family medicine program directors (PDs) was conducted during June and July 2015 to determine if PCMH recognition influences PDs' ability to evaluate training methods and their level of preparedness to evaluate milestones.

RESULTS:
The response rate for the survey was 53.3% (252/473). Nearly two-thirds of the PDs (62.7%) reported that their main FM practice site had earned PCMH recognition. There was no statistical difference between non-PCMH-recognized vs PCMH-recognized programs in how PDs perceived that their program was prepared to assess residents' milestone levels overall (P=0.414). Residents of PCMH-recognized programs were more likely to receive training for team-based care (P=0.009), system improvement plans (P<0.001), root-cause analysis (P=0.002), and health behavior change (P=0.003).

CONCLUSIONS:
PCMH recognition itself did not improve preparedness of FM residency programs to assess milestones. Residents from programs whose main FM practice site is PCMH-recognized are more likely to be trained in the key concepts and tasks associated with the PCMH model, tools that they are expected to utilize extensively after graduation.
Resident-Specific Morbidity Reduced Following ACS NSQIP Data-Driven Quality Program


BACKGROUND:
The Accreditation Council for Graduate Medical Education Milestone Project for general surgery provided a more robust method for developing and tracking residents' competence. This framework enhanced systematic and progressive development of residents' competencies in surgical quality improvement.

STUDY DESIGN:
A 22-month interactive, educational program based on resident-specific surgical outcomes data culminated in a quality improvement project for postgraduate year 4 surgery residents. Self-assessment, quality knowledge test, and resident-specific American College of Surgeons National Surgical Quality Improvement Program Quality In-Training Initiative morbidity were compared before and after the intervention.

RESULTS:
Quality in-training initiative morbidity decreased from 25% (82/325) to 18% (93/517), p = 0.015 despite residents performing more complex cases. All participants achieved level 4 competency (4/4) within the general surgery milestones improvement of care, practice-based learning and improvement competency. Institutional American College of Surgeons National Surgical Quality Improvement Program general surgery morbidity improved from the ninth to the sixth decile. Quality assessment and improvement self-assessment postintervention scores (M = 23.80, SD = 4.97) were not significantly higher than preintervention scores (M = 19.20, SD = 5.26), p = 0.061. Quality Improvement Knowledge Application Tool postintervention test scores (M = 17.4, SD = 4.88), were not significantly higher than pretest scores (M = 13.2, SD = 1.92), p = 0.12.

CONCLUSION:
Sharing validated resident-specific clinical data with participants was associated with improved surgical outcomes. Participating fourth year surgical residents achieved the highest score, a level 4, in the practice based learning and improvement competency of the improvement of care practice domain and observed significantly reduced surgical morbidity for cases in which they participated.
A Multicenter Collaboration for Simulation-Based Assessment of ACGME Milestones in Emergency Medicine


ABSTRACT
In 2014, the six allopathic emergency medicine (EM) residency programs in Chicago established an annual, citywide, simulation-based assessment of all postgraduate year 2 EM residents. The cases and corresponding assessment tools were designed by the simulation directors from each of the participating sites. All assessment tools include critical actions that map directly to numerous EM milestones in 11 different subcompetencies. The 2-hour assessments provide opportunities for residents to lead resuscitations of critically ill patients and demonstrate procedural skills, using mannequins and task trainers respectively. More than 80 residents participate annually and their assessment experiences are essentially identical across testing sites. The assessments are completed electronically and comparative performance data are immediately available to program directors.
Comparing Real-Time versus Delayed Video Assessments for Evaluating ACGME Sub-Competency Milestones in Simulated Patient Care Environments


BACKGROUND:
Background Simulation is an effective method for creating objective summative assessments of resident trainees. Real-time assessment (RTA) in simulated patient care environments is logistically challenging, especially when evaluating a large group of residents in multiple simulation scenarios. To date, there is very little data comparing RTA with delayed (hours, days, or weeks later) video-based assessment (DA) for simulation-based assessments of Accreditation Council for Graduate Medical Education (ACGME) sub-competency milestones. We hypothesized that sub-competency milestone evaluation scores obtained from DA, via audio-video recordings, are equivalent to the scores obtained from RTA.

METHODS:
Forty-one anesthesiology residents were evaluated in three separate simulated scenarios, representing different ACGME sub-competency milestones. All scenarios had one faculty member perform RTA and two additional faculty members perform DA. Subsequently, the scores generated by RTA were compared with the average scores generated by DA. Variance component analysis was conducted to assess the amount of variation in scores attributable to residents and raters.

RESULTS:
Paired t-tests showed no significant difference in scores between RTA and averaged DA for all cases. Cases 1, 2, and 3 showed an intraclass correlation coefficient (ICC) of 0.67, 0.85, and 0.50 for agreement between RTA scores and averaged DA scores, respectively. Analysis of variance of the scores assigned by the three raters showed a small proportion of variance attributable to raters (4% to 15%).

CONCLUSIONS:
The results demonstrate that video-based delayed assessment is as reliable as real-time assessment, as both assessment methods yielded comparable scores. Based on a department's needs or logistical constraints, our findings support the use of either real-time or delayed video evaluation for assessing milestones in a simulated patient care environment.
Simulation for Assessment of Milestones in Emergency Medicine Residents


OBJECTIVES:
All residency programs in the United States are required to report their residents' progress on the milestones to the Accreditation Council for Graduate Medical Education (ACGME) biannually. Since the development and institution of this competency-based assessment framework, residency programs have been attempting to ascertain the best ways to assess resident performance on these metrics. Simulation was recommended by the ACGME as one method of assessment for many of the milestone subcompetencies. We developed three simulation scenarios with scenario-specific milestone-based assessment tools. We aimed to gather validity evidence for this tool.

METHODS:
We conducted a prospective observational study to investigate the validity evidence for three mannequin-based simulation scenarios for assessing individual residents on emergency medicine (EM) milestones. The subcompetencies (i.e., patient care [PC]1, PC2, PC3) included were identified via a modified Delphi technique using a group of experienced EM simulationists. The scenario-specific checklist (CL) items were designed based on the individual milestone items within each EM subcompetency chosen for assessment and reviewed by experienced EM simulationists. Two independent live raters who were EM faculty at the respective study sites scored each scenario following brief rater training. The inter-rater reliability (IRR) of the assessment tool was determined by measuring intraclass correlation coefficient (ICC) for the sum of the CL items as well as the global rating scales (GRSs) for each scenario. Comparing GRS and CL scores between various postgraduate year (PGY) levels was performed with analysis of variance.

RESULTS:
Eight subcompetencies were chosen to assess with three simulation cases, using 118 subjects. Evidence of test content, internal structure, response process, and relations with other variables were found. The ICCs for the sum of the CL items and the GRSs were >0.8 for all cases, with one exception (clinical management GRS = 0.74 in sepsis case). The sum of CL items and GRSs (p < 0.05) discriminated between PGY levels on all cases. However, when the specific CL items were mapped back to milestones in various proficiency levels, the milestones in the higher proficiency levels (level 3 [L3] and 4 [L4]) did not often discriminate between various PGY levels. L3 milestone items discriminated between PGY levels on five of 12 occasions they were assessed, and L4 items discriminated only two of 12 times they were assessed.

CONCLUSION:
Three simulation cases with scenario-specific assessment tools allowed evaluation of EM residents on proficiency L1 to L4 within eight of the EM milestone subcompetencies. Evidence of test content, internal structure, response process, and relations with other variables were found. Good to excellent IRR and the ability to discriminate between various PGY levels was found for both the sum of CL items and the GRSs. However, there was a lack of a positive relationship between advancing PGY level and the completion of higher-level milestone items (L3 and L4).
myTIPreport and Training for Independent Practice: A Tool for Real-Time Workplace Feedback for Milestones and Procedural Skills


BACKGROUND:
Few tools currently exist for effective, accessible delivery of real-time, workplace feedback in the clinical setting.

OBJECTIVE:
We developed and implemented a real-time, web-based tool for performance-based feedback in the clinical environment.

METHODS:
The tool (myTIPreport) was designed for performance-based feedback to learners on the Accreditation Council for Graduate Medical Education (ACGME) Milestones and procedural skills. "TIP" stands for "Training for Independent Practice." We implemented myTIPreport in obstetrics and gynecology (Ob-Gyn) and female pelvic medicine and reconstructive surgery (FPMRS) programs between November 2014 and May 2015. Residents, fellows, teachers, and program directors completed preimplementation and postimplementation surveys on their perceptions of feedback.

RESULTS:
Preimplementation surveys were completed by 656 participants of a total of 980 learners and teachers in 19 programs (12 Ob-Gyn and 7 FPMRS). This represented 72% (273 of 378) of learners and 64% (383 of 602) of teachers. Seventy percent of participants (381 of 546) reported having their own individual processes for real-time feedback; the majority (79%, 340 of 430) described these processes as informal discussions. Over 6 months, one-third of teachers and two-thirds of learners used the myTIPreport tool a total of 4311 times. Milestone feedback was recorded 944 times, and procedural feedback was recorded 3367 times. Feedback addressed all ACGME Milestones and procedures programmed into myTIPreport. Most program directors reported that tool implementation was successful.

CONCLUSIONS:
The majority of learners successfully received workplace feedback using myTIPreport. This web-based tool, incorporating procedures and ACGME Milestones, may be an important transition from other feedback formats.
Programmatic Assessment in Emergency Medicine: Implementation of Best Practices


BACKGROUND:
Programmatic assessment is the intentional collection of key data from multiple sources for both assessment of learning and assessment for learning.

OBJECTIVE:
We developed a system of programmatic assessment (PA) to identify competency progression (summative) and assessment for learning to assist residents in their formative development.

METHODS:
The programmatic assessment was designed iteratively from 2014 through 2016. All assessments were first categorized by competency domain and source of assessment. The number of assessment modalities for each competency domain was collected. These multisource assessments were then mapped by program leadership to the milestones to develop a master PA blueprint. A resident learning management system provided the platform for aggregating formative and summative data, allowing residents and faculty ongoing access to guide learning and assessment. A key component of programmatic assessment was to support resident integration of assessment information through feedback by faculty after shifts and during monthly formal assessments, semiannual resident reviews, and summative judgments by the Clinical Competency Committee.

RESULTS:
Through the PA, the 6 competency domains are assessed through multiple modalities: patient care (22 different assessments), professionalism (18), systems-based practice (17), interprofessional and communication skills (16), medical knowledge (11), and practice-based learning and improvement (6). Each assessment provides feedback to the resident in various formats. Our programmatic assessment has been utilized for more than 2 years with iterative improvements.

CONCLUSIONS:
The implementation of programmatic assessment allowed our program to organize diverse, multisourced feedback to drive both formative and summative assessments.
Do End-of-Rotation and End-of-Shift Assessments Inform Clinical Competency Committees’ (CCC) Decisions?


INTRODUCTION:
Clinical Competency Committees (CCC) require reliable, objective data to inform decisions regarding assignment of milestone proficiency levels, which must be reported to the Accreditation Council for Graduate Medical Education. After the development of two new assessment methods, the end-of-shift (EOS) assessment and the end-of-rotation (EOR) assessment, we sought to evaluate their performance. We report data on the concordance between these assessments, as well as how each informs the final proficiency level determined in biannual CCC meetings. We hypothesized that there would be a high concordance level between the two assessment methods, including concordance of both the EOS and EOR with the final proficiency level designation by the CCC.

METHODS:
The residency program is an urban academic four-year emergency medicine residency with 48 residents. After their shifts in the emergency department (ED), residents handed out EOS assessment forms asking about individual milestones from 15 subcompetencies to supervising physicians, as well as triggered electronic EOR-doctor (EORd) assessments to supervising doctors and EOR-nurse (EORn) to nurses they had worked with after each two-week ED block. EORd assessments contained the full proficiency level scale from 16 subcompetencies, while EORn assessments contained four subcompetencies. Data reports were generated after each six-month assessment period and data was aggregated. We calculated Spearman's rank order correlations for correlations between assessment types and between assessments and final CCC proficiency levels.

RESULTS:
Over 24 months, 5,234 assessments were completed. The strongest correlations with CCC proficiency levels were the EORd for the immediate six-month assessment period prior ($r_s$ 0.71- 0.84), and the CCC proficiency levels from the previous six-months ($r_s$ 0.83-0.92). EOS assessments had weaker correlations ($r_s$ 0.49 to 0.62), as did EORn ($r_s$ 0.4 to 0.73).

CONCLUSION:
End-of-rotation assessments completed by supervising doctors are most highly correlated with final CCC proficiency level designations, while end-of-shift assessments and end-of-rotation assessments by nurses did not correlate strongly with final CCC proficiency levels, both with overestimation of levels noted. Every level of proficiency the CCC assigned appears to be highly correlated with the designated level in the immediate six-month period, perhaps implying CCC members are biased by previous level assignments.
**Straight Line Scoring by Clinical Competency Committees Using Emergency Medicine Milestones**


**BACKGROUND:**
In 2013, milestone ratings became a reporting requirement for emergency medicine (EM) residency programs. Programs rate each resident in the fall and spring on 23 milestone subcompetencies.

**OBJECTIVE:**
This study examined the incidence of straight line scoring (SLS) for EM Milestone ratings, defined as a resident being assessed the same score across the milestone subcompetencies.

**METHODS:**
This descriptive analysis measured the frequencies of SLS for all Accreditation Council for Graduate Medical Education (ACGME)-accredited EM programs during the 2015-2016 academic year. Outcomes were the frequency of SLS in the fall and spring milestone assessments, changes in the number of SLS reports, and reporting trends. Chi-square analysis compared nominal variables.

**RESULTS:**
There were 6257 residents in the fall and 6588 in the spring. Milestone scores were reported for 6173 EM residents in the fall (99% of 6257) and spring (94% of 6588). In the fall, 93% (5753 residents) did not receive SLS ratings and 420 (7%) did, with no significant difference compared with the spring (5776 [94%] versus 397 [6%]). Subgroup analysis showed higher SLS results for residents' first ratings (183 of 2136 versus 237 of 4220, $P < .0001$) and for their final ratings (200 of 2019 versus 197 of 4354, $P < .0001$). Twenty percent of programs submitted 10% or more SLS ratings, and a small percentage submitted more than 50% of ratings as SLS.

**CONCLUSIONS:**
Most programs did not submit SLS ratings. Because of the statistical improbability of SLS, any SLS ratings reduce the validity assertions of the milestone assessments.
Comprehensive Assessment of Struggling Learners Referred to a Graduate Medical Education Remediation Program


BACKGROUND:
Implementation of the Next Accreditation System has provided a standardized framework for identifying learners not meeting milestones, but there is as yet no corresponding framework for remediation.

OBJECTIVE:
We developed a comprehensive assessment process that allows correct diagnosis of a struggling learner's deficit(s) to promote successful remediation.

METHODS:
At the University of Pennsylvania, resident learners within the Department of Medicine who are not meeting milestones are referred to the Early Intervention Remediation Committee (EIRC). The EIRC, composed of 14 faculty members with expertise in remediation, uses a standardized process to assess learners' deficits. These faculty members categorize primary deficits as follows: medical knowledge, clinical reasoning, organization and efficiency, professionalism, and communication skills. The standardized process of assessment includes an analysis of the learner's file, direct communication with evaluators, an interview focused on learner perception of the problem, screening for underlying medical or psychosocial issues, and a review of systems for deficits in the 6 core competencies. Participants were surveyed after participating in this process.

RESULTS:
Over a 2-year period, the EIRC assessed and developed remediation plans for 4% of learners (14 of a total 342). Following remediation and reassessment, the identified problems were satisfactorily resolved in all cases with no disciplinary action. While the process was time intensive, an average of 45 hours per learner, the majority of faculty and residents rated it as positive and beneficial.

CONCLUSIONS:
This structured assessment process identifies targeted areas for remediation and adds to the tools available to Clinical Competency Committees.
BACKGROUND:
The Family Medicine (FM) Milestones are competency-based assessments of residents in key dimensions relevant to practice in the specialty. Residency programs use the milestones in semiannual reviews of resident performance from the time of entry into the program to graduation.

OBJECTIVE:
Using a national sample, we investigated the relationship of FM competency-based assessments to resident progress and the complementarity of milestones with knowledge-based assessments in FM residencies.

METHODS:
We used midyear and end-of-year milestone ratings for all FM residents in Accreditation Council for Graduate Medical Education-accredited programs during academic years 2014-2015 and 2015-2016. The milestones contain 22 items across 6 competencies. We created a summative index across the milestones. The American Board of Family Medicine database provided resident demographics and in-training examination (ITE) scores. We linked information to the milestone data.

RESULTS:
The sample encompassed 6630 FM residents. The summative milestone index increased, on average, for each cohort (postgraduate year 1 [PGY-1] to PGY-2 and PGY-2 to PGY-3) at each assessment. The correlation between the milestone index that excluded the medical knowledge milestone and ITE scores was $r = .195 \ (P < .001)$ for PGY-1 to PGY-2 cohort and $r = .254 \ (P < .001)$ for PGY-2 to PGY-3 cohort. For both cohorts, ITE scores and composite milestone assessments were higher for residents who advanced than for those who did not.

CONCLUSIONS:
Competency-based assessment using the milestones for FM residents seems to be a viable multidimensional tool to assess the successful progression of residents.
Two-Year Experience Implementing a Curriculum to Improve Residents' Patient-Centered Communication Skills


OBJECTIVES:
Surgery milestones from The Accreditation Council for Graduate Medical Education have encouraged a focus on training and assessment of residents' nontechnical skills, including communication. We describe our 2-year experience implementing a simulation-based curriculum, results of annual communication performance assessments, and resident evaluations.

DESIGN:
Eight quarterly modules were conducted on various communication topics. Former patient volunteers served as simulation participants (SP) who completed annual assessments using the Communication Assessment Tool (CAT). During these 2 modules, communication skills were assessed in the following standardized scenarios: (1) delivering bad news to a caregiver of a patient with postoperative intracerebral hemorrhage and (2) primary care gallstone referral with contraindications for cholecystectomy. SP-CAT ratings were evaluated for correlations by individual and associations with trainee and SP characteristics. Surgical patient experience surveys are evaluated during the curriculum.

SETTING:
Independent academic medical center surgical simulation center.

PARTICIPANTS:
Twenty-five surgery residents per year in 2015 to 2017.

RESULTS:
Residents have practiced skills in a variety of scenarios including bad news delivery, medical error disclosure, empathic communication, and end-of-life conversations. Residents report positive learning experiences from the curriculum (90% graded all modules A/A+). Confidence ratings rose following each module (p < 0.001) and in the second year (p < 0.001). Annual assessments yielded insights into skills level, and relationships to resident confidence levels and traits. Communication scores were not associated with resident gender or postgraduate year. Over the course of the curriculum implementation, surgical patients have reported that doctors provided explanations with improved clarity (p = 0.042).

CONCLUSIONS:
The simulation-based SP-CAT has shown initial evidence of usability, content validity, relationships to observed communication behaviors and residents' skills confidence. Evaluations of different scenarios may not be correlated for individuals over time. The communication curriculum paralleled improvements in patient experience concerning surgeons' clear explanations. An ongoing surgery resident communication curriculum has numerous educational, assessment, and institutional benefits.
Impact of an Event Reporting System on Resident Complication Reporting in Plastic Surgery Training: Addressing an ACGME and Plastic Surgery Milestone Project Core Competency


BACKGROUND:
The Accreditation Council for Graduate Medical Education and Plastic Surgery Milestone Project has identified practice-based learning and improvement, which involves systematically analyzing current practices and implementing changes, as a core competency in residency education. In surgical care, complication reporting is an essential component of practice-based learning and improvement as complications are analyzed in morbidity and mortality conference for quality improvement. Unfortunately, current methods for capturing a comprehensive profile of complications may significantly underestimate the true occurrence of complications. Therefore, the objectives of this study are to evaluate an intervention for complication reporting and compare this to current practice, in a plastic surgery training program.

METHODS:
This is a preintervention and postintervention study evaluating resident reporting of complications on a plastic surgery service. The intervention was an online event reporting system developed by department leadership and patient safety experts. The cohorts consisted of all patients undergoing surgery during two separate 3-month blocks bridged by an implementation period. A trained reviewer recorded complications, and this served as the reference standard. Fisher's exact test was used for binary comparisons.

RESULTS:
There were 32 complications detected in 219 patients from June to August of 2015 and 35 complications in 202 patients from October to December of 2015. The proportion of complications reported in the preintervention group was nine of 32 (28.1 percent). After the intervention, this significantly increased to 32 of 35 (91.4 percent) (p < 0.001).

CONCLUSION:
An intervention using an event reporting system, supported by departmental leadership, led to significant improvements in complication reporting by plastic surgery residents.
Implementation of a Service-Specific Template Integrating Objective Structured Clinical Examination and Accreditation Council for Graduate Medical Education Milestones: One Institution's Experience


ABSTRACT
The Accreditation Council for Graduate Medical Education (ACGME) orthopaedic milestones require detailed, frequent resident evaluations. This institution desired a cost-effective objective structured clinical examination (OSCE) to facilitate these evaluations. Data were collected as a prospective, uncontrolled observational study. The OSCE was completed by residents entering and exiting the foot and ankle rotation during postgraduate years 2 and 4. Physician assistants functioned as standardized patients. Statistical analyses were performed using paired and independent t tests. The OSCE was implemented using reliable, low-cost modalities and has facilitated milestones evaluations. Preliminary data show 4th-year residents performed higher in prerotation global assessment with a standardized patient and written exam (p < .03). Second-year residents showed improvement in the written exam on rotation completion (p = .03). Using this methodology, institutions may establish similar cost-effective OSCEs as feasible evaluative solutions to satisfy milestone requirements. The authors believe this tool may be modified for any specialty. (Journal of Surgical Orthopaedic Advances 26(4):257-261, 2017).
A Checklist to Help Faculty Assess ACGME Milestones in a Video-Recorded OSCE


BACKGROUND:
Faculty members need to assess resident performance using the Accreditation Council for Graduate Medical Education Milestones.

OBJECTIVE:
In this randomized study we used an objective structured clinical examination (OSCE) around the disclosure of an adverse event to determine whether use of a checklist improved the quality of milestone assessments by faculty.

METHODS:
In 2013, a total of 20 anesthesiology faculty members from 3 institutions were randomized to 2 groups to assess 5 videos of trainees demonstrating advancing levels of competency on the OSCE. One group used milestones alone, and the other used milestones plus a 13-item checklist with behavioral anchors based on ideal performance. We classified faculty ratings as either correct or incorrect with regard to the competency level demonstrated in each video, and then used logistic regression analysis to assess the effect of checklist use on the odds of correct classification.

RESULTS:
Thirteen of 20 faculty members rated assessing performance using milestones alone as difficult or very difficult. Checklist use was associated with significantly greater odds of correct classification at entry level (odds ratio [OR] = 9.2, 95% confidence interval [CI] 4.0-21.2) and at junior level (OR = 2.7, 95% CI 1.3-5.7) performance. For performance at other competency levels checklist use did not affect the odds of correct classification.

CONCLUSIONS:
A majority of anesthesiology faculty members reported difficulty with assessing a videotaped OSCE of error disclosure using milestones as primary assessment tools. Use of the checklist assisted in correct assessments at the entry and junior levels.
Development of a Global Health Milestones Tool for Learners in Emergency Medicine: A Pilot Project


OBJECTIVES:
In medical education and training, increasing numbers of institutions and learners are participating in global health experiences. Within the context of competency-based education and assessment methodologies, a standardized assessment tool may prove valuable to all of the aforementioned stakeholders. Milestones are now used as the standard for trainee assessment in graduate medical education. Thus, the development of a similar, milestone-based tool was undertaken, with learners in emergency medicine (EM) and global health in mind.

METHODS:
The Global Emergency Medicine Think Tank Education Working Group convened at the 2016 Society for Academic Medicine Annual Meeting in New Orleans, Louisiana. Using the Interprofessional Global Health Competencies published by the Consortium of Universities for Global Health's Education Committee as a foundation, the working group developed individual milestones based on the 11 stated domains. An iterative review process was implemented by teams focused on each domain to develop a final product.

RESULTS:
Milestones were developed in each of the 11 domains, with five competency levels for each domain. Specific learning resources were identified for each competency level and assessment methodologies were aligned with the milestones framework. The Global Health Milestones Tool for learners in EM is designed for continuous usage by learners and mentors across a career.

CONCLUSIONS:
This Global Health Milestones Tool for learners in EM may prove valuable to numerous stakeholders. The next steps include a formalized pilot program for testing the tool's validity and usability across training programs, as well as an assessment of perceived utility and applicability by collaborating colleagues working in training sites abroad.
Mapping Residency Global Health Experiences to the ACGME Family Medicine Milestones


BACKGROUND AND OBJECTIVES:
Global health (GH) experiences are a unique part of family medicine (FM) training that offer an opportunity for residents to demonstrate development across a multitude of the milestones recently implemented by the Accreditation Council for Graduate Medical Education (ACGME). The GH experience presents an opportunity for resident development, and including a component of written reflection can provide tangible evidence of development in areas that can be difficult to assess.

METHODS:
A mixed methods approach was used to integrate quantitative (frequency) data with qualitative content from the written reflections of 12 of our FM residents who participated in GH experiences.

RESULTS:
Written reflections touched on each of the 22 milestones, although some milestones were noted more frequently than others. The most commonly identified milestones fell within the competency areas of systems-based practice, professionalism, and practice-based learning and improvement. Our qualitative approach allowed us to gain an appreciation of the unique experiences that demonstrated growth across the various milestones.

CONCLUSIONS:
We conclude that any program that offers GH experiences should incorporate some form of written reflection to maximize resident growth and offer evaluative faculty a window into that development.
The Impact of Using Mean versus Mode When Assessing Resident Competency


BACKGROUND:
The Accreditation Council for Graduate Medical Education Milestone Project was implemented in 2014 to standardize assessments and progression of residents. While it is recommended that milestones not be used as tools for direct assessments of resident competency, many programs have used or adapted milestone tools for this purpose.

OBJECTIVE:
We sought to explore use of the most frequent milestone level at which a resident was evaluated (ie, the mode), and compared this to the standard practice of using the arithmetic mean for summarizing performance.

METHODS:
We reviewed all Family Medicine Milestone evaluations from 1 program for the first 2 academic years of milestone implementation. Mean and mode scores were calculated across 24 unique residents, 841 evaluation forms, and 5897 measurements. The proportion of overestimation errors (where the mean is at least 0.5 larger than the mode) and underestimation errors (where the mean is at least 0.5 less than the mode) were then compared across resident training year and subcompetency.

RESULTS:
For the 24 residents, an estimation error occurred in 175 of 792 of the comparisons (22%). Of these errors, 118 (67%) were overestimation errors. First-year residents accounted for 55% (96 of 175) of all estimation errors. All subcompetencies had some estimation errors, with 6 having greater than 5%.

CONCLUSIONS:
If the trend for using the milestones as stand-alone assessment tools is to continue, aggregating data by using frequency distributions and mode would be a more stable and appropriate approach given their nominal or, at best, ordinal nature.
Rapid Web-Based Platform for Assessment of Orthopedic Surgery Patient Care Milestones: A 2-Year Validation


OBJECTIVE:
To determine the validity, feasibility, and responsiveness of a new web-based platform for rapid milestone-based evaluations of orthopedic surgery residents.

SETTING:
Single academic medical center, including a trauma center and pediatrics tertiary hospital.

PARTICIPANTS:
Forty residents (PG1-5) in an orthopedic residency program and their faculty evaluators.

METHODS:
Residents and faculty were trained and supported in the use of a novel trainee-initiated web-based evaluation system. Residents were encouraged to use the system to track progress on patient care subcompetencies. Two years of prospectively collected data were reviewed from residents at an academic program. The primary outcome was Spearman's rank correlation between postgraduate year (PGY) and competency level achieved as a measure of validity. Secondary outcomes assessed feasibility, resident self-evaluation versus faculty evaluation, the distributions among subcompetencies, and responsiveness over time.

RESULTS:
Between February 2014 and February 2016, 856 orthopedic surgery patient care subcompetency evaluations were completed (1.2 evaluations per day). Residents promptly requested feedback after a procedure (median = 0 days, interquartile range: 0-2), and faculty responded within 2 days in 51% (median = 2 days, interquartile range: 0-13). Primary outcome showed a correlation between PGY and competency level (r = 0.78, p < 0.001), with significant differences in competency among PGYs (p <0.001 by Kruskal-Wallis rank sum test). Self-evaluations by residents substantially agreed with faculty-assigned competency level (weighted Cohen's K = 0.72, p < 0.001). Resident classes beginning the study as PGY1, 2, and 3 separately demonstrated gains in competency over time (Spearman's rank correlation 0.39, 0.60, 0.59, respectively, each p < 0.001). There was significant variance in the number of evaluations submitted per subcompetency (median = 43, range: 6-113) and competency level assigned (p<0.01).

CONCLUSIONS:
Rapid tracking of trainee competency with milestone-based evaluations in a learner-centered mobile platform demonstrated validity, feasibility, and responsiveness. Next Accreditation System-mandated data may be efficiently collected and used for trainee and program self-study.
Comparison of Male vs Female Resident Milestone Evaluations by Faculty during Emergency Medicine Residency Training


IMPORTANCE:
Although implicit bias in medical training has long been suspected, it has been difficult to study using objective measures, and the influence of sex and gender in the evaluation of medical trainees is unknown. The emergency medicine (EM) milestones provide a standardized framework for longitudinal resident assessment, allowing for analysis of resident performance across all years and programs at a scope and level of detail never previously possible.

OBJECTIVE:
To compare faculty-observed training milestone attainment of male vs female residency training.

DESIGN, SETTING, AND PARTICIPANTS:
This multicenter, longitudinal, retrospective cohort study took place at 8 community and academic EM training programs across the United States from July 1, 2013, to July 1, 2015, using a real-time, mobile-based, direct-observation evaluation tool. The study examined 33 456 direct-observation subcompetency evaluations of 359 EM residents by 285 faculty members.

MAIN OUTCOMES AND MEASURES:
Milestone attainment for male and female EM residents as observed by male and female faculty throughout residency and analyzed using multilevel mixed-effects linear regression modeling.

RESULTS:
A total of 33 456 direct-observation evaluations were collected from 359 EM residents (237 men [66.0%] and 122 women [34.0%]) by 285 faculty members (194 men [68.1%] and 91 women [31.9%]) during the study period. Female and male residents achieved similar milestone levels during the first year of residency. However, the rate of milestone attainment was 12.7% (0.07 levels per year) higher for male residents through all of residency (95% CI, 0.04-0.09). By graduation, men scored approximately 0.15 milestone levels higher than women, which is equivalent to 3 to 4 months of additional training, given that the average resident gains approximately 0.52 levels per year using our model (95% CI, 0.49-0.54). No statistically significant differences in scores were found based on faculty evaluator gender (effect size difference, 0.02 milestone levels; 95% CI for males, -0.09 to 0.11) or evaluator-evaluatee gender pairing (effect size difference, -0.02 milestone levels; 95% CI for interaction, -0.05 to 0.01).

CONCLUSIONS AND RELEVANCE:
Although male and female residents receive similar evaluations at the beginning of residency, the rate of milestone attainment throughout training was higher for male than female residents across all EM subcompetencies, leading to a gender gap in evaluations that continues until graduation. Faculty should be cognizant of possible gender bias when evaluating medical trainees.
Does One Size Fit All? Examining the Application of Neurosurgery Residency Milestones Developed in the USA to a Taiwanese Culture


BACKGROUND:
The Milestone Project was launched in 2009, charging specialties to develop specific educational accomplishments required to establish clinical competency. The milestone assessment method was first introduced to Taiwan in 2013 and prior to applying milestone assessments to our medical education system, the validity and reliability of these questionnaires needed to be evaluated.

METHOD:
Twenty neurosurgical faculty members representing 3 clinical divisions and all 4 branch institutes completed milestone questionnaires for 26 residents semiannually resulting in 435 resident assessments being collected and analyzed.

RESULTS:
Cronbach's a, KR-20, and Kendall's W were used to show acceptable reliability and validity. Rater consistencies for non-skilled parts found that rater consistency progressively improved over time. Not all raters were able to assess the residents for the skilled parts resulting in non-assessable rates ranging from 9.5% to 89.4%. For both non-skilled and skilled items, milestone level as assessed by the staff improved as the resident progressed from R3 to R6 in the residency program and showed that the milestone achievement level for an R3 was lower than that of an R6.

CONCLUSION:
Milestone assessments have high reliability and may be a helpful assessment tool. Although milestone assessment can provide thorough feedback concerning performance and the content of the training program, they may not perfectly suit all residency-training programs, especially in different countries or different cultures. Modifications should be done before applying milestones to different areas so that the results can truly reflect the progress and condition of the training and learning process.
Evaluating Surgical Residents Quickly and Easily against the Milestones Using Electronic Formative Feedback


OBJECTIVE:
This study was conducted to assess the effectiveness of a newly implemented electronic web-based review system created at our institution for evaluating resident performance relative to established milestones.

DESIGN:
Retrospective review of data collected from a survey of general surgery faculty and residents.

SETTING:
Tertiary care teaching hospital system and independent academic medical center.

PARTICIPANTS:
A total of 12 general surgery faculty and 17 general surgery residents participated in this study. The survey queried the level of satisfaction before and after the adoption of QuickNotes using several statements scored on a 5-point scale, with 1 being the lowest rating as "not satisfied," and 5 being the highest rating as "completely satisfied."

RESULTS:
The weighted average improvements from pre- to post-QuickNotes implementation for the faculty responding to the survey ranged from 10% to 40%; weighted average improvements for the residents responding to the survey ranged from 5% to 73%. For the survey of faculty, both sets of weighted averages tended to be higher than the weighted average for the resident's survey responses. The highest rated topic was the faculty's level of satisfaction with the "frequency to provide feedback" with a post-QuickNotes implementation weighted average of 4.25, closely followed by the residents' level of satisfaction with the "evaluation includes positive feedback" with a post-QuickNotes implementation weighted average of 4.24. The most notable increases in weighted averages from preimplementation to postimplementation were noted for "overall satisfaction" (20% increase for faculty, 37% for residents), "reflects actual criteria that matter" (36% increase for faculty, 73% for residents), faculty "opportunity for follow-up" (increase of 40%), resident "reflects overall trends" (increase of 37%), and resident "provides new information about my performance" (increase of 37%).

CONCLUSIONS:
Our institutional adoption of QuickNotes into the resident evaluation process has been associated with an overall increased level of satisfaction in the evaluation process by both faculty and residents. The design of QuickNotes facilitates its integration into the resident training environment, as it is web based, easy to use, and has no additional cost over the standard New Innovations subscription. Although it is designed to capture snapshots of trainee behavior and performance, monthly reports through QuickNotes can be used effectively in conjunction with the more traditional end-of-rotation evaluations to show trends, identify areas of strength that should be reinforced, demonstrate areas needing improvement, allow for a more tailored individual education plan to be developed, and permit a more accurate determination of milestone progression.
Implementation of a Needs-Based, Online Feedback Tool for Anesthesia Residents With Subsequent Mapping of the Feedback to the ACGME Milestones


BACKGROUND:
Optimizing feedback that residents receive from faculty is important for learning. The goals of this study were to (1) conduct focus groups of anesthesia residents to define what constitutes optimal feedback; (2) develop, test, and implement a web-based feedback tool; and (3) then map the contents of the written comments collected on the feedback tool to the Accreditation Council for Graduate Medical Education (ACGME) anesthesiology milestones.

METHODS:
All 72 anesthesia residents in the program were invited to participate in 1 of 5 focus groups scheduled over a 2-month period. Thirty-seven (51%) participated in the focus groups and completed a written survey on previous feedback experiences. On the basis of the focus group input, an initial online feedback tool was pilot-tested with 20 residents and 62 feedback sessions, and then a final feedback tool was deployed to the entire residency to facilitate the feedback process. The completed feedback written entries were mapped onto the 25 ACGME anesthesiology milestones.

RESULTS:
Focus groups revealed 3 major barriers to good feedback: (1) too late such as, for example, at the end of month-long clinical rotations, which was not useful because the feedback was delayed; (2) too general and not specific enough to immediately remedy behavior; and (3) too many in that the large number of evaluations that existed were unhelpful such as those with unclear behavioral anchors compromised the overall feedback culture. Thirty residents (42% of 72 residents in the program) used the final online feedback tool with 121 feedback sessions with 61 attendings on 15 rotations at 3 hospital sites. The number of feedback tool uses per resident averaged 4.03 (standard deviation 5.08, median 2, range 1-21, 25th-75th % quartile 1-4). Feedback tool uses per faculty averaged 1.98 (standard deviation 3.2, median 1, range 1-25, 25th-75th % quartile 1-2). For the feedback question item "specific learning objective demonstrated well by the resident," this yielded 296 milestone-specific responses. The majority (71.3%) were related to the patient care competency, most commonly the anesthetic plan and conduct (35.8%) and airway management (11.1%) milestones; 10.5% were related to the interpersonal and communication skills competency, most commonly the milestones communication with other professionals (4.4%) or with patients and families (4.4%); and 8.4% were related to the practice-based learning and improvement competency, most commonly self-directed learning (6.1%). For the feedback tool item "specific learning objective that resident may improve," 67.0% were related to patient care, most commonly anesthetic plan and conduct (33.5%) followed by use/interpretation of monitoring and equipment (8.5%) and airway management (8.5%); 10.2% were related to practice-based learning and improvement, most commonly self-directed learning (6.8%); and 9.7% were related to the systems-based practice competency.

CONCLUSIONS:
Resident focus groups recommended that feedback be timely and specific and be structured around a tool. A customized online feedback tool was developed and implemented. Mapping of the free-text feedback comments may assist in assessing milestones. Use of the feedback tool was lower than expected, which may indicate that it is just 1 of many implementation steps required for behavioral and culture change to support a learning environment with frequent and useful feedback.
Number of Weeks Rotating in the Emergency Department Has a Greater Effect on Ultrasound Milestone Competency than a Dedicated Ultrasound Rotation


OBJECTIVES:
Ultrasound (US) is vital to modern emergency medicine (EM). Across residencies, there is marked variability in US training. The "goal-directed focused US" part of the Milestones Project states that trainees must correctly acquire and interpret images to achieve a level 3 milestone. Standardized methods by which programs teach these skills have not been established. Our goal was to determine whether residents could achieve level 3 with or without a dedicated US rotation.

METHODS:
Thirty-three first- and second-year residents were assigned to control (no rotation) and intervention (US rotation) groups. The intervention group underwent a 2-week curriculum in vascular access, the aorta, echocardiography, focused assessment with sonography for trauma, and pregnancy. To test acquisition, US-trained emergency medicine physicians administered an objective structured clinical examination. To test interpretation, residents had to identify normal versus abnormal findings. Mixed-model logistic regression tested the association of a US rotation while controlling for confounders: weeks in the emergency department (ED) as a resident, medical school US rotation, and postgraduate years.

RESULTS:
For image acquisition, medical school US rotation and weeks in the ED as a resident were significant (P = .03; P = .04) whereas completion of a US rotation and postgraduate years were not significant. For image interpretation, weeks in the ED as a resident was the only significant predictor of performance (P = .002) whereas completion of a US rotation and medical school US rotation were not significant.

CONCLUSIONS:
To achieve a level 3 milestone, weeks in the ED as a resident were significant for mastering image acquisition and interpretation. A dedicated US rotation did not have a significant effect. A medical school US rotation had a significant effect on image acquisition but not interpretation. Further studies are needed to best assess methods to meet US milestones.
How Effective are New Milestones Assessments at Demonstrating Resident Growth? 1 Year of Data


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.
Piloting the Mobile Medical Milestones Application (M3App©): A Multi-Institution Evaluation


BACKGROUND AND OBJECTIVES:
Competency-based evaluation of the Accreditation Council for Graduate Medical Education (ACGME) Milestones requires the development of new evaluation tools that can better capture learners' behavior. This study describes the implementation and initial assessment of an innovative point-of-care mobile application, the M3App©, linked to the Family Medicine Milestones.

METHODS:
Seven family medicine residency programs in North Carolina implemented the M3App.© Program faculty and residents were surveyed prior to implementation regarding current evaluation methods and their quality and use and acceptability of electronic evaluation tools. Surveys were repeated after implementation for comparison.

RESULTS:
All seven programs successfully implemented the M3App. Most faculty members found the tool well designed, easy to use, beneficial to the quality and efficiency of feedback they provide, and to their knowledge of Milestones. Residents reported significant increases in the volume and quality of written feedback they receive.

CONCLUSIONS:
The M3App provides an efficient, convenient tool for assessing Milestones that can improve the quantity and quality of feedback residents receive from faculty. Improved faculty perception of knowledge of Milestones after M3App implementation suggests that the tool is also effective for faculty development.
Connecting Milestones to Observable Clinical Performance through Standardized Pediatric Cardiology Rotation Evaluations: Report and Reflections on a Novel Pilot Project


**ABSTRACT**

Uniform evaluations were designed and implemented by pediatric cardiology program directors. The evaluation forms incorporated ACGME/ABP subcompetencies into clinical observations. Fellow milestone assessment was achievable through the use of the evaluation forms. Early survey results suggest that these were an improvement over existing tools.
Using the ACGME Milestones for Resident Self-Evaluation and Faculty Engagement


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 a 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 a: 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.
A Multi-Source Feedback Tool for Measuring a Subset of Pediatrics Milestones


BACKGROUND:
The Pediatrics Milestones Assessment Pilot employed a new multisource feedback (MSF) instrument to assess nine Pediatrics Milestones among interns and subinterns in the inpatient context.

OBJECTIVE:
To report validity evidence for the MSF tool for informing milestone classification decisions.

METHODS:
We obtained MSF instruments by different raters per learner per rotation. We present evidence for validity based on the unified validity framework.

RESULTS:
One hundred and ninety two interns and 41 subinterns at 18 Pediatrics residency programs received a total of 1084 MSF forms from faculty (40%), senior residents (34%), nurses (22%), and other staff (4%). Variance in ratings was associated primarily with rater (32%) and learner (22%). The milestone factor structure fit data better than simpler structures. In domains except professionalism, ratings by nurses were significantly lower than those by faculty and ratings by other staff were significantly higher. Ratings were higher when the rater observed the learner for longer periods and had a positive global opinion of the learner. Ratings of interns and subinterns did not differ, except for ratings by senior residents. MSF-based scales correlated with summative milestone scores.

CONCLUSION:
We obtain moderately reliable MSF ratings of interns and subinterns in the inpatient context to inform some milestone assignments.
Entrusting Observable Practice Activities and Milestones over the 36 Months of an Internal Medicine Residency


PURPOSE: Competency-based medical education and milestone reporting have led to increased interest in work-based assessments using entrustment over time as an assessment framework. Little is known about data collected from these assessments during residency. This study describes the results of entrustment of discrete work-based skills over 36 months in the University of Cincinnati internal medicine (IM) residency program.

METHOD: Attending physician and peer/allied health assessors provided entrustment ratings of resident performance on work-based observable practice activities (OPAs) mapped to Accreditation Council for Graduate Medicine Education/American Board of Internal Medicine reporting milestones for IM. These data were translated into milestones data and tracked longitudinally. The authors analyzed data from this new entrustment system’s first 36 months (July 2012-June 2015).

RESULTS: During the 36-month period, assessors made 364,728 milestone assessments (mapped from OPAs) of 189 residents. Residents received an annualized average of 83 assessment encounters, producing means of 3,987 milestone assessments and 4,325 words of narrative assessment. Mean entrustment ratings (range 1-5) from all assessors for all milestones rose from 2.46 for first-month residents to 3.92 for 36th-month residents (r = 0.9252, P < .001). Attending physicians’ entrustment ratings were lower than peer/allied health assessors’ ratings. Medical knowledge and patient care milestones were rated lower than professionalism and interpersonal and communication skills milestones.

CONCLUSIONS: Entrustment of milestones appears to rise progressively over time, with differences by assessor type, competency, milestone, and resident. Further research is needed to elucidate the validity of these data in promotion, remediation, and reporting decisions.
Patients, Nurses, and Physicians Working Together to Develop a Discharge Entrustable Professional Activity Assessment Tool


PROBLEM:
The Accreditation Council for Graduate Medical Education milestones were written by physicians and thus may not reflect all the behaviors necessary for physicians to optimize their performance as a key member of an interprofessional team.

APPROACH:
From April to May 2013, the authors, Educational Research Outcomes Collaborative leaders, assembled interprofessional team discussion groups, including patients or family members, nurses, physician trainees, physician educators, and other staff (optional), at 11 internal medicine (IM) programs. Led by the site’s principal investigator, the groups generated a list of physician behaviors related to the entrustable professional activity (EPA) of a safe and effective discharge of a patient from the hospital, and prioritized those behaviors.

OUTCOMES:
A total of 182 behaviors were listed, with lists consisting of between 10 and 29 behaviors. Overall, the site principal investigators described all participants as emerging from the activity with a new understanding of the complexity of training physicians for the discharge EPA. The authors batched behaviors into six components of a safe and effective discharge: medication reconciliation, discharge summary, patient/caregiver communication, team communication, active collaboration, and anticipation of posthospital needs. Specific, high-priority behavior examples for each component were identified, and an assessment tool for direct observation was developed for the discharge EPA.

NEXT STEPS:
The authors are currently evaluating trainee and educator perceptions of the assessment tool after implementation in 15 IM programs. Additional next steps include developing tools for other EPAs, as well as a broader evaluation of patient outcomes in the era of milestone-based assessment.
Learnings from the Pilot Implementation of Mobile Medical Milestones Application


BACKGROUND:
Implementation of the educational milestones benefits from mobile technology that facilitates ready assessments in the clinical environment. We developed a point-of-care resident evaluation tool, the Mobile Medical Milestones Application (M3App), and piloted it in 8 North Carolina family medicine residency programs.

OBJECTIVE:
We sought to examine variations we found in the use of the tool across programs and explored the experiences of program directors, faculty, and residents to better understand the perceived benefits and challenges of implementing the new tool.

METHODS:
Residents and faculty completed presurveys and postsurveys about the tool and the evaluation process in their program. Program directors were interviewed individually. Interviews and open-ended survey responses were analyzed and coded using the constant comparative method, and responses were tabulated under themes.

RESULTS:
Common perceptions included increased data collection, enhanced efficiency, and increased perceived quality of the information gathered with the M3App. Residents appreciated the timely, high-quality feedback they received. Faculty reported becoming more comfortable with the tool over time, and a more favorable evaluation of the tool was associated with higher utilization. Program directors reported improvements in faculty knowledge of the milestones and resident satisfaction with feedback.

CONCLUSIONS:
Faculty and residents credited the M3App with improving the quality and efficiency of resident feedback. Residents appreciated the frequency, proximity, and specificity of feedback, and faculty reported the app improved their familiarity with the milestones. Implementation challenges included lack of a physician champion and competing demands on faculty time.
Using Milestones as Evaluation Metrics during an Emergency Medicine Clerkship


BACKGROUND:
The Accreditation Council for Graduate Medical Education's (ACGME) Milestones presumes graduating medical students will enter residency proficient at Milestone level 1 for 23 skills. The Next Accreditation System now includes Milestones for each postgraduate specialty, and it is unlikely that schools will document every emergency medicine (EM) applicant's EM-specific skills in their performance evaluation.

OBJECTIVES:
The goals of this research were to determine if assessment of the Milestones was feasible during a medical student clerkship and examine the proportion of medical students performing at Milestone level 1.

METHODS:
This study was conducted at a center with Liaison Committee on Medical Education-approved medical training and a 4-year EM residency. Using traditional clerkship, we studied the feasibility of an ACGME EM Milestones-based clerkship assessment. Data led to redesign of the clerkship and its evaluation process, including all level 1 anchor(s) to add "occasionally" (>60%), "usually" (>80%), and "always" (100%) on a Likert scale to on-shift assessment forms.

RESULTS:
During the feasibility phase (2013-14), 75 students rotated though the clerkship; 55 evaluations were issued and 50 contained the Milestone summary. Eight deficiencies were noted in Milestone 12 and three in Milestone 14. After changes, 49 students rotated under the new evaluation rubric. Of 575 completed on-shift evaluations, 16 Milestone deficiencies were noted. Of 41 institutional evaluations issued, only one student had deficiencies noted, all of which pertained to patient care. All evaluations in this second cohort contained each student's Milestone proficiency.

CONCLUSIONS:
Assessment of the Milestones is feasible. Communication of ACGME EM Milestone proficiency may identify students who require early observation or remediation. The majority of students meet the anchors for the Milestones, suggesting that clerkship assessment with the ACGME EM Milestones does not adequately differentiate students.
Shortening the Miles to the Milestones: Connecting EPA-Based Evaluations to ACGME Milestone Reports for Internal Medicine Residency Programs


ABSTRACT
The Next Accreditation System requires internal medicine training programs to provide the Accreditation Council for Graduate Medical Education (ACGME) with semiannual information about each resident's progress in 22 subcompetency domains. Evaluation of resident "trustworthiness" in performing entrustable professional activities (EPAs) may offer a more tangible assessment construct than evaluations based on expectations of usual progression toward competence. However, translating results from EPA-based evaluations into ACGME milestone progress reports has proven to be challenging because the constructs that underlay these two systems differ. The authors describe a process to bridge the gap between rotation-specific EPA-based evaluations and ACGME milestone reporting. Developed at the University of Washington in 2012 and 2013, this method involves mapping EPA-based evaluation responses to "milestone elements," the narrative descriptions within the columns of each of the 22 internal medicine subcompetencies. As faculty members complete EPA-based evaluations, the mapped milestone elements are automatically marked as "confirmed." Programs can maintain a database that tallies the number of times each milestone element is confirmed for a resident; these data can be used to produce graphical displays of resident progress along the internal medicine milestones. Using this count of milestone elements allows programs to bridge the gap between faculty assessments of residents based on rotation-specific observed activities and semiannual ACGME reports based on the internal medicine milestones. Although potentially useful for all programs, this method is especially beneficial to large programs where clinical competency committee members may not have the opportunity for direct observation of all residents.
Mapping Direct Observations from Objective Structured Clinical Examinations to the Milestones across Specialties


BACKGROUND:
Little is known about residents’ performance on the milestones at the institutional level. Our institution formed a work group to explore this using an institutional-level curriculum and residents’ evaluation of the milestones.

OBJECTIVE:
We assessed whether beginner-level milestones for interpersonal and communication skills (ICS) related to observable behaviors in ICS-focused objective structured clinical examinations (OSCEs) for postgraduate year (PGY) 1 residents across specialties.

METHODS:
The work group compared ICS subcompetencies across 12 programs to identify common beginner-level physician-patient communication milestones. The selected ICS milestone sets were compared for common language with the ICS-OSCE assessment tool—the Kalamazoo Essential Elements of Communication Checklist-Adapted (KEECC-A). To assess whether OSCE scores related to ICS milestone scores, all PGY-1 residents from programs that were part of Next Accreditation System Phase 1 were identified; their OSCE scores from July 2013 to June 2014 and ICS subcompetency scores from December 2014 were compared.

RESULTS:
The milestones for 10 specialties and the transitional year had at least 1 ICS subcompetency that related to physician-patient communication. The language of the ICS beginner-level milestones appears similar to behaviors outlined in the KEECC-A. All 60 residents with complete data received at least a beginner-level ICS subcompetency score and at least a satisfactory score on all 3 OSCEs.

CONCLUSIONS:
The ICS-OSCE scores for PGY-1 residents appear to relate to beginner-level milestones for physician-patient communication across multiple specialties.
Rating the Quality of Entrustable Professional Activities: Content Validation and Associations with the Clinical Context


BACKGROUND:
Entrustable professional activities (EPAs) have been developed to assess resident physicians with respect to Accreditation Council for Graduate Medical Education (ACGME) competencies and milestones. Although the feasibility of using EPAs has been reported, we are unaware of previous validation studies on EPAs and potential associations between EPA quality scores and characteristics of educational programs.

OBJECTIVES:
Our aim was to validate an instrument for assessing the quality of EPAs for assessment of internal medicine residents, and to examine associations between EPA quality scores and features of rotations.

DESIGN:
This was a prospective content validation study to design an instrument to measure the quality of EPAs that were written for assessing internal medicine residents.

PARTICIPANTS:
Residency leadership at Mayo Clinic, Rochester participated in this study. This included the Program Director, Associate program directors and individual rotation directors.

INTERVENTIONS:
The authors reviewed salient literature. Items were developed to reflect domains of EPAs useful for assessment. The instrument underwent further testing and refinement. Each participating rotation director created EPAs that they felt would be meaningful to assess learner performance in their area. These 229 EPAs were then assessed with the QUEPA instrument to rate the quality of each EPA.

MAIN MEASURES:
Performance characteristics of the QUEPA are reported. Quality ratings of EPAs were compared to the primary ACGME competency, inpatient versus outpatient setting and specialty type.

KEY RESULTS:
QUEPA tool scores demonstrated excellent reliability (ICC range 0.72 to 0.94). Higher ratings were given to inpatient versus outpatient (3.88, 3.66; p = 0.03) focused EPAs. Medical knowledge EPAs scored significantly lower than EPAs assessing other competencies (3.34, 4.00; p < 0.0001).

CONCLUSIONS:
The QUEPA tool is supported by good validity evidence and may help in rating the quality of EPAs developed by individual programs. Programs should take care when writing EPAs for the outpatient setting or to assess medical knowledge, as these tended to be rated lower.
Reflections in a Time of Transition: Orthopaedic Faculty and Resident Understanding of Accreditation Schemes and Opinions on Surgical Skills Feedback


INTRODUCTION:
Orthopaedic surgery is one of the first seven specialties that began collecting Milestone data as part of the Accreditation Council for Graduate Medical Education's Next Accreditation System (NAS) rollout. This transition from process-based advancement to outcome-based education is an opportunity to assess resident and faculty understanding of changing paradigms, and opinions about technical skill evaluation.

METHODS:
In a large academic orthopaedic surgery residency program, residents and faculty were anonymously surveyed. A total of 31/32 (97%) residents and 29/53 (55%) faculty responded to Likert scale assessments and provided open-ended responses. An internal end-of-rotation audit was conducted to assess timeliness of evaluations. A mixed-method analysis was utilized, with nonparametric statistical testing and a constant-comparative qualitative method.

RESULTS:
There was greater familiarity with the six core competencies than with Milestones or the NAS (p<0.05). A majority of faculty and residents felt that end-of-rotation evaluations were not adequate for surgical skills feedback. Fifty-eight per cent of residents reported that end-of-rotation evaluations were rarely or never filled out in a timely fashion. An internal audit demonstrated that more than 30% of evaluations were completed over a month after rotation end. Qualitative analysis included themes of resident desire for more face-to-face feedback on technical skills after operative cases, and several barriers to more frequent feedback.

DISCUSSION:
The NAS and outcome-based education have arrived. Residents and faculty need to be educated on this changing paradigm. This transition period is also a window of opportunity to address methods of evaluation and feedback. In our orthopaedic residency, trainees were significantly less satisfied than faculty with the amount of technical and surgical skills feedback being provided to trainees. The quantitative and qualitative analyses converge on one theme: a desire for frequent, explicit, timely feedback after operative cases. To overcome the time-limited clinical environment, feedback tools need to be easily integrated and efficient. Creative solutions may be needed to truly achieve outcome-based graduate medical education.
ENTRUSTABLE PROFESSIONAL ACTIVITIES: TEN THINGS RADIOLOGISTS DO


RATIONALE AND OBJECTIVES:
Learner assessment in medical education has undergone tremendous change over the past two decades. During this time frame, the concept of Entrustable Professional Activities (EPAs) was introduced to guide the faculty when making competency-based decisions on the level of supervision required by trainees. EPAs are gaining momentum in medical education as a basis for decisions related to transitioning from residency training to clinical practice. The purpose of this article is twofold: (1) define EPAs for radiology (EPA-R) and (2) illustrate radiology-specific examples of these EPAs.

MATERIALS AND METHODS:
A multi-institutional work group composed of members of the Alliance of Directors and Vice Chairs of Education in Radiology convened at the 2015 Association of University Radiologists annual meeting to discuss radiology EPAs. The EPAs initially developed by the Accreditation Council for Graduate Medical Education (ACGME) Radiology milestone work group and the resultant ACGME Radiology milestones formed the basis for this discussion.

RESULTS:
A total of 10 radiology EPAs and illustrative vignettes were developed to help radiology educators and trainees better understand milestone assessment and how this translates to the necessary skills and responsibilities of practicing radiologists. Examples of EPA mapping to the ACGME subcompetencies and methods of assessment were included.

CONCLUSIONS:
EPAs offer an opportunity to improve our approach to training by increasing our focus on how we provide appropriate supervision to our residents and assess their progress. In this work, through suggested lists and vignettes, we have attempted to establish the framework for further discussion and development of EPA-Rs.
Comprehensive Observations of Resident Evolution: A Novel Method for Assessing Procedure-Based Residency Training


BACKGROUND:
Assessment of surgical skills in the operating room remains a challenge. Increasing documentation requirements of the Accreditation Council for Graduate Medical Education are necessitating mechanisms to document trainee competence without hindering operative turnover. The authors created a comprehensive electronic resource to facilitate plastic surgery training program compliance with changes mandated by Next Accreditation System Milestones and the ACGME.

METHODS:
In 2013, the authors implemented the Comprehensive Observations of Resident Evolution, or CORE, a Web-based tool to assess plastic surgery residents. It comprises a rapid electronic assessment of resident operating room performance completed after each surgery; a data dashboard displaying graphical summaries of resident progress by case, Milestone, or current procedural terminology code; and an electronic Milestones tracker (MileMarker), which enables ongoing trainee assessments.

RESULTS:
From January through October of 2014, 24 residents completed nearly 1300 Operative Entrustability Assessments. Thirty-eight percent of residents reported more immediate feedback regarding operative performance. The assessment demonstrates construct validity, which distinguishes novice residents from experienced residents. Individual case data identify resident-specific operative strengths and weaknesses. Using assessment data, the first two Clinical Competency Committee reviews were 81 percent and 87 percent shorter than Milestones pilot test site reports (average, 11.5 and 8 minutes versus 60 minutes per resident, respectively).

CONCLUSIONS:
Comprehensive Observations of Resident Evolution is capable of capturing operative performance data on all operating room cases by primary current procedural terminology code. It increases immediate attending/trainee feedback and assessment transparency, enables trainee self-monitoring, and informs end-of-rotation reviews, program-wide assessments, and tailoring of training to address specific needs. It is a valuable resource for tracking resident progress in real-time while maintaining compliance with evolving ACGME requirements.
Driving Care Quality: Aligning Trainee Assessment and Supervision through Practical Application of Entrustable Professional Activities, Competencies, and Milestones


ABSTRACT
To address the long-standing challenge of meaningful trainee assessment, the authors reviewed and expanded on the Accountable Assessment for Quality Care and Supervision (AAQCS) equation. The equation proposes that care quality is the product of the interaction between trainee performance (measured by workplace assessment) and supervision (required level of intervention to ensure care quality) in the context of the environment where the care occurs: Trainee performance × Appropriate supervision = Safe, effective patient-centered care. Assessing trainee performance and matching that performance to “appropriate” supervision, however, is fraught with challenges. The authors suggest a unifying framework that integrates entrustable professional activities (EPAs), competencies, and milestones to inform trainee assessment and supervision, thereby enabling the practical application of the AAQCS equation in the workplace. Because the unit of measure for an EPA is the outcome of whether the trainee can safely and effectively perform the professional activity without supervision, the proposed unifying framework directly aligns with the dependent variable in the AAQCS equation: care quality. The value of applying a unifying framework that integrates EPAs, competencies, and milestones to the AAQCS equation in the clinical learning environment lies in its ability to provide supervisors with a shared mental model of performance expectations for trainees, reducing unwanted variability and improving assessment accuracy; guidance for aligning performance milestones of trainees with the needed level of supervisor intervention to ensure care quality; and substrate for specific feedback to improve the trainee's professional development as a way to ensure future care quality.
A Pilot Study of Orthopaedic Resident Self-Assessment Using a Milestones’ Survey Just Prior to Milestones Implementation


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.
Competency Evaluations in the Next Accreditation System: Contributing to Guidelines and Implications


CONSTRUCT:
This study examines validity evidence of end-of-rotation evaluation scores used to measure competencies and milestones as part of the Next Accreditation System (NAS) of the Accreditation Council for Graduate Medical Education (ACGME).

BACKGROUND:
Since the implementation of the milestones, end-of-rotation evaluations have surfaced as a potentially useful assessment method. However, validity evidence on the use of rotation evaluation scores as part of the NAS has not been studied. This article examines validity evidence for end-of-rotation evaluations that can contribute to developing guidelines that support the NAS.

APPROACH:
Data from 2,701 end-of-rotation evaluations measuring 21 out of 22 Internal Medicine milestones for 142 residents were analyzed (July 2013-June 2014). Descriptive statistics were used to measure the distribution of ratings by evaluators (faculty, n = 116; fellows, n = 59; peer-residents, n = 131), by postgraduate years. Generalizability analysis and higher order confirmatory factor analysis were used to examine the internal structure of ratings. Psychometric implications for combining evaluation scores using composite score reliability were examined.

RESULTS:
Milestone ratings were significantly higher for each subsequent year of training (15/21 milestones). Faculty evaluators had greater variability in ratings across milestones, compared to fellows and residents; faculty ratings were generally correlated with milestone ratings from fellows (r = .45) and residents (r = .25), but lower correlations were found for Professionalism and Interpersonal and Communication Skills. The \( \phi \)-coefficient was .71, indicating good reliability. Internal structure supported a 6-factor solution, corresponding to the hierarchical relationship between the milestones and the 6 core competencies. Evaluation scores corresponding to Patient Care, Medical Knowledge, and Practice-Based Learning and Improvement had higher correlations to milestones reported to the ACGME. Mean evaluation ratings predicted problem residents (odds ratio = 5.82, p < .001).

CONCLUSIONS:
Guidelines for rotation evaluations proposed in this study provide useful solutions that can help program directors make decisions on resident progress and contribute to assessment systems in graduate medical education.
Use of a Formal Assessment Instrument for Evaluation of Resident Operative Skills in Pediatric Neurosurgery


OBJECT:
Currently there is no standardized tool for assessment of neurosurgical resident performance in the operating room. In light of enhanced requirements issued by the Accreditation Council for Graduate Medical Education's Milestone Project and the Matrix Curriculum Project from the Society of Neurological Surgeons, the implementation of such a tool seems essential for objective evaluation of resident competence. Beyond compliance with governing body guidelines, objective assessment tools may be useful to direct early intervention for trainees performing below the level of their peers so that they may be given more hands-on teaching, while strong residents can be encouraged by faculty members to progress to conducting operations more independently with passive supervision. The aims of this study were to implement a validated assessment tool for evaluation of operative skills in pediatric neurosurgery and determine its feasibility and reliability.

METHODS:
All neurosurgery residents completing their pediatric rotation over a 6-month period from January 1, 2014, to June 30, 2014, at the authors' institution were enrolled in this study. For each procedure, residents were evaluated by means of a form, with one copy being completed by the resident and a separate copy being completed by the attending surgeon. The evaluation form was based on the validated Objective Structured Assessment of Technical Skills for Surgery (OSATS) and used a 5-point Likert-type scale with 7 categories: respect for tissue; time and motion; instrument handling; knowledge of instruments; flow of operation; use of assistants; and knowledge of specific procedure. Data were then stratified by faculty versus resident (self-) assessment; postgraduate year level; and difficulty of procedure. Descriptive statistics (means and SDs) were calculated, and the results were compared using the Wilcoxon signed-rank test and Student t-test. A p value < 0.05 was considered statistically significant. RESULTS Six faculty members, 1 fellow, and 8 residents completed evaluations for 299 procedures, including 32 ventriculoperitoneal (VP) shunt revisions, 23 VP shunt placements, 19 endoscopic third ventriculostomies, and 18 craniotomies for tumor resection. There was no significant difference between faculty and resident self-assessment scores overall or in any of the 7 domains scores for each of the involved residents. On self-assessment, senior residents scored themselves significantly higher (p < 0.02) than junior residents overall and in all domains except for "time and motion." Faculty members scored senior residents significantly higher than junior residents only for the "knowledge of instruments" domain (p = 0.05). When procedure difficulty was considered, senior residents’ scores from faculty members were significantly higher (p = 0.04) than the scores given to junior residents for expert procedures only. Senior residents’ self-evaluation scores were significantly higher than those of junior residents for both expert (p = 0.03) and novice (p = 0.006) procedures.

CONCLUSIONS:
OSATS is a feasible and reliable assessment tool for the comprehensive evaluation of neurosurgery resident performance in the operating room. The authors plan to use this tool to assess resident operative skill development and to improve direct resident feedback.
ABSTRACT
Graduate medical education is increasingly focused on patient safety and quality improvement; training programs must adapt their curriculum to address these changes. We propose a novel curriculum for emergency medicine (EM) residency training programs specifically addressing patient safety, systems-based management, and practice-based performance improvement, called "EM Debates." Following implementation of this educational curriculum, we performed a cross-sectional study to evaluate the curriculum through resident self-assessment. Additionally, a cross-sectional study to determine the ED clinical competency committee's (CCC) ability to assess residents on specific competencies was performed. Residents were overall very positive towards the implementation of the debates. Of those participating in a debate, 71% felt that it improved their individual performance within a specific topic, and 100% of those that led a debate felt that they could propose an evidence-based approach to a specific topic. The CCC found that it was easier to assess milestones in patient safety, systems-based management, and practice-based performance improvement (sub-competencies 16, 17, and 19) compared to prior to the implementation of the debates. The debates have been a helpful venue to teach EM residents about patient safety concepts, identifying medical errors, and process improvement.
Gearing Up for Milestones in Surgery: Will Simulation Play a Role?


BACKGROUND:
The Consortium of American College of Surgeons-Accredited Education Institutes was created to promote patient safety through the use of simulation, develop new education and technologies, identify best practices, and encourage research and collaboration.

METHODS:
During the 7th Annual Meeting of the Consortium, leaders from a variety of specialties discussed how simulation is playing a role in the assessment of resident performance within the context of the Milestones of the Accreditation Council for Graduate Medical Education as part of the Next Accreditation System.

CONCLUSION:
This report presents experiences from several viewpoints and supports the utility of simulation for this purpose.
INTRODUCTION:
Emergency medicine (EM) milestones are used to assess residents' progress. While some milestone validity evidence exists, there is a lack of standardized tools available to reliably assess residents. Inherent to this is a concern that we may not be truly measuring what we intend to assess. The purpose of this study was to design a direct observation milestone assessment instrument supported by validity and reliability evidence. In addition, such a tool would further lend validity evidence to the EM milestones by demonstrating their accurate measurement.

METHODS:
This was a multi-center, prospective, observational validity study conducted at eight institutions. The Critical Care Direct Observation Tool (CDOT) was created to assess EM residents during resuscitations. This tool was designed using a modified Delphi method focused on content, response process, and internal structure validity. Paying special attention to content validity, the CDOT was developed by an expert panel, maintaining the use of the EM milestone wording. We built response process and internal consistency by piloting and revising the instrument. Raters were faculty who routinely assess residents on the milestones. A brief training video on utilization of the instrument was completed by all. Raters used the CDOT to assess simulated videos of three residents at different stages of training in a critical care scenario. We measured reliability using Fleiss' kappa and interclass correlations.

RESULTS:
Two versions of the CDOT were used: one used the milestone levels as global rating scales with anchors, and the second reflected a current trend of a checklist response system. Although the raters who used the CDOT routinely rate residents in their practice, they did not score the residents' performances in the videos comparably, which led to poor reliability. The Fleiss' kappa of each of the items measured on both versions of the CDOT was near zero.

CONCLUSION:
The validity and reliability of the current EM milestone assessment tools have yet to be determined. This study is a rigorous attempt to collect validity evidence in the development of a direct observation assessment instrument. However, despite strict attention to validity evidence, inter-rater reliability was low. The potential sources of reducible variance include rater- and instrument-based error. Based on this study, there may be concerns for the reliability of other EM milestone assessment tools that are currently in use.
Pathology Milestones: Assessing Clinical Competency by Committee


ABSTRACT
All Accreditation Council for Graduate Medical Education accredited pathology residency training programs are now required to evaluate residents using the new Pathology Milestones assessment tool. Similar to implementation of the 6 Accreditation Council for Graduate Medical Education competencies a decade ago, there have been challenges in implementation of the new milestones for many residency programs. The pathology department at the University of Iowa has implemented a process that divides the labor of the task in rating residents while also maintaining consistency in the process. The process is described in detail, and some initial trends in milestone evaluation are described and discussed. Our experience indicates that thoughtful implementation of the Pathology Milestones can provide programs with valuable information that can inform curricular changes.
Navigating the Next Accreditation System: A Dashboard for the Milestones


INTRODUCTION:
In July 2014, all residency programs accredited by the Accreditation Council for Graduate Medical Education (ACGME) were enrolled in a new system called the Next Accreditation System. Residency programs may not be clear on how best to comply with these new accreditation requirements. Large amounts of data must be collected, evaluated, and submitted twice a year to the council's Web-based data collection system. One challenge is that the new "end-of-rotation" evaluations must reflect specialty-specific milestones, on which many faculty members are not well versed. Like other residency programs, we tried to address the challenges using our local resources.

METHODS:
We used our existing electronic goals and objectives for each rotation coupled with appropriate end-of-rotation evaluations reflecting the specialty-specific milestones through a process of editing and mapping.

RESULTS:
Data extracted from these evaluations were added to an interactive dashboard that also contained evaluations on additional program-specific modifiers of residents' performance. A resident's final overall performance was visually represented on a plot graph. The novel dashboard included features to save evaluations for future comparisons and to track residents' progress during their entire training. It proved simple to use and was able to reduce the time needed for each resident evaluation to 5 to 10 minutes.

CONCLUSION:
This tool has made it much easier and less challenging for the members of our Clinical Competency Committee to start deliberation about each resident's performance.
Interprofessional Collaboration Milestones: Advocating for Common Assessment Criteria in Graduate Medical Education


BACKGROUND:
Milestone-based assessments of resident physicians inform critical decisions regarding resident competence and advancement. Thus, it is essential that milestone evaluations are based upon strong validity evidence and that consistent evaluation criteria are used across residency programs. A common approach to assessment of interprofessional collaboration milestones is particularly important since standardized measures of individual resident competence in interprofessional collaboration have not been established.

DISCUSSION:
We propose that assessments of interprofessional collaboration in graduate medical education meet common criteria, namely, these assessments should: 1) measure competency of an individual resident, 2) occur in the context of an interprofessional team, 3) be ascertained via direct observation of the resident, 4) be performed in a real-world clinical practice setting (such as a hospital ward, outpatient clinic, or operating room). We present the evidence-based rationale for these criteria and cite examples of published assessment instruments that fulfill one or more of the criteria, however further research is needed to ensure fidelity of assessments. The proposed criteria may assist residency educators as they endeavor to provide robust and consistent assessments of interprofessional collaboration milestones.
Development and Validation of an Assessment of Regional Anesthesia Ultrasound Interpretation Skills


BACKGROUND:
Interpretation of ultrasound images and knowledge of anatomy are essential skills for ultrasound-guided peripheral nerve blocks. Competency-based educational models promoted by the Accreditation Council for Graduate Medical Education require the development of assessment tools for the achievement of different competency milestones to demonstrate the longitudinal development of skills that occur during training.

METHODS:
A rigorous study guided by psychometric principles was undertaken to identify and validate the domains and items in an assessment of ultrasound interpretation skills for regional anesthesia. A survey of residents, academic faculty, and community anesthesiologists, as well as video recordings of experts teaching ultrasound-guided peripheral nerve blocks, was used to develop short video clips with accompanying multiple choice-style questions. Four rounds of pilot testing produced a 50-question assessment that was subsequently administered online to residents, fellows, and faculty from multiple institutions.

RESULTS:
Test results from 90 participants were analyzed with Item Response Theory model fitting indicating that a 47-item subset of the test fits the model well (P = 0.11). There was a significant linear relation between expected and predicted item difficulty (P < 0.001). Overall test scores increased linearly with higher levels of formal anesthesia training, regional anesthesia training, number of ultrasound-guided blocks performed per year, and a self-rating of regional anesthesia skill (all P < 0.001).

CONCLUSIONS:
This study provides evidence for the reliability, content validity, and construct validity of a 47-item multiple choice-style online test of ultrasound interpretation skills for regional anesthesia, which can be used as an assessment of competency milestone achievement in anesthesiology training.
Developing a Comprehensive Resident Education Evaluation System in the Era of Milestone Assessment


OBJECTIVES:
In an effort to move training programs toward competency-based education, the Accreditation Council for Graduate Medical Education (ACGME) introduced the Next Accreditation System (NAS), which organizes specific milestones regarding resident skills, knowledge, and abilities along a continuum. In order to foster innovation and creativity, the ACGME has provided programs with minimal guidelines regarding the optimal way to approach these milestones.

METHODS:
The education team at UT Southwestern embraced the milestones and developed a process in which performance assessment methods were critically evaluated, mapped onto an extrapolated performance list corresponding to the areas required by the ACGME milestones, and filled gaps in the previous system by modifying evaluation tools and creating new program components.

RESULTS:
Although the authors are early in the evolution of applying the new milestones system, this approach has thus far allowed them to comprehensively evaluate the residents and the program in an efficient and effective fashion, with notable improvements compared to the prior approach.

CONCLUSIONS:
The authors hope that these experiences can inform others embarking upon similar journeys with the milestones.
Implementation of Nephrology Subspecialty Curricular Milestones


ABSTRACT

Beginning in the 2014-2015 training year, the US Accreditation Council for Graduate Medical Education (ACGME) required that nephrology Clinical Competency Committees assess fellows' progress toward 23 subcompetency "context nonspecific" internal medicine subspecialty milestones. Fellows' advancement toward the "ready for unsupervised practice" target milestone now is tracked in each of the 6 competencies: Patient Care, Medical Knowledge, Professionalism, Interpersonal Communication Skills, Practice-Based Learning and Improvement, and Systems-Based Practice. Nephrology program directors and subspecialty societies must define nephrology-specific "curricular milestones," mapped to the nonspecific ACGME milestones. Although the ACGME goal is to produce data that can discriminate between successful and underperforming training programs, the approach is at risk to produce biased, inaccurate, and unhelpful information. We map the ACGME internal medicine subspecialty milestones to our previously published nephrology-specific milestone schema and describe entrustable professional activities and other objective assessment tools that inform milestone decisions. Mapping our schema onto the ACGME subspecialty milestone reporting form allows comparison with the ACGME subspecialty milestones and the curricular milestones developed by the American Society of Nephrology Program Directors. Clinical Competency Committees may easily adapt and directly translate milestone decisions reached using our schema onto the ACGME internal medicine subspecialty competency milestone-reporting format.
Piloting a Structured Practice Audit to Assess ACGME Milestones in Written Handoff Communication in Internal Medicine


BACKGROUND:
Written communication skills are integral to patient care handoffs. Residency programs require feasible assessment tools that provide timely formative and summative feedback, ideally linked to the Accreditation Council for Graduate Medical Education Milestones.

OBJECTIVE:
We describe the use of 1 such tool-UPDATED-to assess written handoff communication skills in internal medicine interns.

METHODS:
During 2012-2013, the authors piloted a structured practice audit at 1 academic institution to audit written sign-outs completed by 45 interns, using the UPDATED tool, which scores 7 aspects of sign-out communication linked to milestones. Intern sign-outs were audited by trained faculty members throughout the year. Results were incorporated into intern performance reviews and Clinical Competency Committees.

RESULTS:
A total of 136 sign-outs were audited (averaging 3.1 audits per intern). In the first trimester, 14 interns (31%) had satisfactory audit results. Five interns (11%) had critical deficiencies and received immediate feedback, and the remaining 26 (58%) were assigned future audits due to missing audits or unsatisfactory scores. In the second trimester, 21 interns (68%) had satisfactory results, 1 had critical deficiencies, and 9 (29%) required future audits. Nine of the 10 remaining interns in the final trimester had satisfactory audits. Faculty time was estimated at 10 to 15 minutes per sign-out audited.

CONCLUSIONS:
The UPDATED audit is a milestone-based tool that can be used to assess written sign-out communication skills in internal medicine residency programs. Future work is planned to adapt the tool for use by senior supervisory residents to appraise sign-outs in real time.
Use of Emergency Medicine Milestones as Items on End-of-Shift Evaluations Results in Overestimates of Residents' Proficiency Level


BACKGROUND:
The emergency medicine milestones were developed to provide more objective resident assessment than current methods. However, little is known about the best practices for applying the milestones in resident assessment.

OBJECTIVE:
We examined the utility of end-of-shift evaluations (ESEs) constructed using the milestones in resident assessment.

METHODS:
We developed 14 daily ESEs, each of which included 9 or 10 emergency medicine milestones. Postgraduate year (PGY)-1 and PGY-2 residents were assessed on milestone levels 1 through 3; PGY-3 and PGY-4 residents were assessed on levels 3 through 5. Each milestone was rated on a nominal scale (yes, no, or not applicable). The Clinical Competency Committee combined the ESE data with data from other assessments to determine each resident's proficiency level for the emergency medicine subcompetencies. We used descriptive statistics to summarize resident ESEs and milestone levels. We analyzed differences in ESE score across PGY levels using t tests and analyses of variance.

RESULTS:
Faculty completed 763 ESEs on 33 residents with a range of 2 to 54 (median=22) ESEs per resident. Faculty rarely (8%, 372 of 4633) rated a resident as not achieving a milestone on the ESEs. Analyses of variance revealed that ESE scores on level 3 milestones did not differ significantly by PGY level. There was poor agreement between ESE scores and Clinical Competency Committee ratings.

CONCLUSIONS:
The ESEs constructed using the milestones resulted in grade or milestone inflation. Our results do not support using milestones as a stand-alone assessment tool.
Clinical Assessment and Management Examination—Outpatient (CAMEO): Its Validity and Use in a Surgical Milestones Paradigm


OBJECTIVES:
Clinical Assessment and Management Examination—Outpatient (CAMEO) is a metric for evaluating the clinical performance of surgery residents. The aim of this study was to investigate the measurement characteristics of CAMEO and propose how it might be used as an evaluation tool within the general surgery milestones project.

DESIGN:
A total of 117 CAMEO evaluations were gathered and used for analysis. Internal consistency reliability was estimated, and item characteristics were explored. A Kruskal-Wallis procedure was performed to discern how well the instrument discriminated between training levels. An exploratory factor analysis was also conducted to understand the dimensionality of the evaluation.

SETTING:
CAMEO evaluations were collected from 2 departments of surgery geographically located in the Midwestern United States. Combined, the participating academic institutions graduate approximately 18 general surgery residents per year.

PARTICIPANTS:
In this retrospective data analysis, the number of evaluations per resident ranged from 1 to 7, and evaluations were collected from 2006 to 2013. For the purpose of data analysis, residents were classified as interns (postgraduate year 1 [PGY1]), juniors (PGY2-3), or seniors (PGY4-5).

RESULTS:
CAMEO scores were found to have high internal consistency (Cronbach's α = 0.96), and all items were highly correlated (≥ 0.86) to composite CAMEO scores. Scores discriminated between senior residents (PGY4-5) and lower level residents (PGY1-3). Per an exploratory factor analysis, CAMEO was revealed to measure a single dimension of "clinical competence."

CONCLUSIONS:
The findings of this research aligned with related literature and verified that CAMEO scores have desirable measurement properties, making CAMEO an attractive resource for evaluating the clinical performance of surgery residents.
A Faculty Development Program to Reduce Rater Error on Milestone-Based Assessments

Raj JM, Thorn, PM. Journal of Graduate Medical Education. December 2014. doi: http://dx.doi.org/10.4300/JGME-D-14-00161.1

BACKGROUND:
Rater errors, such as halo/reverse halo, range restriction, and leniency errors, are frequently cited as threats to the validity of resident assessment by faculty.

OBJECTIVE:
We studied whether participation in faculty development on the use of a new Milestone-based assessment tool reduced rater error for participants compared to individuals who did not participate.

METHODS:
We reviewed evaluations of resident Milestones completed by faculty at the end of rotations between July 2012 and June 2013. The 2 Milestones in each competency with the greatest number of ratings were selected for analysis.

RESULTS:
A total of 412 evaluations were analyzed, including 217 completed by faculty who participated in the development activity, and 240 completed by nonparticipant faculty. All evaluations that contained identical scores for all Milestones (16%) were completed by nonparticipant faculty ($\chi^2 = 37.498$, P, .001). Faculty who had participated in development assigned a wider range of scores and lower minimum scores to residents, and provided the highest ratings for residents less frequently (P, .001) than nonparticipants.

CONCLUSIONS:
Faculty who participated in education about the Milestones demonstrated significantly less halo, range restriction, and leniency errors than faculty members who did not participate. These findings support a recommendation to develop a cadre of “core faculty” by training them in the use of Milestone assessment tools, and making them responsible for a significant portion of resident assessments.
Immediate Surgical Skills Feedback in the Operating Room Using "SurF" Cards.


BACKGROUND:
Ensuring residents develop operative skills requires application of the principles of guided learning, deliberate practice, and directed feedback.

OBJECTIVE:
We sought to create and implement a tool to promote procedural "key" step review and immediate feedback on surgical skills, and examined faculty and resident satisfaction with surgical skills feedback.

METHODS:
We created surgical skills feedback (SurF) cards for 8 gynecologic procedures. Faculty/fellows and residents completed prestudy surveys querying frequency of preoperative key step review and satisfaction with surgical skill feedback, a SurF card each time 1 of 8 procedures was performed, and poststudy surveys to evaluate for changes.

RESULTS:
Prestudy surveys were completed by 31 faculty/fellows and 20 residents, with 55% (17 of 31) of the faculty/fellows and 5% (1 of 20) of the residents reporting key step review before surgery. All reported low satisfaction rates with feedback frequency, quality, and timeliness. After implementation of SurF cards, preoperative key step review occurred in 78% (82 of 105) of the procedures. Twenty-one faculty/fellows (68%) and 16 residents (80%) completed our poststudy survey. Faculty/fellows reported statistically similar key step review (n = 15 [71%], P = .23), while residents reported that key step review had significantly improved (n = 6 [38%], P = .01). Resident satisfaction with feedback frequency (5% to 50%, P = .002) and quality (15% to 50%, P = .02) increased significantly.

CONCLUSIONS:
The SurF cards we developed facilitated procedural key step review, were associated with significantly improved resident satisfaction with surgical feedback, and could prove helpful with outcomes assessments, such as Accreditation Council for Graduate Medical Education-required documentation of Milestone attainment.
Assessment of Resident Operative Performance Using a Real-Time Mobile Web System: Preparing for the Milestone Age


OBJECTIVE:
To satisfy trainees' operative competency requirements while improving feedback validity and timeliness using a mobile Web-based platform.

DESIGN:
The Southern Illinois University Operative Performance Rating Scale (OPRS) was embedded into a website formatted for mobile devices. From March 2013 to February 2014, faculty members were instructed to complete the OPRS form while providing verbal feedback to the operating resident at the conclusion of each procedure. Submitted data were compiled automatically within a secure Web-based spreadsheet. Conventional end-of-rotation performance (CERP) evaluations filed 2006 to 2013 and OPRS performance scores were compared by year of training using serial and independent-samples t tests. The mean CERP scores and OPRS overall resident operative performance scores were directly compared using a linear regression model. OPRS mobile site analytics were reviewed using a Web-based reporting program.

SETTING:
Large university-based general surgery residency program.

PARTICIPANTS:
General Surgery faculty used the mobile Web OPRS system to rate resident performance. Residents and the program director reviewed evaluations semiannually.

RESULTS:
Over the study period, 18 faculty members and 37 residents logged 176 operations using the mobile OPRS system. There were 334 total OPRS website visits. Median time to complete an evaluation was 45 minutes from the end of the operation, and faculty spent an average of 134 seconds on the site to enter 1 assessment. In the 38,506 CERP evaluations reviewed, mean performance scores showed a positive linear trend of 2% change per year of training (p = 0.001). OPRS overall resident operative performance scores showed a significant linear (p = 0.001), quadratic (p = 0.001), and cubic (p = 0.003) trend of change per year of clinical training, reflecting the resident operative experience in our training program. Differences between postgraduate year-1 and postgraduate year-5 overall performance scores were greater with the OPRS (mean = 0.96, CI: 0.55-1.38) than with CERP measures (mean = 0.37, CI: 0.34-0.41).

Additionally, there were consistent increases in each of the OPRS subcategories.

CONCLUSIONS:
In contrast to CERPs, the OPRS fully satisfies the Accreditation Council for Graduate Medical Education and American Board of Surgery operative assessment requirements. The mobile Web platform provides a convenient interface, broad accessibility, automatic data compilation, and compatibility with common database and statistical software. Our mobile OPRS system encourages candid feedback dialog and generates a comprehensive review of individual and group-wide operative proficiency in real time.
Entrustable Professional Activities: Making Sense of the Emergency Medicine Milestones


BACKGROUND:
The Next Accreditation System (NAS) is being implemented by the Accreditation Council for Graduate Medical Education with seven specialties, including Emergency Medicine (EM), which began in July 2013. The NAS represents a more structured method of accreditation, with dependence on outcomes and less emphasis on educational process. A key component of the NAS is the individual resident semiannual reporting of the Milestone proficiency levels for all sub-competencies, which are more specific areas of domain for the general competencies. All specialties are struggling to some extent with developing assessment mechanisms for the Milestones. At the heart of this struggle is the conceptualization of the Milestones themselves—descriptors of the individual. In practice, faculty assess clinical care provided to the patient by the resident. This creates difficulty for faculty to assign a resident to a specific sub-competency proficiency level when their focus has been on assessment of clinical care.

OBJECTIVES:
The objectives of this article include the discussion of whether Entrustable Professional Activities (EPAs) could be defined and linked to milestones in a way that, once implemented, could inform Clinical Competency Committees of the Milestone proficiency reporting.

DISCUSSION:
EPAs are units of professional work, or clinical care that may help translate aspects of clinical care into Milestone proficiencies. This article explores EPAs in depth, and discusses how EPAs may be used within EM as one method of assigning proficiency levels to residents.

CONCLUSIONS:
EPAs may be a useful tool to inform Milestone proficiency placement of residents. Because EPAs are based on clinical descriptions rather than individual physician descriptions, there may be less faculty development needed for Milestone sub-competency assessment.
Entrustment and Mapping of Observable Practice Activities for Resident Assessment


ABSTRACT
Entrustable Professional Activities (EPAs) and the Next Accreditation System reporting milestones reduce general competencies into smaller evaluable parts. However, some EPAs and reporting milestones may be too broad to use as direct assessment tools. We describe our internal medicine residency curriculum and assessment system, which uses entrustment and mapping of observable practice activities (OPAs) for resident assessment. We created discrete OPAs for each resident rotation and learning experience. In combination, these serve as curricular foundation and tools for assessment. OPA performance is measured via a 5-point entrustment scale, and mapped to milestones and EPAs. Entrustment ratings of OPAs provide an opportunity for immediate structured feedback of specific clinical skills, and mapping OPAs to milestones and EPAs can be used for longitudinal assessment, promotion decisions, and reporting. Direct assessment and demonstration of progressive entrustment of trainee skill over time are important goals for all training programs. Systems that use OPAs mapped to milestones and EPAs provide the opportunity for achieving both, but require validation.
Assessing Competency in Practice-Based Learning: A Foundation for Milestones in Learning Portfolio Entries


BACKGROUND:
Graduate medical education is undergoing a dramatic shift toward competency-based assessment of learners. Competency assessment requires clear definitions of competency and validated assessment methods. The purpose of this study is to identify criteria used by surgical educators to judge competence in Practice-Based Learning and Improvement (PBL&I) as demonstrated in learning portfolios.

METHODS:
A total of 6 surgical learning and instructional portfolio entries served as documents to be assessed by 3 senior surgical educators. These faculty members were asked to rate and then identify criteria used to assess PBL&I competency. Individual interviews and group discussions were conducted, recorded, and transcribed to serve as the study dataset. Analysis was performed using qualitative methodology to identify themes for the purpose of defining competence in PBL&I. The assessment themes derived are presented with narrative examples to describe the progression of competency.

RESULTS:
The collaborative coding process resulted in identification of 7 themes associated with competency in PBL&I related to surgical learning and instructional portfolio entries: (1) self-awareness regarding effect of actions; (2) identification and thorough description of learning goals; (3) cases used as catalyst for reflection; (4) reconceptualization with appropriate use and critique of cited literature; (5) communication skills/completeness of entry template; (6) description of future behavioral change; and (7) engagement in process--identifies as personally relevant.

CONCLUSIONS:
The identified themes are consistent with and complement other criteria emerging from reflective practice literature and experiential learning theory. This study provides a foundation for further development of a tool for assessing learner portfolios consistent with the Accreditation Council for Graduate Medical Education's Next Accreditation System requirements.
Programmatic Assessment of Level 1 Milestones in Incoming Interns


OBJECTIVES:
With the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System, emergency medicine (EM) residency programs will be required to report residents' progress through the EM milestones. The milestones include five progressively advancing skill levels, with Level 1 defining the skill set of a medical school graduate and Level 5, that of an attending physician. The ACGME stresses that multiple forms of assessment should be used to ensure capture of the multifaceted competencies. The objective of this study was to determine the feasibility and results of programmatic assessment of Level 1 milestones using multisource assessments for incoming EM interns in July.

METHODS:
The study population was interns starting in 2012 and 2013. Interns' Level 1 milestone assessment was done with four distinct methods: 1) the postgraduate orientation assessment (POA) by the Graduate Medical Education Office for all incoming interns (this multistation examination covers nine of the EM milestones and includes standardized patient cases, task completion, and computer-based stations); 2) direct observation of patient encounters by core faculty using a milestones-based clinical skills competency checklist; 3) the global monthly assessment at the end of the intern orientation month that was updated to reflect the EM milestones; and 4) faculty assessment during procedural labs. These occurred during the July orientation month that included the POA, clinical shifts, didactic sessions, and procedure labs.

RESULTS:
In the POA, interns were competent in 48% to 93% of the milestones assessed. Overall, competency was 70% to 80%, with low scores noted in aseptic technique (patient care Milestone 13 [PC13]) and written and verbal hand-off (interpersonal communications skills [ICS]2). In overall communication, 70% of interns demonstrated competency. In excess of 80% demonstrated competency in critical values interpretation (PC3), informed consent (PC9), pain assessment (PC11), and geriatric functional assessment (PC3). On direct observation, almost all Level 1 milestones were achieved (93% to 100%); however, only 78% of interns achieved competency in pharmacotherapy (PC5). On global monthly evaluations, all interns met Level 1 milestones.

CONCLUSIONS:
A multisource assessment of EM milestones is feasible and useful to determine Level 1 milestones achievement for incoming interns. A structured assessment program, used in conjunction with more traditional forms of evaluation such as global monthly evaluations and direct observation, is useful for identifying deficits in new trainees and may be able inform the creation of early intervention programs.
Complexity in Graduate Medical Education: A Collaborative Education Agenda for Internal Medicine and Geriatric Medicine


ABSTRACT
Internal medicine residents today face significant challenges in caring for an increasingly complex patient population within ever-changing education and health care environments. As a result, medical educators, health care system leaders, payers, and patients are demanding change and accountability in graduate medical education (GME). A 2012 Society of General Internal Medicine (SGIM) retreat identified medical education as an area for collaboration between internal medicine and geriatric medicine. The authors first determined a short-term research agenda for resident education by mapping selected internal medicine reporting milestones to geriatrics competencies, and listing available sample learner assessment tools. Next, the authors proposed a strategy for long-term collaboration in three priority areas in clinical medicine that are challenging for residents today: (1) team-based care, (2) transitions and readmissions, and (3) multi-morbidity. The short-term agenda focuses on learner assessment, while the long-term agenda allows for program evaluation and improvement. This model of collaboration in medical education combines the resources and expertise of internal medicine and geriatric medicine educators with the goal of increasing innovation and improving outcomes in GME targeting the needs of our residents and their patients.
Putting the Pediatrics Milestones into Practice: A Consensus Roadmap and Resource Analysis


ABSTRACT

The Accreditation Council for Graduate Medical Education has partnered with member boards of the American Board of Medical Specialties to initiate the next steps in advancing competency-based assessment in residency programs. This initiative, known as the Milestone Project, is a paradigm shift from traditional assessment efforts and requires all pediatrics residency programs to report individual resident progression along a series of 4 to 5 developmental levels of performance, or milestones, for individual competencies every 6 months beginning in June 2014. The effort required to successfully make this shift is tremendous given the number of training programs, training institutions, and trainees. However, it holds great promise for achieving training outcomes that align with patient needs; developing a valid, reliable, and meaningful way to track residents' development; and providing trainees with a roadmap for learning. Recognizing the resources needed to implement this new system, the authors, all residency program leaders, provide their consensus view of the components necessary for implementing and sustaining this effort, including resource estimates for completing this work. The authors have identified 4 domains: (1) Program Review and Development of Stakeholders and Participants, (2) Assessment Methods and Validation, (3) Data and Assessment System Development, and (4) Summative Assessment and Feedback.

This work can serve as a starting point and framework for collaboration with program, department, and institutional leaders to identify and garner necessary resources and plan for local and national efforts that will ensure successful transition to milestones-based assessment.
Integrating the NAS Milestones and Handheld Technology to Improve Residency Training and Assessment


OBJECTIVE:
To incorporate the use of an intuitive and robust assessment tool in conjunction with the Next Accreditation System Milestones to maximize opportunities for trainee performance feedback and continuous trainee assessment, with the long-term goal of increasing the rate of performance improvement and mastery of knowledge and surgical skills.

DESIGN:
Pilot study.

SETTING:
Johns Hopkins Medicine, Baltimore, MD. Primary, tertiary, and quaternary clinical care; institutional environment.

PARTICIPANTS:
Experimental group: two randomly selected postgraduate year-1 integrated training program residents per year for 2 consecutive years from the Department of Plastic and Reconstructive Surgery.

CONTROL GROUP:
Traditionally trained residents from the integrated training program in the Department of Plastic and Reconstructive Surgery. Study duration: 7 years (until residents complete residency training).

ANTICIPATED RESULTS:
This assessment strategy would create large amounts of informative data on trainees, which can be cross-referenced to determine trainee progress. Assessment data would be collected continuously from all faculty surgeons. Comparisons of faculty and resident self-assessments would facilitate resident evaluations. Ease of use of the data collection structure would improve faculty evaluation compliance and timely resident case report completion.

CONCLUSIONS:
Improving the efficiency and efficacy of competency documentation is critical. Using portable technologies is an intuitive way to improve the trainee assessment process. We anticipate that this 2-pronged approach to trainee assessments would quickly provide large amounts of informative data to better assess trainee progress and inform Milestone assessments in a manner that facilitates immediate feedback. Assessments of faculty and resident satisfaction would help us further refine the assessment process as needed. If successful, this format could easily be implemented by other training programs.
Initial Performance of a Modified Milestones Global Evaluation Tool for Semiannual Evaluation of Residents by Faculty


OBJECTIVES:
To determine whether faculty could successfully evaluate residents using a competency-based modified Milestones global evaluation tool.

DESIGN:
A program's leadership team modified a draft Surgery Milestones Working Group summative global assessment instrument into a modified Milestones tool (MMT) for local use during faculty meetings devoted to semiannual resident review. Residents were scored on 15 items spanning all competencies using an 8-point graphic response scale; unstructured comments also were solicited. Arithmetic means were computed at the resident and postgraduate year cohort levels for items and competency item sets. Score ranges (highest minus lowest score) were calculated; variability was termed "low" (range <2.0 points), "moderate" (range = 2.0), or "high" (range >2.0). A subset of "low" was designated "small" (1.0-1.9). Trends were sought among item, competency, and total Milestones scores. MMT correlations with examination scores and multisource (360°) assessments were explored. The success of implementing MMT was judged using published criteria for educational assessment methods.

SETTING:
Fully accredited, independently sponsored residency.

PARTICIPANTS:
Program leaders and 22 faculty members (71% voluntary, mean 12y of experience).

RESULTS:
Twenty-six residents were assessed, yielding 7 to 13 evaluations for MMT per categorical resident and 3 to 6 per preliminary trainee. Scores spanned the entire response scale. All MMT evaluations included narrative comments. Individual resident score variability was low (96% within competencies and 92% across competencies). Subset analysis showed that small variations were common (35% within competencies and 54% across competencies). Postgraduate year cohort variability was higher (61% moderate or high within competencies and 50% across competencies). Cohort scores at the item, competency, and total score levels exhibited rising trajectories, suggesting MMT construct validity. MMT scores did not demonstrate concurrent validity, correlating poorly with other metrics. The MMT met multiple criteria for good assessment.

CONCLUSIONS:
A modified Milestones global evaluation tool can be successfully adopted for semiannual assessments of resident performance by volunteer faculty members.
Early Feedback on the Use of the Internal Medicine Reporting Milestones in Assessment of Resident Performance


BACKGROUND:
The educational milestones were designed as a criterion-based framework for assessing resident progression on the 6 Accreditation Council for Graduate Medical Education competencies.

OBJECTIVE:
We obtained feedback on, and assessed the construct validity and perceived feasibility and utility of, draft Internal Medicine Milestones for Patient Care and Systems-Based Practice.

METHODS:
All participants in our mixed-methods study were members of competency committees in internal medicine residency programs. An initial survey assessed participant and program demographics; focus groups obtained feedback on the draft milestones and explored their perceived utility in resident assessment, and an exit survey elicited input on the value of the draft milestones in resident assessment. Surveys were tabulated using descriptive statistics. Conventional content analysis method was used to assess the focus group data.

RESULTS:
Thirty-four participants from 17 programs completed surveys and participated in 1 of 6 focus groups. Overall, the milestones were perceived as useful in formative and summative assessment of residents. Participants raised concerns about the length and complexity of some draft milestones and suggested specific changes. The focus groups also identified a need for faculty development. In the exit survey, most participants agreed that the Patient Care and Systems-Based Practice Milestones would help competency committees assess trainee progress toward independent practice.

CONCLUSIONS:
Draft reporting milestones for 2 competencies demonstrated significant construct validity in both the content and response process and the perceived utility for the assessment of resident performance. To ensure success, additional feedback from the internal medicine community and faculty development will be necessary.
Playing with Curricular Milestones in the Educational Sandbox: Q-Sort Results from an Internal Medicine Educational Collaborative


PURPOSE:
In competency-based medical education, the focus of assessment is on learner demonstration of predefined outcomes or competencies. One strategy being used in internal medicine (IM) is applying curricular milestones to assessment and reporting milestones to competence determination. The authors report a practical method for identifying sets of curricular milestones for assessment of a landmark, or a point where a resident can be entrusted with increased responsibility.

METHOD:
Thirteen IM residency programs joined in an educational collaborative to apply curricular milestones to training. The authors developed a game using Q-sort methodology to identify high-priority milestones for the landmark "Ready for indirect supervision in essential ambulatory care" (EsAMB). During May to December 2010, the programs'ambulatory faculty participated in the Q-sort game to prioritize 22 milestones for EsAMB. The authors analyzed the data to identify the top 8 milestones.

RESULTS:
In total, 149 faculty units (1-4 faculty each) participated. There was strong agreement on the top eight milestones; six had more than 92% agreement across programs, and five had 75% agreement across all faculty units. During the Q-sort game, faculty engaged in dynamic discussion about milestones and expressed interest in applying the game to other milestones and educational settings.

CONCLUSIONS:
The Q-sort game enabled diverse programs to prioritize curricular milestones with interprogram and interparticipant consistency. A Q-sort exercise is an engaging and playful way to address milestones in medical education and may provide a practical first step toward using milestones in the real-world educational setting.
CCCs – Structure and Function
Assessment Methods and Resource Requirements for Milestone Reporting by an Emergency Medicine Clinical Competency Committee


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) introduced milestones for Emergency Medicine (EM) in 2012. Clinical Competency Committees (CCC) are tasked with assessing residents on milestones and reporting them to the ACGME. Appropriate workflows for CCCs are not well defined.

OBJECTIVE:
Our objective was to compare different approaches to milestone assessment by a CCC, quantify resource requirements for each and to identify the most efficient workflow.

DESIGN:
Three distinct processes for rendering milestone assessments were compared: Full milestone assessments (FMA) utilizing all available resident assessment data, Ad-hoc milestone assessments (AMA) created by multiple expert educators using their personal assessment of resident performance, Self-assessments (SMA) completed by residents. FMA were selected as the theoretical gold standard. Intraclass correlation coefficients were used to analyze for agreement between different assessment methods. Kendall's coefficient was used to assess the inter-rater agreement for the AMA.

RESULTS:
All 13 second-year residents and 7 educational faculty of an urban EM Residency Program participated in the study in 2013. Substantial or better agreement between FMA and AMA was seen for 8 of the 23 total subcompetencies (PC4, PC8, PC9, PC11, MK, PROF2, ICS2, SBP2), and for 1 subcompetency (SBP1) between FMA and SMA. Multiple AMA for individual residents demonstrated substantial or better interobserver agreement in 3 subcompetencies (PC1, PC2, and PROF2). FMA took longer to complete compared to AMA (80.9 vs. 5.3 min, p < 0.001).

CONCLUSIONS:
Using AMA to evaluate residents on the milestones takes significantly less time than FMA. However, AMA and SMA agree with FMA on only 8 and 1 subcompetencies, respectively. An estimated 23.5 h of faculty time are required each month to fulfill the requirement for semiannual reporting for a residency with 42 trainees.
Program Director Perceptions of Usefulness of the Accreditation Council for Graduate Medical Education Milestones System for Urology Resident Evaluation


OBJECTIVES:
To assess the application and perceived usefulness of the Accreditation Council for Graduation Medical Education (ACGME) Milestones system for resident evaluation among urology program directors (PDs).

METHOD:
We conducted an online survey of 133 urology PDs. The survey addressed several domains: (1) demographic information, (2) logistics and implementation of the faculty Clinical Competency Committee (CCC) meetings, and (3) perceived overall effectiveness and usefulness of the Milestones assessments.

RESULTS:
Eighty-eight responses were obtained (66% response rate). A total of 42/88 programs (48%) described the Milestones as very or somewhat unhelpful in resident evaluation, with a comparable proportion (44%) responding Milestones assessments never or almost never accurately distinguished between residents. Respondents felt higher scores on all domains of the Milestones were completely or somewhat uncorrelated to higher inservice exam scores (58%), with a smaller fraction (49%) deeming they were not predictive of board passage rates. Overall, 30% of respondents answered neutrally as to whether they felt the Milestones format has led to better resident formative feedback, and 35% were neutral as to the implications of this system toward promoting professional development.

CONCLUSIONS:
The ACGME Milestones system for resident evaluation was initiated to create a uniform competency-based assessment system; however, a sizable proportion of urology PDs in our cohort did not find the Milestones system helpful or accurate in assessing residents or predicting future successes. Given the Milestones system is still in its infancy, the utility of this system within urology has yet to be fully assessed.
Influence of Clinical Competency Committee Review Process on Summative Resident Assessment Decisions


BACKGROUND:
Clinical Competency Committees (CCCs) are charged with making summative assessment decisions about residents.

OBJECTIVE:
We explored how review processes CCC members utilize influence their decisions regarding residents' milestone levels and supervisory roles.

METHODS:
We conducted a multisite longitudinal prospective observational cohort study at 14 pediatrics residency programs during academic year 2015-2016. Individual CCC members biannually reported characteristics of their review process and Accreditation Council for Graduate Medical Education milestone levels and recommended supervisory role categorizations assigned to residents. Relationships among characteristics of CCC member reviews, mean milestone levels, and supervisory role categorizations were analyzed using mixed-effects linear regression, reported as mean differences with 95% confidence intervals (CIs), and Bayesian mixed-effects ordinal regression, reported as odds ratios (ORs) and 95% credible intervals (CrIs).

RESULTS:
A total of 155 CCC members participated. Members who provided milestones or other professional development feedback after CCC meetings assigned significantly lower mean milestone levels (mean 1.4 points; CI -2.2 to -0.6; \( P < .001 \)) and were significantly less likely to recommend supervisory responsibility in any setting (OR = 0.23, CrI 0.05-0.83) compared with CCC members who did not. Members recommended less supervisory responsibility when they reviewed more residents (OR = 0.96, 95% CrI 0.94-0.99) and participated in more review cycles (OR = 0.22, 95% CrI 0.07-0.63).

CONCLUSIONS:
This study explored the association between characteristics of individual CCC member reviews and their summative assessment decisions about residents. Further study is needed to gain deeper understanding of factors influencing CCC members' summative assessment decisions.
OBJECTIVE: This single-specialty, multi-institutional study aimed to determine 1) the association between milestone ratings for individual competencies and average milestone ratings (AMRs) and 2) the association between AMRs and recommended supervisory role categorizations made by individual clinical competency committee (CCC) members.

METHODS: During the 2015-16 academic year, CCC members at 14 pediatric residencies reported milestone ratings for 21 competencies and recommended supervisory role categories (may not supervise, may supervise in some settings, may supervise in all settings) for residents they reviewed. An exploratory factor analysis of competencies was conducted. The associations among individual competencies, the AMR, and supervisory role categorizations were determined by computing bivariate correlations. The relationship between AMRs and recommended supervisory role categorizations was examined using an ordinal mixed logistic regression model.

RESULTS: Of the 155 CCC members, 68 completed both milestone assignments and supervision categorizations for 451 residents. Factor analysis of individual competencies controlling for clustering of residents in raters and sites resulted in a single-factor solution (cumulative variance: 0.75). All individual competencies had large positive correlations with the AMR (correlation coefficient: 0.84-0.93), except for two professionalism competencies (Prof1: 0.63 and Prof4: 0.65). When combined across training year and time points, the AMR and supervisory role categorization had a moderately positive correlation (0.56).

CONCLUSIONS: This exploratory study identified a modest correlation between average milestone ratings and supervisory role categorization. Convergence of competencies on a single factor deserves further exploration, with possible rater effects warranting attention.
A Multicenter Prospective Comparison of the Accreditation Council for Graduate Medical Education Milestones: Clinical Competency Committee vs. Resident Self-Assessment


OBJECTIVE:
The Accreditation Council for Graduate Medical Education requires accredited residency programs to implement competency-based assessments of medical trainees based upon nationally established Milestones. Clinical competency committees (CCC) are required to prepare biannual reports using the Milestones and ensure reporting to the Accreditation Council for Graduate Medical Education. Previous research demonstrated a strong correlation between CCC and resident scores on the Milestones at 1 institution. We sought to evaluate a national sampling of general surgery residency programs and hypothesized that CCC and resident assessments are similar.

DESIGN:
Details regarding the makeup and process of each CCC were obtained. Major disparities were defined as an absolute mean difference of $\geq 0.5$ on the 4-point scale. A negative assessment disparity indicated that the residents evaluated themselves at a lower level than did the CCC. Statistical analysis included Wilcoxon rank sum and Sign tests.

SETTING:
CCCs and categorical general surgery residents from 15 residency programs completed the Milestones document independently during the spring of 2016.

RESULTS:
Overall, 334 residents were included; 44 (13%) and 43 (13%) residents scored themselves $\geq 0.5$ points higher and lower than the CCC, respectively. Female residents scored themselves a mean of 0.08 points lower, and male residents scored themselves a mean of 0.03 points higher than the CCC. Median assessment differences for postgraduate year (PGY) 1-5 were 0.03 (range: -0.94 to 1.28), -0.11 (range: -1.22 to 1.22), -0.08 (range: -1.28 to 0.81), 0.02 (range: -0.91 to 1.00), and -0.19 (range: -1.16 to 0.50), respectively. Residents in university vs. independent programs had higher rates of negative assessment differences in medical knowledge (15% vs. 6%; $P = 0.015$), patient care (17% vs. 5%; $P = 0.002$), professionalism (23% vs. 14%; $P = 0.013$), and system-based practice (18% vs. 9%; $P = 0.031$) competencies. Major assessment disparities by sex or PGY were similar among individual competencies.

CONCLUSIONS:
Surgery residents in this national cohort demonstrated self-awareness when compared to assessments by their respective CCCs. This was independent of program type, sex, or level of training. PGY 5 residents, female residents, and those from university programs consistently rated themselves lower than the CCC, but these were not major disparities and the significance of this is unclear.
How Do Clinical Competency Committees Use Different Sources of Data to Assess Residents' Performance on the Internal Medicine Milestones? A Mixed Methods Pilot Study


PURPOSE:
This study examines how Clinical Competency Committees (CCCs) synthesize assessment data to make judgments about residents' clinical performances.

METHODS:
Between 2014 and 2015, after four six-month reporting periods to the Accreditation Council for Graduate Medical Education (ACGME), 7 of 16 CCC faculty at Rush University Medical Center completed questionnaires focused on their perspectives about rating residents on their achievement of the milestones and participated in a focus group. Qualitative data were analyzed using grounded theory. Milestones ratings for two six-month ACGME reporting cycles (n = 100 categorical residents) were also analyzed.

RESULTS:
CCC members weighted resident rotation ratings highest (weight = 37%), followed by faculty rotation comments (weight = 27%) and personal experience with residents (weight = 14%) for making judgments about learner's milestone levels. Three assessment issues were identified from qualitative analyses: (1) "design issues" (e.g. problems with available data or lack thereof); (2) "synthesis issues" (e.g. factors influencing ratings and decision-making processes) and (3) "impact issues" (e.g. how CCC generated milestones ratings are used).

CONCLUSIONS:
Identifying factors that affect assessment at all stages of the CCC process can contribute to improving assessment systems, including support for faculty development for CCCs. Recognizing challenges in synthesizing first and second-hand assessment data is an important step in understanding the CCC decision-making process.
The Effect and Use of Milestones in the Assessment of Neurological Surgery Residents and Residency Programs


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.
Ensuring Resident Competence: A Narrative Review of the Literature on Group Decision Making to Inform the Work of Clinical Competency Committees


INTRODUCTION:
The expectation for graduate medical education programs to ensure that trainees are progressing toward competence for unsupervised practice prompted requirements for a committee to make decisions regarding residents' progress, termed a clinical competency committee (CCC). The literature on the composition of these committees and how they share information and render decisions can inform the work of CCCs by highlighting vulnerabilities and best practices. Objective We conducted a narrative review of the literature on group decision making that can help characterize the work of CCCs, including how they are populated and how they use information.

METHODS:
English language studies of group decision making in medical education, psychology, and organizational behavior were used.

RESULTS:
The results highlighted 2 major themes. Group member composition showcased the value placed on the complementarity of members' experience and lessons they had learned about performance review through their teaching and committee work. Group processes revealed strengths and limitations in groups' understanding of their work, leader role, and information-sharing procedures. Time pressure was a threat to the quality of group work.

CONCLUSIONS:
Implications of the findings include the risks for committees that arise with homogeneous membership, limitations to available resident performance information, and processes that arise through experience rather than deriving from a well-articulated purpose of their work. Recommendations are presented to maximize the effectiveness of CCC processes, including their membership and access to, and interpretation of, information to yield evidence-based, well-reasoned judgments.
How Do Emergency Medicine Residency Programs Structure their Clinical Competency Committees? A Survey


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) recently has mandated the formation of a clinical competency committee (CCC) to evaluate residents across the newly defined milestone continuum. The ACGME has been nonproscriptive of how these CCCs are to be structured in order to provide flexibility to the programs.

OBJECTIVES:
No best practices for the formation of CCCs currently exist. We seek to determine common structures of CCCs recently formed in the Council of Emergency Medicine Residency Directors (CORD) member programs and identify unique structures that have been developed.

METHODS:
In this descriptive study, an 18-question survey was distributed via the CORD listserv in the late fall of 2013. Each member program was asked questions about the structure of its CCC. These responses were analyzed with simple descriptive statistics.

RESULTS:
A total of 116 of the 160 programs responded, giving a 73% response rate. Of responders, most (71.6%) CCCs are chaired by the associate or assistant program director, while a small number (14.7%) are chaired by a core faculty member. Program directors (PDs) chair 12.1% of CCCs. Most CCCs are attended by the PD (85.3%) and selected core faculty members (78.5%), leaving the remaining committees attended by any core faculty. Voting members of the CCC consist of the residency leadership either with the PD (53.9%) or without the PD (36.5%) as a voting member. CCCs have an average attendance of 7.4 members with a range of three to 15 members. Of respondents, 53.1% of CCCs meet quarterly while 37% meet monthly. The majority of programs (76.4%) report a system to match residents with a faculty mentor or advisor. Of respondents, 36% include the resident's faculty mentor or advisor to discuss a particular resident. Milestone summaries (determination of level for each milestone) are the primary focus of discussion (93.8%), utilizing multiple sources of information.

CONCLUSIONS:
The substantial variability and diversity found in our CORD survey of CCC structure and function suggest that there are myriad strategies that residency programs can use to match individual program needs and resources to requirements of the ACGME. Identifying a single protocol for CCC structure and development may prove challenging.
Pathology Milestones: Assessing Clinical Competency by Committee


ABSTRACT
All Accreditation Council for Graduate Medical Education accredited pathology residency training programs are now required to evaluate residents using the new Pathology Milestones assessment tool. Similar to implementation of the 6 Accreditation Council for Graduate Medical Education competencies a decade ago, there have been challenges in implementation of the new milestones for many residency programs. The pathology department at the University of Iowa has implemented a process that divides the labor of the task in rating residents while also maintaining consistency in the process. The process is described in detail, and some initial trends in milestone evaluation are described and discussed. Our experience indicates that thoughtful implementation of the Pathology Milestones can provide programs with valuable information that can inform curricular changes.
Navigating the Next Accreditation System: A Dashboard for the Milestones


INTRODUCTION:
In July 2014, all residency programs accredited by the Accreditation Council for Graduate Medical Education (ACGME) were enrolled in a new system called the Next Accreditation System. Residency programs may not be clear on how best to comply with these new accreditation requirements. Large amounts of data must be collected, evaluated, and submitted twice a year to the council's Web-based data collection system. One challenge is that the new "end-of-rotation" evaluations must reflect specialty-specific milestones, on which many faculty members are not well versed. Like other residency programs, we tried to address the challenges using our local resources.

METHODS:
We used our existing electronic goals and objectives for each rotation coupled with appropriate end-of-rotation evaluations reflecting the specialty-specific milestones through a process of editing and mapping.

RESULTS:
Data extracted from these evaluations were added to an interactive dashboard that also contained evaluations on additional program-specific modifiers of residents' performance. A resident's final overall performance was visually represented on a plot graph. The novel dashboard included features to save evaluations for future comparisons and to track residents' progress during their entire training. It proved simple to use and was able to reduce the time needed for each resident evaluation to 5 to 10 minutes.

CONCLUSION:
This tool has made it much easier and less challenging for the members of our Clinical Competency Committee to start deliberation about each resident's performance.

PURPOSE:
Clinical competency committees (Milestones: a rapid assessment method for the Clinical Competency Committee.

INTRODUCTION:
Educational milestones are now used to assess the developmental progress of all U.S. graduate medical residents during training. Twice annually, each program's Clinical Competency Committee (CCC) makes these determinations and reports its findings to the Accreditation Council for Graduate Medical Education (ACGME). The ideal way to conduct the CCC is not known. After finding that deliberations reliant upon the new milestones were time intensive, our internal medicine residency program tested an approach designed to produce rapid but accurate assessments.

MATERIAL AND METHODS:
For this study, we modified our usual CCC process to include pre-meeting faculty ratings of resident milestones progress with in-meeting reconciliation of their ratings. Data were considered largely via standard report and presented in a pre-arranged pattern. Participants were surveyed regarding their perceptions of data management strategies and use of milestones. Reliability of competence assessments was estimated by comparing pre-/post- intervention class rank lists produced by individual committee members with a master class rank list produced by the collective CCC after full deliberation.

RESULTS:
Use of the study CCC approach reduced committee deliberation time from 25 min to 9 min per resident (p<0.001). Committee members believed milestones improved their ability to identify and assess expected elements of competency development (p = 0.026). Individual committee member assessments of trainee progress agreed well with collective CCC assessments.

CONCLUSIONS:
Modification of the clinical competency process to include pre-meeting competence ratings with in-meeting reconciliation of these ratings led to shorter deliberation times, improved evaluator satisfaction and resulted in reliable milestone assessments.
Implementation of Nephrology Subspecialty Curricular Milestones


ABSTRACT
Beginning in the 2014-2015 training year, the US Accreditation Council for Graduate Medical Education (ACGME) required that nephrology Clinical Competency Committees assess fellows' progress toward 23 subcompetency "context nonspecific" internal medicine subspecialty milestones. Fellows' advancement toward the "ready for unsupervised practice" target milestone now is tracked in each of the 6 competencies: Patient Care, Medical Knowledge, Professionalism, Interpersonal Communication Skills, Practice-Based Learning and Improvement, and Systems-Based Practice. Nephrology program directors and subspecialty societies must define nephrology-specific "curricular milestones," mapped to the nonspecific ACGME milestones. Although the ACGME goal is to produce data that can discriminate between successful and underperforming training programs, the approach is at risk to produce biased, inaccurate, and unhelpful information. We map the ACGME internal medicine subspecialty milestones to our previously published nephrology-specific milestone schema and describe entrustable professional activities and other objective assessment tools that inform milestone decisions. Mapping our schema onto the ACGME subspecialty milestone reporting form allows comparison with the ACGME subspecialty milestones and the curricular milestones developed by the American Society of Nephrology Program Directors. Clinical Competency Committees may easily adapt and directly translate milestone decisions reached using our schema onto the ACGME internal medicine subspecialty competency milestone-reporting format.
A Systematic Approach toward Building a Fully Operational Clinical Competency Committee


BACKGROUND:
The Accreditation Council for Graduate Medical Education has offered minimal guidelines for the creation and implementation of clinical competency committees (CCCs). As surgical residency programs may differ greatly in terms of size and structure, requirements that are too specific throughout the process could place some programs at a great disadvantage.

OBJECTIVE:
The purpose of this article is to address some of the common considerations all surgery residency programs will face. The creation of standard operating procedures for the CCCs will allow each committee to develop internal consistency, improve productivity, maintain efficiency and quality control, facilitate training of new committee members, and cross-train other faculty and residents on the key processes to provide transparency.

METHODS:
This article offers recommendations on the 3 key areas of CCC implementation: the prereview, resident milestone review, and the postreview processes. Specific components related to shifting culture, committee membership and terms, assessing available evidence, and review dissemination are outlined, and example scenarios are provided throughout the article.

CONCLUSION:
With the implementation of CCCs and the milestones project, residency programs have an opportunity to improve the overall quality of decision making regarding residents' promotion to the next training level or independent practice. CCCs will undoubtedly be confronted with numerous challenges, as they implement the milestones project and are faced with the need to make multiple changes. Therefore, implementing milestones should be viewed as a goal to be accomplished over the long term.
How to Use Milestones Data to Improve your Program
Moving toward Milestone-Based Assessment in Osteopathic Manipulative Medicine


ABSTRACT

Osteopathic medicine is continuing to move toward competency-based education at undergraduate and graduate levels. Competencies and Entrustable Professional Activities (EPAs) have been implemented to guide educators on the skills and abilities that osteopathic medical students and residents should be able to perform as physicians. Unfortunately, many of these skills have not been well described, and the threshold of "competence" or "entrustability" for each of these tasks remains elusive. The author presents an approach to measuring competence in the domain of osteopathic manipulative medicine using a milestone rubric to assess skills related to osteopathic screening, diagnosis, technique, and explanation. This rubric can be applied to all levels of osteopathic training and across many diagnostic and treatment modalities. Clearly defining and assessing the individual skills composing competence in osteopathic manipulative medicine will be increasingly important as medical education continues to evolve and modernize.
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.
Analysis of Milestone-Based End-of-Rotation Evaluations for Ten Residents Completing a Three-Year Anesthesiology Residency


INTRODUCTION:
Faculty are required to assess the development of residents using educational milestones. This descriptive study examined the end-of-rotation milestone-based evaluations of anesthesiology residents by rotation faculty directors. The goals were to measure: (1) how many of the 25 Accreditation Council for Graduate Medical Education (ACGME) anesthesiology subcompetency milestones were included in each of the residency's rotations evaluations, (2) the percentage of evaluations sent to the rotation director that were actually completed by the director, (3) the length of time between the end of the residents' rotations and completion of the evaluations, (4) the frequency of straight line scoring, defined as the resident receiving the same milestone level score for all subcompetencies on the evaluation, and (5) how often a resident received a score below a Level 4 in at least one subcompetency in the three months prior to graduating.

METHODS:
In 2013, the directors for each the 24 anesthesia rotations in the Stanford University School of Medicine Anesthesiology Residency Program created new milestone-based evaluations to be used at the end of rotations to evaluate residents. The directors selected the subcompetencies from the list released by the ACGME that were most appropriate for their rotation. End-of-rotation evaluations for the post graduate year (PGY)-2 to PGY-4 from July 1, 2014 to June 30, 2017 were retrospectively analyzed for a sample of 10 residents randomly selected from 22 residents in the graduating class.

RESULTS:
The mean number of subcompetencies evaluated by each of the 24 rotations in the residency equaled 17.88 (standard deviation (SD): 3.39, range 10-24, median 18.5) from the available possible total of 25 subcompetencies. Three subcompetencies (medical knowledge, communication with patients and families, and coordination of patient care within the healthcare system) were included in the evaluation instruments of all 24 rotations. The three least frequently listed subcompetencies were: "acute, chronic, and cancer-related pain consultation/management" (25% of rotations had this on the end-of-rotation evaluation), "triage and management of critically ill patient in non-operative setting" (33%), and "education of patient, families, students, residents, and others" (38%). Overall, 418 end of rotation evaluations were issued and 341 (82%) completed, with 63% completed within one month, 22% between month one and two, and 15% after two months. The frequency of straight line scoring varied, from never occurring (0%) in three rotations to always occurring (100%) in two rotations, with an overall average of 51% (SD: 33%). Sixty-one percent of straight line scoring corresponded to the residents' postgraduate year whereby, for example, a post-graduate year two resident received an ACGME Level 2 proficiency for all subcompetencies. Thirty-one percent of the straight line scoring was higher than the resident's year of training (e.g., a PGY-2 received Level 3 or higher for all the subcompetencies). The remaining 7% of straight line scoring was below the expected level for the year of training. Three of seven residents had at least one subcompetency rated as below a Level 4 on one of the evaluations during the three months prior to finishing residency.

CONCLUSION:
Formal analysis of a residency program's end-of-rotation milestone evaluations may uncover opportunities to improve competency-based evaluations.
Competencies, Milestones, and a Level of Supervision Scale for Entrustable Professional Activities for Scholarship


PROBLEM:
Scholarship is an important element of both undergraduate and graduate medical education, and scholarly activity is required for all pediatric fellows. However, despite the creation of entrustable professional activities (EPA) for scholarship, the specific progressive levels of performance and the appropriate level of supervision for a given performance level have not been defined. The authors developed competencies and milestones for the scholarship EPA to provide a framework for assessment across the continuum; a level of supervision scale was also developed.

APPROACH:
The Vitae Researcher Development Framework served as a template to create the competencies and milestones for the scholarship EPA. Beginning in September 2015 and using a modified Delphi approach, three distinct drafts were circulated to individuals with expertise in various types of scholarship until broad agreement was achieved. Then, in October 2016, the Steering Committee of the Subspecialty Pediatrics Investigator Network created a level of supervision scale, modeled after one it had previously developed.

OUTCOMES:
Eight competencies were identified as important in making entrustment decisions related to scholarship. For each competency, five milestone levels that span the continuum from novice to expert were created. A supervision scale with five progressive levels of entrustment was also created.

NEXT STEPS:
Next steps include a study to obtain validity evidence for the supervision scale and determine the correlation between milestone and supervision levels. These competencies, milestones, and supervision levels can potentially serve as a roadmap for trainees and junior faculty and also play a role in the assessment of physician-scientists.
Resident-Specific Morbidity Reduced Following ACS NSQIP Data-Driven Quality Program


BACKGROUND:
The Accreditation Council for Graduate Medical Education Milestone Project for general surgery provided a more robust method for developing and tracking residents' competence. This framework enhanced systematic and progressive development of residents' competencies in surgical quality improvement.

STUDY DESIGN:
A 22-month interactive, educational program based on resident-specific surgical outcomes data culminated in a quality improvement project for postgraduate year 4 surgery residents. Self-assessment, quality knowledge test, and resident-specific American College of Surgeons National Surgical Quality Improvement Program Quality In-Training Initiative morbidity were compared before and after the intervention.

RESULTS:
Quality in-training initiative morbidity decreased from 25% (82/325) to 18% (93/517), p = 0.015 despite residents performing more complex cases. All participants achieved level 4 competency (4/4) within the general surgery milestones improvement of care, practice-based learning and improvement competency. Institutional American College of Surgeons National Surgical Quality Improvement Program general surgery morbidity improved from the ninth to the sixth decile. Quality assessment and improvement self-assessment postintervention scores (M = 23.80, SD = 4.97) were not significantly higher than preintervention scores (M = 19.20, SD = 5.26), p = 0.061. Quality Improvement Knowledge Application Tool postintervention test scores (M = 17.4, SD = 4.88), were not significantly higher than pretest scores (M = 13.2, SD = 1.92), p = 0.12.

CONCLUSION:
Sharing validated resident-specific clinical data with participants was associated with improved surgical outcomes. Participating fourth year surgical residents achieved the highest score, a level 4, in the practice based learning and improvement competency of the improvement of care practice domain and observed significantly reduced surgical morbidity for cases in which they participated.
Education Research: The Current State of Neurophysiology Education in Selected Neurology Residency Programs


OBJECTIVE:
Prior research has illustrated there is a knowledge gap in neurology residents' neurophysiology education (EEG and EMG), and we sought to understand whether this is still an issue and to recognize the barriers in order to create solutions and improve education.

METHODS:
Surveys were developed for adult neurology residents and one for program directors asking about confidence in neurophysiology knowledge, percent of graduates reaching level 4 ACGME (American Council of Graduate Medical Education) milestones in EEG and EMG, methods of learning used, interest in the subjects, and suggestions for improvements.

RESULTS:
Twenty-six program directors (19% responder rate) and 55 residents (from at least 16 different programs) completed the survey. Program directors thought that 85% of graduating residents met level 4 milestones in EEG and only 75% in EMG. Structured rotations and more time allocated to education of these topics were frequent barriers mentioned. Postgraduate year 4 residents were 60% and 67% confident in EEG and 64%, 59%, and 62.3% in EMG level 4 milestones. Learning to read EEGs was considered important throughout residents' training; however, this interest and value decreased over time with EMG.

CONCLUSION:
In our study, program directors suspect up to a quarter of residents may graduate not meeting level 4 ACGME milestones, and residents expressed lack of confidence in these areas. The educational methods used to instruct residents in EEG and EMG were similar as were the barriers they face across programs. This information hopefully will help fuel curriculum design and interest in these important neurology techniques.
INTRODUCTION:
The Accreditation Council for Graduate Medical Education (ACGME) is the governing body responsible for accrediting graduate medical training programme in the USA. The Emergency Medicine Milestones (EM-Milestones) were developed by the ACGME and American Board of Emergency Medicine as a guide and monitoring tool for the knowledge, skills, abilities and experiences to be acquired during training. Alumni surveys have been reported as a valuable resource for training programme to identify areas for improvement; however, there are few studies regarding programme improvement in emergency medicine. We aimed to use the EM-Milestones, adapted as an alumni self-assessment survey, to identify areas for training programme improvement.

METHODS:
This study was conducted at an urban, academic affiliated, community hospital in New York city with an emergency medicine training programme consisting of 30 residents over 3 years. Alumni of our emergency medicine training programme were sent an EM-Milestones-based self-assessment survey. Participants evaluated their ability in each EM-Milestones subcompetency on a Likert scale. Data were analysed using descriptive statistics.

RESULTS:
Response rate was 74% (69/93). Alumni reported achieving the target performance in 5/6 general competencies, with Systems-Based Practice falling below the target performance. The survey further identified 6/23 subcompetencies (Pharmacotherapy, Ultrasound, Wound Management, Patient Safety, Systems-Based Management and Technology) falling below the target performance level.

DISCUSSION:
Alumni self-evaluation of competence using the EM-Milestones provides valuable information concerning confidence to practice independently; these data, coupled with regular milestone evaluation of existing trainees, can identify problem areas and provide a blueprint for targeted programme improvement.
OBJECTIVE: To evaluate whether burnout was associated with emotional intelligence and job performance in surgical residents.

DESIGN: General surgery residents at a single institution were surveyed using the Maslach Burnout Inventory (MBI) and trait EI questionnaire (TEIQ-SF). Burnout was defined as scoring in 2 of the 3 following domains; Emotional Exhaustion (high), Depersonalization (high), and Personal Accomplishment (low). Job performance was evaluated using faculty evaluations of clinical competency-based surgical milestones and standardized test scores including the American Board of Surgery In-Training Exam (ABSITE) and the United States Medical Licensing Examination (USMLE) Step 3. USMLE Step 1 and USMLE Step 2, which were taken prior to residency training, were included to examine possible associations of burnout with USMLE examinations. Statistical comparison was made using Pearson correlation and simple linear regression adjusting for PGY level.

SETTING: This study was conducted at the University of Alabama at Birmingham (UAB) general surgery residency program.

PARTICIPANTS: All current and incoming general surgery residents at UAB were invited to participate in this study.

RESULTS: Forty residents participated in the survey (response rate 77%). Ten residents, evenly distributed from incoming residents to PGY-4, had burnout (25%). Mean global EI was lower in residents with burnout versus those without burnout (3.71 vs 3.9, p = 0.02). Of the 4 facets of EI, mean self-control values were lower in residents with burnout versus those without burnout (3.3 vs 4.06, p < 0.01). Each component of burnout was associated with global EI, with the strongest correlation being with personal accomplishment (r = 0.64; p < 0.01). Residents with burnout did not have significantly different mean scores for USMLE Step 1 (229 vs 237, p = 0.12), Step 2 (248 vs 251, p = 0.56), Step 3 (223 vs 222, p = 0.97), or ABSITE percentile (44.6 vs 58, p = 0.33) compared to residents without burnout. Personal accomplishment was associated with ABSITE percentile scores (r = 0.35; p = 0.049). None of the 16 surgical milestone scores were significantly associated with burnout.

CONCLUSIONS: Burnout is present in surgery residents and associated with emotional intelligence. There was no association of burnout with USMLE scores, ABSITE percentile, or surgical milestones. Traditional methods of assessing resident performance may not be capturing burnout and strategies to reduce burnout should consider targeting emotional intelligence.
Identifying Gaps in the Performance of Pediatric Trainees Who Receive Marginal/Unsatisfactory Ratings


PURPOSE:
To perform a derivation study to determine in which subcompetencies marginal/unsatisfactory pediatric residents had the greatest deficits compared with their satisfactorily performing peers and which subcompetencies best discriminated between marginal/unsatisfactory and satisfactorily performing residents.

METHOD:
Multi-institutional cohort study of all 21 milestones (rated on four or five levels) reported to the Accreditation Council for Graduate Medical Education, and global marginal/unsatisfactory versus satisfactory performance reported to the American Board of Pediatrics. Data were gathered in 2013-2014. For each level of training (postgraduate year [PGY] 1, 2, and 3), mean differences between milestone levels of residents with marginal/unsatisfactory and satisfactory performance adjusted for clustering by program and C-statistics (area under receiver operating characteristic curve) were calculated. A Bonferroni-corrected significance threshold of .0007963 was used to account for multiple comparisons.

RESULTS:
Milestone and overall performance evaluations for 1,704 pediatric residents in 41 programs were obtained. For PGY1s, two subcompetencies had almost a one-point difference in milestone levels between marginal/unsatisfactory and satisfactory trainees and outstanding discrimination (≥0.90): organize/prioritize (0.93; C-statistic: 0.91) and transfer of care (0.97; C-statistic: 0.90). The largest difference between marginal/unsatisfactory and satisfactory PGY2s was trustworthiness (0.78). The largest differences between marginal/unsatisfactory and satisfactory PGY3s were ethical behavior (1.17), incorporating feedback (1.03), and professionalization (0.96). For PGY2s and PGY3s, no subcompetencies had outstanding discrimination.

CONCLUSIONS:
Marginal/unsatisfactory pediatric residents had different subcompetency gaps at different training levels. While PGY1s may have global deficits, senior residents may have different performance deficiencies requiring individualized counseling and targeted performance improvement plans.
Comprehensive Assessment of Struggling Learners Referred to a Graduate Medical Education Remediation Program


BACKGROUND: Implementation of the Next Accreditation System has provided a standardized framework for identifying learners not meeting milestones, but there is as yet no corresponding framework for remediation.

OBJECTIVE: We developed a comprehensive assessment process that allows correct diagnosis of a struggling learner's deficit(s) to promote successful remediation.

METHODS: At the University of Pennsylvania, resident learners within the Department of Medicine who are not meeting milestones are referred to the Early Intervention Remediation Committee (EIRC). The EIRC, composed of 14 faculty members with expertise in remediation, uses a standardized process to assess learners' deficits. These faculty members categorize primary deficits as follows: medical knowledge, clinical reasoning, organization and efficiency, professionalism, and communication skills. The standardized process of assessment includes an analysis of the learner's file, direct communication with evaluators, an interview focused on learner perception of the problem, screening for underlying medical or psychosocial issues, and a review of systems for deficits in the 6 core competencies. Participants were surveyed after participating in this process.

RESULTS: Over a 2-year period, the EIRC assessed and developed remediation plans for 4% of learners (14 of a total 342). Following remediation and reassessment, the identified problems were satisfactorily resolved in all cases with no disciplinary action. While the process was time intensive, an average of 45 hours per learner, the majority of faculty and residents rated it as positive and beneficial.

CONCLUSIONS: This structured assessment process identifies targeted areas for remediation and adds to the tools available to Clinical Competency Committees.
Impact of an Event Reporting System on Resident Complication Reporting in Plastic Surgery Training: Addressing an ACGME and Plastic Surgery Milestone Project Core Competency


BACKGROUND:
The Accreditation Council for Graduate Medical Education and Plastic Surgery Milestone Project has identified practice-based learning and improvement, which involves systematically analyzing current practices and implementing changes, as a core competency in residency education. In surgical care, complication reporting is an essential component of practice-based learning and improvement as complications are analyzed in morbidity and mortality conference for quality improvement. Unfortunately, current methods for capturing a comprehensive profile of complications may significantly underestimate the true occurrence of complications. Therefore, the objectives of this study are to evaluate an intervention for complication reporting and compare this to current practice, in a plastic surgery training program.

METHODS:
This is a preintervention and postintervention study evaluating resident reporting of complications on a plastic surgery service. The intervention was an online event reporting system developed by department leadership and patient safety experts. The cohorts consisted of all patients undergoing surgery during two separate 3-month blocks bridged by an implementation period. A trained reviewer recorded complications, and this served as the reference standard. Fisher's exact test was used for binary comparisons.

RESULTS:
There were 32 complications detected in 219 patients from June to August of 2015 and 35 complications in 202 patients from October to December of 2015. The proportion of complications reported in the preintervention group was nine of 32 (28.1 percent). After the intervention, this significantly increased to 32 of 35 (91.4 percent) (p < 0.001).

CONCLUSION:
An intervention using an event reporting system, supported by departmental leadership, led to significant improvements in complication reporting by plastic surgery residents.
Using the Teaching Perspectives Inventory as an Introduction to a Residents-as-Teachers Curriculum


OBJECTIVE:
The Anesthesiology Milestone Project includes a milestone for assessing the teaching attributes of residents within the competency of Practice-based Learning and Improvement. We intend to develop a Residents-as-Teachers educational curriculum to assist our residents in successfully achieving this milestone. The goal of this study is to identify the specific teaching perspectives and intentions of our residents and to promote residents' comprehension of their own teaching philosophy.

METHODS:
We invited our residents to complete the Teaching Perspective Inventory (TPI) and a follow-up survey to gather information regarding dominant and recessive teaching perspectives, their intended career pathway, and their view of the importance of understanding teaching perspectives.

RESULTS:
The two most common dominant teaching perspectives are apprenticeship and nurturing for residents who are planning a career in both academic medicine and private practice. A greater percentage of residents planning an academic career agree that identifying their teaching perspective is beneficial to their role as a clinical educator, compared to those anticipating a career in private practice.

CONCLUSIONS:
Based on this pilot data, our Residents-as-Teachers curriculum will include instruction of educational strategies specifically designed towards the apprenticeship and nurturing perspectives.
Towards Consensus: Training in Procedural Skills for Diagnostic Radiology Residents – Current Opinions of Residents and Faculty at a Large Academic Center


RATIONALE AND OBJECTIVES:
The Diagnostic Radiology Milestones Project provides a framework for measuring resident competence in radiologic procedures, but there are limited data available to assist in developing these guidelines. We performed a survey of current radiology residents and faculty at our institution as a first step toward obtaining data for this purpose. The survey addressed attitudes toward procedural standardization and procedures that trainees should be competent by the end of residency.

MATERIALS AND METHODS:
Current residents and faculty members were surveyed about whether or not there should be standardization of procedural training, in which procedures residents should achieve competency, and the number of times a procedure needs to be performed to achieve competency.

RESULTS:
Survey data were received from 60 study participants with an overall response rate of 32%. Sixty-five percent of respondents thought that procedural training should be standardized. Standardization of procedural training would include both the list of procedures that trainees should be competent in at the end of residency and the standard minimum number of procedures to achieve competency. Procedures that both residents and faculty agreed are important in which to achieve competency included central line/port procedures; CT-guided abdominal, thoracic, and musculoskeletal procedures; minor fluoroscopic-guided procedures; general fluoroscopy; peripheral line placements; and US-guided abdominal procedures. For most of these categories, most respondents believed that these procedures needed to be performed 6-20 times to achieve competency.

CONCLUSION:
Both resident and faculty respondents agreed that procedural training should be standardized during residency, and competence in specific procedures should be achieved at the completion of residency. Although this study is limited to a single institution, our data may provide assistance in developing future guidelines for standardizing image-guided procedure training. Future studies could be expanded to create a national consensus regarding the implementation of the Diagnostic Radiology Milestones Project.
How Do Clinical Competency Committees Use Different Sources of Data to Assess Residents' Performance on the Internal Medicine Milestones? A Mixed Methods Pilot Study


PURPOSE:
This study examines how Clinical Competency Committees (CCCs) synthesize assessment data to make judgments about residents' clinical performances.

METHODS:
Between 2014 and 2015, after four six-month reporting periods to the Accreditation Council for Graduate Medical Education (ACGME), 7 of 16 CCC faculty at Rush University Medical Center completed questionnaires focused on their perspectives about rating residents on their achievement of the milestones and participated in a focus group. Qualitative data were analyzed using grounded theory. Milestones ratings for two six-month ACGME reporting cycles (n = 100 categorical residents) were also analyzed.

RESULTS:
CCC members weighted resident rotation ratings highest (weight = 37%), followed by faculty rotation comments (weight = 27%) and personal experience with residents (weight = 14%) for making judgments about learner's milestone levels. Three assessment issues were identified from qualitative analyses: (1) "design issues" (e.g. problems with available data or lack thereof); (2) "synthesis issues" (e.g. factors influencing ratings and decision-making processes) and (3) "impact issues" (e.g., how CCC generated milestones ratings are used).

CONCLUSIONS:
Identifying factors that affect assessment at all stages of the CCC process can contribute to improving assessment systems, including support for faculty development for CCCs. Recognizing challenges in synthesizing first and second-hand assessment data is an important step in understanding the CCC decision-making process.
Mindfulness, Burnout, and Effects on Performance Evaluations in Internal Medicine Residents


PURPOSE:
Burnout has been documented at high levels in medical residents with negative effects on performance. Some dispositional qualities, like mindfulness, may protect against burnout. The purpose of the present study was to assess burnout prevalence among internal medicine residents at a single institution, examine the relationship between mindfulness and burnout, and provide preliminary findings on the relation between burnout and performance evaluations in internal medicine residents.

METHODS:
Residents (n = 38) completed validated measures of burnout at three time points separated by 2 months and a validated measure of dispositional mindfulness at baseline. Program director end-of-year performance evaluations were also obtained on 22 milestones used to evaluate internal medicine resident performance; notably, these milestones have not yet been validated for research purposes; therefore, the investigation here is exploratory.

RESULTS:
Overall, 71.1% (n = 27) of the residents met criteria for burnout during the study. Lower scores on the "acting with awareness" facet of dispositional mindfulness significantly predicted meeting burnout criteria $\chi^2(5) = 11.88, p = 0.04$. Lastly, meeting burnout criteria significantly predicted performance on three of the performance milestones, with positive effects on milestones from the "system-based practices" and "professionalism" domains and negative effects on a milestone from the "patient care" domain.

CONCLUSION:
Burnout rates were high in this sample of internal medicine residents and rates were consistent with other reports of burnout during medical residency. Dispositional mindfulness was supported as a protective factor against burnout. Importantly, results from the exploratory investigation of the relationship between burnout and resident evaluations suggested that burnout may improve performance on some domains of resident evaluations while compromising performance on other domains. Implications and directions for future research are discussed.
The Effect and Use of Milestones in the Assessment of Neurological Surgery Residents and Residency Programs


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.
Critical Deficiency Ratings in Milestone Assessment: A Review and Case Study


PURPOSE:
The Accreditation Council for Graduate Medical Education (ACGME) requires programs to report learner progress using specialty-specific milestones. It is unclear how milestones can best identify critical deficiencies (CDs) in trainee performance. Specialties developed milestones independently of one another; not every specialty included CDs within milestones ratings. This study examined the proportion of ACGME milestone sets that include CD ratings, and describes one residency program's experiences using CD ratings in assessment.

METHOD:
The authors reviewed ACGME milestones for all 99 specialties in November 2015, determining which rating scales contained CDs. The authors also reviewed three years of data (July 2012-June 2015) from the University of Cincinnati Medical Center (UCMC) internal medicine residency assessment system based on observable practice activities mapped to ACGME milestones. Data were analyzed by postgraduate year, assessor type, rotation, academic year, and core competency. The Mantel-Haenszel chi-square test was used to test for changes over time.

RESULTS:
Specialties demonstrated heterogeneity in accounting for CDs in ACGME milestones, with 22% (22/99) of specialties having no language describing CDs in milestones assessment. Thirty-three percent (63/189) of UCMC internal medicine residents received at least one CD rating, with CDs accounting for 0.18% (668/364,728) of all assessment ratings. The authors identified CDs across multiple core competencies and rotations.

CONCLUSIONS:
Despite some specialties not accounting for CDs in milestone assessment, UCMC's experience demonstrates that a significant proportion of residents may be rated as having a CD during training. Identification of CDs may allow programs to develop remediation and improvement plans.
Evaluating Surgical Residents Quickly and Easily Against the Milestones Using Electronic Formative Feedback


OBJECTIVE:
This study was conducted to assess the effectiveness of a newly implemented electronic web-based review system created at our institution for evaluating resident performance relative to established milestones.

DESIGN:
Retrospective review of data collected from a survey of general surgery faculty and residents.

SETTING:
Tertiary care teaching hospital system and independent academic medical center.

PARTICIPANTS:
A total of 12 general surgery faculty and 17 general surgery residents participated in this study. The survey queried the level of satisfaction before and after the adoption of QuickNotes using several statements scored on a 5-point scale, with 1 being the lowest rating as "not satisfied," and 5 being the highest rating as "completely satisfied."

RESULTS:
The weighted average improvements from pre- to post-QuickNotes implementation for the faculty responding to the survey ranged from 10% to 40%; weighted average improvements for the residents responding to the survey ranged from 5% to 73%. For the survey of faculty, both sets of weighted averages tended to be higher than the weighted average for the resident's survey responses. The highest rated topic was the faculty's level of satisfaction with the "frequency to provide feedback" with a post-QuickNotes implementation weighted average of 4.25, closely followed by the residents' level of satisfaction with the "evaluation includes positive feedback" with a post-QuickNotes implementation weighted average of 4.24. The most notable increases in weighted averages from preimplementation to postimplementation were noted for "overall satisfaction" (20% increase for faculty, 37% for residents), "reflects actual criteria that matter" (36% increase for faculty, 73% for residents), faculty "opportunity for follow-up" (increase of 40%), resident "reflects overall trends" (increase of 37%), and resident "provides new information about my performance" (increase of 37%).

CONCLUSIONS:
Our institutional adoption of QuickNotes into the resident evaluation process has been associated with an overall increased level of satisfaction in the evaluation process by both faculty and residents. The design of QuickNotes facilitates its integration into the resident training environment, as it is web based, easy to use, and has no additional cost over the standard New Innovations subscription. Although it is designed to capture snapshots of trainee behavior and performance, monthly reports through QuickNotes can be used effectively in conjunction with the more traditional end-of-rotation evaluations to show trends, identify areas of strength that should be reinforced, demonstrate areas needing improvement, allow for a more tailored individual education plan to be developed, and permit a more accurate determination of milestone progression.
Competent for Unsupervised Practice: Use of Pediatric Residency Training Milestones to Assess Readiness


PURPOSE:
To describe clinical skills progression during pediatric residency using the distribution of pediatric milestone assessments by subcompetency and year of training and to determine reasonable milestone expectations at time of graduation.

METHOD:
Multi-institutional cohort study of the milestones reported to the Accreditation Council for Graduate Medical Education for all 21 pediatric subcompetencies. Most subcompetencies were measured using five milestone levels (1 = novice, 2 = advanced beginner, 3 = competent, 4 = proficient, 5 = master); 3 subcompetencies had only four levels defined.

RESULTS:
Milestone assessments for 2,030 pediatric residents in 47 programs during academic year 2013-2014 were obtained. There was significant variation in end-of-year milestone ratings for residents within each level of training, which decreased as training level increased. Most (78.9%; 434/550) graduating third-year pediatric residents received a milestone rating of ≥ 3 in all 21 subcompetencies; fewer (21.1%; 116/550) received a rating of ≥ 4 in all subcompetencies. Across all training levels, professionalism and interpersonal communication skills were rated highest; quality improvement was rated lowest.

CONCLUSIONS:
Trainees entered residency with a wide range of skills. As they advanced, skill variability within a training level decreased. Most graduating pediatric residents were still advancing on the milestone continuum toward proficiency and mastery, and an expectation of milestone ratings ≥ 4 in all categories upon graduation is unrealistic; milestone ratings ≥ 3 upon graduation may be more realistic. Understanding current pediatric residents' and graduates' skills can help to identify key areas that should be specifically targeted during training.
Milestones: A Rapid Assessment Method for the Clinical Competency Committee


INTRODUCTION:
Educational milestones are now used to assess the developmental progress of all U.S. graduate medical residents during training. Twice annually, each program's Clinical Competency Committee (CCC) makes these determinations and reports its findings to the Accreditation Council for Graduate Medical Education (ACGME). The ideal way to conduct the CCC is not known. After finding that deliberations reliant upon the new milestones were time intensive, our internal medicine residency program tested an approach designed to produce rapid but accurate assessments.

MATERIAL AND METHODS:
For this study, we modified our usual CCC process to include pre-meeting faculty ratings of resident milestones progress with in-meeting reconciliation of their ratings. Data were considered largely via standard report and presented in a pre-arranged pattern. Participants were surveyed regarding their perceptions of data management strategies and use of milestones. Reliability of competence assessments was estimated by comparing pre-/post- intervention class rank lists produced by individual committee members with a master class rank list produced by the collective CCC after full deliberation.

RESULTS:
Use of the study CCC approach reduced committee deliberation time from 25 min to 9 min per resident ($p < 0.001$). Committee members believed milestones improved their ability to identify and assess expected elements of competency development ($p = 0.026$). Individual committee member assessments of trainee progress agreed well with collective CCC assessments.

CONCLUSIONS:
Modification of the clinical competency process to include pre-meeting competence ratings with in-meeting reconciliation of these ratings led to shorter deliberation times, improved evaluator satisfaction and resulted in reliable milestone assessments.
OBJECTIVE:
The Accreditation Council for Graduate Medical Education requires accredited general surgery residencies to implement competency-based developmental outcomes in resident evaluations. Overall, 16 milestones are evaluated by a clinical competency committee (CCC). The milestones span 8 domains of surgical practice and 6 Accreditation Council for Graduate Medical Education clinical competencies. The highest level suggests preparedness for independent practice. Our objective was to compare self-assessments and committee evaluations within the milestone framework.

STUDY DESIGN:
All residents underwent semiannual evaluations from 2013 to 2015. Residents independently completed a self-assessment using the milestones. The CCC completed the milestones document using resident evaluations and consensus opinion of committee members. Assessment differences were calculated for each evaluation. A negative value indicated that the residents evaluated themselves at a lower level than the committee. Major assessment disparities were defined as >0.5 on a 4-point scale.

SETTING:
An independent academic medical center.

PARTICIPANTS:
General surgery residents.

RESULTS:
Overall, 20 residents participated; 7 were female. In total, 5 (7%) evaluations had a mean overall assessment difference >0.5, whereas 6 (8%) had a difference <-0.5. Residents evaluated themselves lower than the committee with a median assessment difference of -0.06 [-0.25 to 0.16] (p = 0.041). Evaluations were similar across surgical domains. Negative self-evaluations were more common for medical knowledge (-0.25 [-0.25 to 0.25], p = 0.025). Female residents had 2% positive and 13% negative major assessment disparity rates versus 10% positive and 9% negative rates among male residents. Postgraduate year III residents had 12% positive and 4% negative major disparity rates; all other years had higher negative than positive rates.

CONCLUSIONS:
Surgery residents within our program demonstrated adequate self-awareness, with most self-evaluations falling within a half level of the CCC report. This self-awareness was consistent across surgical domains and most clinical competencies. Residents perceived a lower level of medical knowledge than the CCC. Subgroup analysis revealed interesting trends in the effects of sex, postgraduate year level, and academic year timing, which will take additional study to fully delineate.
Using Milestones as Evaluation Metrics during an Emergency Medicine Clerkship


BACKGROUND:
The Accreditation Council for Graduate Medical Education's (ACGME) Milestones presumes graduating medical students will enter residency proficient at Milestone level 1 for 23 skills. The Next Accreditation System now includes Milestones for each postgraduate specialty, and it is unlikely that schools will document every emergency medicine (EM) applicant's EM-specific skills in their performance evaluation.

OBJECTIVES:
The goals of this research were to determine if assessment of the Milestones was feasible during a medical student clerkship and examine the proportion of medical students performing at Milestone Level 1.

METHODS:
This study was conducted at a center with Liaison Committee on Medical Education-approved medical training and a 4-year EM residency. Using traditional clerkship, we studied the feasibility of an ACGME EM Milestones-based clerkship assessment. Data led to redesign of the clerkship and its evaluation process, including all level 1 anchor(s) to add "occasionally" (>60%), "usually" (>80%), and "always" (100%) on a Likert scale to on-shift assessment forms.

RESULTS:
During the feasibility phase (2013-14), 75 students rotated though the clerkship; 55 evaluations were issued and 50 contained the Milestone summary. Eight deficiencies were noted in Milestone 12 and three in Milestone 14. After changes, 49 students rotated under the new evaluation rubric. Of 575 completed on-shift evaluations, 16 Milestone deficiencies were noted. Of 41 institutional evaluations issued, only one student had deficiencies noted, all of which pertained to patient care. All evaluations in this second cohort contained each student's Milestone proficiency.

CONCLUSIONS:
Assessment of the Milestones is feasible. Communication of ACGME EM Milestone proficiency may identify students who require early observation or remediation. The majority of students meet the anchors for the Milestones, suggesting that clerkship assessment with the ACGME EM Milestones does not adequately differentiate students.
A First Look at the Accreditation Council for Graduate Medical Education Anesthesiology Milestones: Implementation of Self-Evaluation in a Large Residency Program


STUDY OBJECTIVE:
The objective was to determine if there is a correlation between resident postgraduate year (PGY) of training and self-evaluation of performance using the Accreditation Council for Graduate Medical Education milestones.

DESIGN:
Survey.

SETTING:
Residency program at a large academic center.

PATIENTS:
Residents and Faculty Clinical Competency Committee (CCC).

INTERVENTIONS:
None.

MEASUREMENTS:
Resident and CCC milestone scores.

MAIN RESULTS:
Correlation coefficients for average score for each milestone vs PGY level ranged from 0.80 for receiving and giving feedback to 0.95 for anesthetic choice and conduct. All milestones showed a relatively linear relationship with PGY of training, and none were found to be consistently reached very late or very early in training. When examining variation across the scores for the individual residents, the distributions for PGY-2 and -3 appeared to be wider than those for PGY-1 and -4. The intraclass correlation coefficients ranged from 0.718 to 0.928.

CONCLUSIONS:
There was a remarkable degree of consistency in the relationship between level of training and resident self-assessment score for every milestone, as well as strong agreement between the resident and CCC faculty scores. Examination of the variance in the scores, when interpreted in light of our particular training program’s characteristics, suggests that the milestones accurately reflect the progression in skill across the residency. In addition, given the concordance between the self-evaluation scores and the CCC faculty scores, self-evaluation may be a reasonable starting point as programs begin the daunting task of determining scores for each of the 25 milestones as part of the biannual evaluation process.
A Survey of Ultrasound Milestone Incorporation into Emergency Medicine Training Programs


OBJECTIVES:
With the introduction of the Emergency Medicine Milestone Project in 2013, residencies now assess emergency ultrasound (US) skills at regular intervals. However, it is unclear how programs are implementing the emergency US milestones and assessing competency. With the use of the milestone tool, a survey was distributed to emergency US educators to determine when programs are providing emergency US education, when residents are expected to attain competency, and whether the milestones reflect their expectations of trainees.

METHODS:
We conducted a prospective cross-sectional survey study distributed electronically to designated emergency US experts at 169 programs. Participants were queried on education and competency evaluation within the context of the milestones by designating a postgraduate year when the 5 milestone levels were taught and competency was expected. Survey findings were reported as percentages of total respondents from descriptive statistics.

RESULTS:
Responses were received from 53% of programs, and 99% were familiar with the milestones. Most programs provide level 1 (88%) and 2 (85%) instruction during postgraduate year 1. Most programs expect level 1 competency before residency (61%) and expect mastery of level 2 by the end of postgraduate year 1 (60%). Sixty-two percent believe the milestones do not accurately reflect their expectations, citing insufficient minimum scan numbers, lack of specificity, and unattainable level 5 requirements.

CONCLUSIONS:
There is substantial variability in the frequency and methods of competency evaluation using the emergency US milestones. However, most responders agree that residents should obtain level 2 competency by postgraduate year 1. Variation exists regarding what year and what skills define level 3 or greater competency.
The Problem Resident Behavior Guide: Strategies for Remediation


ABSTRACT

In 2012, the ACGME supplemented the core competencies with outcomes-based milestones for resident performance within the six competency domains. These milestones address the knowledge, skills, abilities, attitudes, and experiences that a resident is expected to progress through during the course of training. Even prior to the initiation of the milestones, there was a paucity of EM literature addressing the remediation of problem resident behaviors and there remain few readily accessible tools to aid in the implementation of a remediation plan. The goal of the "Problem Resident Behavior Guide" is to provide specific strategies for resident remediation based on deficiencies identified within the framework of the EM milestones. The "Problem Resident Behavior Guide" is a written instructional manual that provides concrete examples of remediation strategies to address specific milestone deficiencies. The more than 200 strategies stem from the experiences of the authors who have professional experience at three different academic hospitals and emergency medicine residency programs, supplemented by recommendations from educational leaders as well as utilization of valuable education adjuncts, such as focused simulation exercises, lecture preparation, and themed ED shifts. Most recommendations require active participation by the resident with guidance by faculty to achieve the remediation expectations. The ACGME outcomes-based milestones aid in the identification of deficiencies with regards to resident performance without providing recommendations on remediation. The Problem Resident Behavior Guide can therefore have a significant impact by filling in this gap.
Driving Care Quality: Aligning Trainee Assessment and Supervision through Practical Application of Entrustable Professional Activities, Competencies, and Milestones


ABSTRACT
To address the long-standing challenge of meaningful trainee assessment, the authors reviewed and expanded on the Accountable Assessment for Quality Care and Supervision (AAQCS) equation. The equation proposes that care quality is the product of the interaction between trainee performance (measured by workplace assessment) and supervision (required level of intervention to ensure care quality) in the context of the environment where the care occurs: Trainee performance × Appropriate supervision = Safe, effective patient-centered care. Assessing trainee performance and matching that performance to "appropriate" supervision, however, is fraught with challenges. The authors suggest a unifying framework that integrates entrustable professional activities (EPAs), competencies, and milestones to inform trainee assessment and supervision, thereby enabling the practical application of the AAQCS equation in the workplace. Because the unit of measure for an EPA is the outcome of whether the trainee can safely and effectively perform the professional activity without supervision, the proposed unifying framework directly aligns with the dependent variable in the AAQCS equation: care quality. The value of applying a unifying framework that integrates EPAs, competencies, and milestones to the AAQCS equation in the clinical learning environment lies in its ability to provide supervisors with a shared mental model of performance expectations for trainees, reducing unwanted variability and improving assessment accuracy; guidance for aligning performance milestones of trainees with the needed level of supervisor intervention to ensure care quality; and substrate for specific feedback to improve the trainee's professional development as a way to ensure future care quality.
A Milestone-Based Evaluation System-The Cure for Grade Inflation?


PURPOSE:
Controversy exists over the optimal use of the Milestones in the process of resident Evaluation and feedback. We sought to evaluate the performance of a Milestones-based feedback system in comparison to a traditional model.

METHODS:
The traditional evaluation system (TES) consisted of a generic 16-item survey using a 5-point Likert scale ranging from 1 to 5, and a free-text comments section. The Milestones-based evaluation system (MBES) was launched in July 2014, ranging from 0 to 4. Individual milestones were mapped to rotations based on resident educational goals by postgraduate year (PGY). The MBES consisted of a survey with a maximum of 7 items, followed by a free-text comment section. Within each evaluation system, an overall composite score was calculated for each categorical general surgical resident. To scale the 2 systems for comparison, TES scores were adjusted downward by 1 point. Descriptive statistics were performed. Univariate analysis was performed with the Wilcoxon signed-rank test. A test for trend across PGY was used for the MBES only.

RESULTS:
In the traditional system, the median score was 3.66 (range: 3.2-4.0). There was no meaningful difference in the median score by PGY. In the new system, the median score was 2.69 (range: 1.5-3.7, p < 0.01). The median score differed across PGY and increased by PGY of training (p < 0.01). There was an increase in differences between median scores by PGY.

CONCLUSIONS:
On using the milestones to facilitate faculty evaluation of resident knowledge and skill, there was a trend in increasing score by PGY of training. In the MBES, scores could be used to better discriminate resident skill and knowledge levels and resulted in improved differentiation in scoring by PGY. The use of the milestones as a basis for evaluation enabled the program to provide more meaningful feedback to residents and represents an improvement in surgical education.
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.
Development and Evaluation of Standardized Narrative Cases Depicting the General Surgery Professionalism Milestones


PURPOSE:
Residency programs now are required to use educational milestones, which has led to the need for new methods of assessment. The literature suggests that narrative cases are a promising tool to track residents' progress. This study demonstrates the process for developing and evaluating narrative cases representing the five levels of the professionalism milestones.

METHOD:
In 2013, the authors identified 28 behaviors in the Accreditation Council for Graduate Medical Education general surgery professionalism milestones. They modified previously published narrative cases to fit these behaviors. To evaluate the quality of these cases, the authors developed a 28-item, five-point scale instrument, which 29 interdisciplinary faculty completed. The authors compared the faculty ratings by narrative case and specialty with the authors' initial rankings of the cases by milestone level. They used t tests and analysis of variance to compare mean scores across specialties.

RESULTS:
The authors developed 10 narrative cases, 2 for each of the 5 milestone levels. Each case contained at least 20 of the 28 behaviors identified in the milestones. Mean faculty ratings matched the milestone levels. Reliability was good (G coefficient = 0.86, phi coefficient = 0.85), indicating consistency in raters' ability to determine the proper milestone level for each case.

CONCLUSIONS:
The authors demonstrate a process for using specialty-specific milestones to develop narrative cases that map to a spectrum of professionalism behaviors. This process can be applied to other competencies and specialties to facilitate faculty awareness of resident performance descriptors and provide a frame of reference for milestones assessment.
Development and Validation of an Assessment of Regional Anesthesia Ultrasound Interpretation Skills


**BACKGROUND:**
Interpretation of ultrasound images and knowledge of anatomy are essential skills for ultrasound-guided peripheral nerve blocks. Competency-based educational models promoted by the Accreditation Council for Graduate Medical Education require the development of assessment tools for the achievement of different competency milestones to demonstrate the longitudinal development of skills that occur during training.

**METHODS:**
A rigorous study guided by psychometric principles was undertaken to identify and validate the domains and items in an assessment of ultrasound interpretation skills for regional anesthesia. A survey of residents, academic faculty, and community anesthesiologists, as well as video recordings of experts teaching ultrasound-guided peripheral nerve blocks, was used to develop short video clips with accompanying multiple choice-style questions. Four rounds of pilot testing produced a 50-question assessment that was subsequently administered online to residents, fellows, and faculty from multiple institutions.

**RESULTS:**
Test results from 90 participants were analyzed with Item Response Theory model fitting indicating that a 47-item subset of the test fits the model well (P = 0.11). There was a significant linear relation between expected and predicted item difficulty (P < 0.001). Overall test scores increased linearly with higher levels of formal anesthesia training, regional anesthesia training, number of ultrasound-guided blocks performed per year, and a self-rating of regional anesthesia skill (all P < 0.001).

**CONCLUSIONS:**
This study provides evidence for the reliability, content validity, and construct validity of a 47-item multiple choice-style online test of ultrasound interpretation skills for regional anesthesia, which can be used as an assessment of competency milestone achievement in anesthesiology training.
Use of Emergency Medicine Milestones as Items on End-of-Shift Evaluations Results in Overestimates of Residents' Proficiency Level


BACKGROUND:
The emergency medicine milestones were developed to provide more objective resident assessment than current methods. However, little is known about the best practices for applying the milestones in resident assessment.

OBJECTIVE:
We examined the utility of end-of-shift evaluations (ESEs) constructed using the milestones in resident assessment.

METHODS:
We developed 14 daily ESEs, each of which included 9 or 10 emergency medicine milestones. Postgraduate year (PGY)-1 and PGY-2 residents were assessed on milestone levels 1 through 3; PGY-3 and PGY-4 residents were assessed on levels 3 through 5. Each milestone was rated on a nominal scale (yes, no, or not applicable). The Clinical Competency Committee combined the ESE data with data from other assessments to determine each resident's proficiency level for the emergency medicine subcompetencies. We used descriptive statistics to summarize resident ESEs and milestone levels. We analyzed differences in ESE score across PGY levels using t tests and analyses of variance.

RESULTS:
Faculty completed 763 ESEs on 33 residents with a range of 2 to 54 (median=22) ESEs per resident. Faculty rarely (8%, 372 of 4633) rated a resident as not achieving a milestone on the ESEs. Analyses of variance revealed that ESE scores on level 3 milestones did not differ significantly by PGY level. There was poor agreement between ESE scores and Clinical Competency Committee ratings.

CONCLUSIONS:
The ESEs constructed using the milestones resulted in grade or milestone inflation. Our results do not support using milestones as a stand-alone assessment tool.
Realizing the Promise of Competency-Based Medical Education


ABSTRACT
Competency-based medical education (CBME) places a premium on both educational and clinical outcomes. The Milestones component of the Next Accreditation System represents a fundamental change in medical education in the United States and is part of the drive to realize the full promise of CBME. The Milestones framework provides a descriptive blueprint in each specialty to guide curriculum development and assessment practices. From the beginning of the Outcomes project in 1999, the Accreditation Council for Graduate Medical Education and the larger medical education community recognized the importance of improving their approach to assessment. Work-based assessments, which rely heavily on the observations and judgments of clinical faculty, are central to a competency-based approach. The direct observation of learners and the provision of robust feedback have always been recognized as critical components of medical education, but CBME systems further elevate their importance. Without effective and frequent direct observation, coaching, and feedback, the full potential of CBME and the Milestones cannot be achieved. Furthermore, simply using the Milestones as end-of-rotation evaluations to "check the box" to meet requirements undermines the intent of an outcomes-based accreditation system. In this Commentary, the author explores these challenges, addressing the concerns raised by Williams and colleagues in their Commentary. Meeting the assessment challenges of the Milestones will require a renewed commitment from institutions to meet the profession's "special obligations" to patients and learners. All stakeholders in graduate medical education must commit to a professional system of self-regulation to prepare highly competent physicians to fulfill this social contract.
Milestone-Based Assessments Are Superior to Likert-Type Assessments in Illustrating Trainee Progression


BACKGROUND:
The Pediatrics Milestone Project uses behavioral anchors, narrative descriptions of observable behaviors, to describe learner progression through the Accreditation Council for Graduate Medical Education competencies. Starting June 2014, pediatrics programs were required to submit milestone reports for their trainees semiannually. Likert-type scale assessment tools were not designed to inform milestone reporting, creating a challenge for Clinical Competency Committees.

OBJECTIVE:
To determine if milestone-based assessments better stratify trainees by training level compared to Likert-type assessments.

METHODS:
We compared assessment results for 3 subcompetencies after changing from a 5-point Likert scale to milestone-based behavioral anchors in July 2013. Program leadership evaluated the new system by (1) comparing PGY-1 mean scores on Likert-type versus milestone-based assessments; and (2) comparing mean scores on the Likert-type versus milestone-based assessments across PGY levels.

RESULTS:
Mean scores for PGY-1 residents were significantly higher on the prior year’s Likert-type assessments than milestone-based assessments for all 3 subcompetencies (P, .01). Stratification by PGY level was not observed with Likert-type assessments (eg, interpersonal and communication skills 1 [ICS1] mean score for PGY-1, 3.99 versus PGY-3, 3.98; P 5 .98). In contrast, milestone-based assessments demonstrated stratification by PGY level (eg, the ICS1 mean score was 3.06 for PGY-1, 3.83 for PGY-2, and 3.99 for PGY-3; P,.01 for PGY-1 versus PGY-3). Significantly different means by trainee level were noted across 21 subcompetencies on milestone-based assessments (P, .01 for PGY-1 versus PGY-3).

CONCLUSIONS:
Initial results indicate milestone-based assessments stratify trainee performance by level better than Likert-type assessments. Average PGY-level scores from milestone-based assessments may ultimately provide guidance for determining whether trainees are progressing at the expected pace.
Impact on House Staff Evaluation Scores When Changing from a Dreyfus- to a Milestone-Based Evaluation Model: One Internal Medicine Residency Program’s Findings


PURPOSE:
As graduate medical education (GME) moves into the Next Accreditation System (NAS), programs must take a critical look at their current models of evaluation and assess how well they align with reporting outcomes. Our objective was to assess the impact on house staff evaluation scores when transitioning from a Dreyfus-based model of evaluation to a Milestone-based model of evaluation. Milestones are a key component of the NAS.

METHOD:
We analyzed all end of rotation evaluations of house staff completed by faculty for academic years 2010-2011 (pre-Dreyfus model) and 2011-2012 (post-Milestone model) in one large university-based internal medicine residency training program. Main measures included change in PGY-level average score; slope, range, and separation of average scores across all six Accreditation Council for Graduate Medical Education (ACGME) competencies.

RESULTS:
Transitioning from a Dreyfus-based model to a Milestone-based model resulted in a larger separation in the scores between our three post-graduate year classes, a steeper progression of scores in the PGY-1 class, a wider use of the 5-point scale on our global end of rotation evaluation form, and a downward shift in the PGY-1 scores and an upward shift in the PGY-3 scores.

CONCLUSIONS:
For faculty trained in both models of assessment, the Milestone-based model had greater discriminatory ability as evidenced by the larger separation in the scores for all the classes, in particular the PGY-1 class.
The Milestones Passport: A Learner-Centered Application of the Milestone Framework to Prompt Real-Time Feedback in the Emergency Department


BACKGROUND:
In July 2013, emergency medicine residency programs implemented the Milestone assessment as part of the Next Accreditation System.

OBJECTIVE:
We hypothesized that applying the Milestone framework to real-time feedback in the emergency department (ED) could affect current feedback processes and culture. We describe the development and implementation of a Milestone-based, learner-centered intervention designed to prompt real-time feedback in the ED.

METHODS:
We developed and implemented the Milestones Passport, a feedback intervention incorporating subcompetencies, in our residency program in July 2013. Our primary outcomes were feasibility, including faculty and staff time and costs, number of documented feedback encounters in the first 2 months of implementation, and user-reported time required to complete the intervention. We also assessed learner and faculty acceptability.

RESULTS:
Development and implementation of the Milestones Passport required 10 hours of program coordinator time, 120 hours of software developer time, and 20 hours of faculty time. Twenty- eight residents and 34 faculty members generated 257 Milestones Passport feedback encounters. Most residents and faculty reported that the encounters required fewer than 5 minutes to complete, and 48% (12 of 25) of the residents and 68% (19 of 28) of faculty reported satisfaction with the Milestones Passport intervention. Faculty satisfaction with overall feedback in the ED improved after the intervention (93% versus 54%, \( P = .003 \)), whereas resident satisfaction with feedback did not change significantly.

CONCLUSIONS:
The Milestones Passport feedback intervention was feasible and acceptable to users; however, learner satisfaction with the Milestone assessment in the ED was modest.
Programmatic Assessment of Level 1 Milestones in Incoming Interns


OBJECTIVES:
With the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System, emergency medicine (EM) residency programs will be required to report residents' progress through the EM milestones. The milestones include five progressively advancing skill levels, with Level 1 defining the skill set of a medical school graduate and Level 5, that of an attending physician. The ACGME stresses that multiple forms of assessment should be used to ensure capture of the multifaceted competencies. The objective of this study was to determine the feasibility and results of programmatic assessment of Level 1 milestones using multisource assessments for incoming EM interns in July.

METHODS:
The study population was interns starting in 2012 and 2013. Interns' Level 1 milestone assessment was done with four distinct methods: 1) the postgraduate orientation assessment (POA) by the Graduate Medical Education Office for all incoming interns (this multistation examination covers nine of the EM milestones and includes standardized patient cases, task completion, and computer-based stations); 2) direct observation of patient encounters by core faculty using a milestones-based clinical skills competency checklist; 3) the global monthly assessment at the end of the intern orientation month that was updated to reflect the EM milestones; and 4) faculty assessment during procedural labs. These occurred during the July orientation month that included the POA, clinical shifts, didactic sessions, and procedure labs.

RESULTS:
In the POA, interns were competent in 48% to 93% of the milestones assessed. Overall, competency was 70% to 80%, with low scores noted in aseptic technique (patient care Milestone 13 [PC13]) and written and verbal hand-off (interpersonal communications skills [ICS]2). In overall communication, 70% of interns demonstrated competency. In excess of 80% demonstrated competency in critical values interpretation (PC3), informed consent (PC9), pain assessment (PC11), and geriatric functional assessment (PC3). On direct observation, almost all Level 1 milestones were achieved (93% to 100%); however, only 78% of interns achieved competency in pharmacotherapy (PC5). On global monthly evaluations, all interns met Level 1 milestones.

CONCLUSIONS:
A multisource assessment of EM milestones is feasible and useful to determine Level 1 milestones achievement for incoming interns. A structured assessment program, used in conjunction with more traditional forms of evaluation such as global monthly evaluations and direct observation, is useful for identifying deficits in new trainees and may be able inform the creation of early intervention programs.
Operationalizing the Internal Medicine Milestones—An Early Status Report


BACKGROUND:
The internal medicine milestones were developed to advance outcomes-based residency training and will play an important role in the next accreditation system.

INNOVATION:
As an element of our program's participation in the internal medicine educational innovations project, we implemented a milestones-based evaluation process in our general medicine and pulmonary-critical care rotations on July 1, 2010.

MEASURES:
Outcomes assessed included survey-rated acceptability to participating faculty, residents, and clinical competency committee members.

RESULTS:
Faculty and residents agreed that the milestones promoted a common understanding of what knowledge, skills, and attitudes should be displayed at particular points in residents' professional development and enhanced evaluators' ability to provide specific performance feedback. Most residents and faculty members agreed that the milestones promoted fairness and uniformity in the evaluation process. Clinical competency committee members agreed the milestones improved the quality of information available for deliberations and resulted in more uniform promotion standards. Faculty rated the use of too many milestones per form/tool at a mean of 7.3 (where 1 was minimally problematic, and 10 was maximally problematic) and the potential for evaluator fatigue (mean, 8.2) as the most significant challenges to the use of milestones. Eight of 12 faculty members would recommend milestones in other programs; 4 were uncertain.

CONCLUSIONS:
Despite logistical challenges, educators and trainees found that milestones promoted a common understanding of what knowledge, skills and attitudes should be displayed at particular stages of training; permitted greater specificity in performance feedback; and enhanced uniformity and fairness in promotion decisions.
The Pediatrics Milestones: Initial Evidence for their Use as Learning Road Maps for Residents


OBJECTIVE:
As the next step in competency-based medical education, the Pediatrics Milestone Project seeks to provide a learner-centered approach to training and assessment. To help accomplish this goal, this study sought to determine how pediatric residents understand, interpret, and respond to the Pediatrics Milestones.

METHODS:
Cognitive interviews with 48 pediatric residents from all training levels at 2 training programs were conducted. Each participant reviewed one Pediatrics Milestone document (PMD). Eight total Pediatrics Milestones, chosen for their range of complexity, length, competency domain, and primary author, were included in this study. Six residents, 2 from each year of residency training, reviewed each PMD. Interviews were transcribed and coded using inductive methods, and codes were grouped into themes that emerged.

RESULTS:
Four major themes emerged through coding and analysis: 1) the participants' degree of understanding of the PMDs is sufficient, often deep; 2) the etiology of participants' understanding is rooted in their experiences; 3) there are qualities of the PMD that may contribute to or detract from understanding; and 4) participants apply their understanding by noting the PMD describes a developmental progression that can provide a road map for learning. Additionally, we learned that residents are generally comfortable being placed in the middle of a series of developmental milestones. Two minor themes focusing on interest and practicality were also identified.

CONCLUSIONS:
This study provides initial evidence for the Pediatrics Milestones as learner-centered documents that can be used for orientation, education, formative feedback, and, ultimately, assessment.
Using Milestones to Guide Curriculum
Case-Based Simulation Empowering Pediatric Residents to Communicate about Diagnostic Uncertainty


ABSTRACT
Background Uncertainty is ubiquitous in medical practice. The Pediatrics Milestones from the Accreditation Council on Graduate Medical Education state that advanced learners should acknowledge and communicate about clinical uncertainty. If uncertainty is not acknowledged, patient care may suffer. There are no described curricula specifically aimed to improve learners' ability to acknowledge and discuss clinical uncertainty. We describe an educational intervention designed to fill this gap. Methods Second-year pediatric residents engaged in a two-phase simulation-based educational intervention designed to improve their ability to communicate about diagnostic uncertainty with patients and caregivers. In each phase, residents engaged in two simulated cases and debriefs. Performance was assessed after each simulated patient encounter using standardized metrics, along with learner perceptions of the experience. Results Residents' skills in communicating with patients and families about diagnostic uncertainty improved after this intervention (mean score post 3.84 vs. 3.28 pre on a five-point Likert scale, p<0.001). Residents rated the experience as relevant, challenging and positive. Conclusions This prospective study suggests that a simulation-based intervention was effective in improving resident physicians' skills in communicating about diagnostic uncertainty with patients and families. Further study is needed to determine how learners perform in real clinical environments.
Psychotherapy Competency Milestones: An Exploratory Pilot of CBT and Psychodynamic Psychotherapy Skills Acquisition in Junior Psychiatry Residents


OBJECTIVE:
Psychiatry residents train in Psychodynamic Psychotherapy and Cognitive Behavioral Therapy (CBT), evidence-supported treatments used in mental health care that can facilitate clinical reasoning, foster therapeutic alliances, and improve clinical outcomes. However, empirically derived milestones are needed to evaluate competency. This exploratory pilot examined changes over 1 year of training in junior psychiatry residents' competency milestone elements in Psychodynamic Psychotherapy and CBT.

METHODS:
Seventy-nine randomly selected audio-recorded sessions from differing phases of Psychodynamic Psychotherapy and CBT with five junior residents and ten patients were rated using the Psychotherapy Process Q-sort (PQS).

RESULTS:
In both treatments, patient engagement with attention to in-session emotions improved. In CBT, residents were directive, supported patients' self-efficacy, emphasized patients' accepting responsibility for their problems, discussed homework such as thought records, and focused on termination in the concluding sessions. In Psychodynamic Psychotherapy, residents attended to emotional arousal and linked patients' feelings or perceptions to past situations or behavior. Growth and hierarchical linear modeling differentiated these treatments, with CBT v. Psychodynamic adherence to PQS modality-specific ideal elements being 52% v. 19%.

CONCLUSION:
Teaching and observation using empirically derived observable psychotherapy practice behaviors is feasible and can be used to assess milestone elements for competency-based education of psychiatry trainees.
Introducing a Curriculum in Ethics and Professionalism for Dermatology Residencies


ABSTRACT
There is general agreement on what constitutes ethical reasoning and professional behavior, but standardized methods to teach these skills in dermatology residency are currently unavailable. We introduce a model curriculum designed to impart the knowledge and skills to meet the Accreditation Council for Graduate Medical Education Dermatology Milestones for Professionalism over a 3-year cycle.
Development of Hospice and Palliative Medicine Knowledge and Skills for Emergency Medicine Residents: Using the Accreditation Council for Graduate Medical Education Milestone Framework


OBJECTIVES:
Emergency medicine (EM) physicians commonly care for patients with serious life-limiting illness. Hospice and palliative medicine (HPM) is a subspecialty pathway of EM. Although a subspecialty level of practice requires additional training, primary-level skills of HPM such as effective communication and symptom management are part of routine clinical care and expected of EM residents. However, unlike EM residency curricula in disciplines like trauma and ultrasound, there is no nationally defined HPM curriculum for EM resident training. An expert consensus group was convened with the aim of defining content areas and competencies for HPM primary-level practice in the ED setting. Our overall objective was to develop HPM milestones within a competency framework that is relevant to the practice of EM.

METHODS:
The American College of Emergency Physicians Palliative Medicine Section assembled a committee that included academic EM faculty, community EM physicians, EM residents, and nurses, all with interest and expertise in curricular design and palliative medicine.

RESULTS:
The committee peer reviewed and assessed HPM content for validity and importance to EM residency training. A topic list was developed with three domains: provider skill set, clinical recognition of HPM needs, and logistic understanding related to HPM in the ED. The group also developed milestones in HPM-EM to identify relevant knowledge, skills, and behaviors using the framework modeled after the Accreditation Council for Graduate Medical Education (ACGME) EM milestones. This framework was chosen to make the product as user-friendly and familiar as possible to facilitate use by EM educators.

CONCLUSIONS:
Educators in EM residency programs now have access to HPM content areas and milestones relevant to EM practice that can be used for curriculum development in EM residency programs. The HPM-EM skills/competencies presented herein are structured in a familiar milestone framework that is modeled after the widely accepted ACGME EM milestones.
Long-Term Retention of Musculoskeletal Ultrasound Training during Residency


ABSTRACT
The Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Physical Medicine and Rehabilitation (ABPMR) developed milestones for evaluation of resident physicians that include proper musculoskeletal ultrasound (MSUS) examination of major joints. To date, there have been no published data demonstrating acquisition and retention of these skills and correlation with the milestone evaluation. The investigators developed and implemented a curriculum in musculoskeletal ultrasound examination for Physical Medicine and Rehabilitation (PM&R) residents at a large academic medical center. The investigators chose six joints for training and evaluation: ankle, elbow, hip, knee, shoulder and wrist/hand. The program included: 1) didactic lectures on anatomy and ultrasound technique; 2) peer-led demonstrations of the procedure on a standardized patient (SP); 3) individual practice on SPs; 4) faculty observation and feedback; 5) review sessions and additional practice; and, 6) assessment of skills in an objective structured clinical examination (OSCE). From 2013-2017, 30 PM&R residents were trained and evaluated. The results, based on OSCE scores, showed that the majority of residents achieved the appropriate level of competency for their year. A blended, standardized curriculum in MSUS instruction with assessment by an OSCE, can be used to evaluate MSUS skills, and can help align this education with residency milestones.
Advancing Simulation-Based Education in Pain Medicine


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) has recently implemented milestones and competencies as a framework for training fellows in Pain Medicine, but individual programs are left to create educational platforms and assessment tools that meet ACGME standards.

OBJECTIVES:
In this article, we discuss the concept of milestone-based competencies and the inherent challenges for implementation in pain medicine. We consider simulation-based education (SBE) as a potential tool for the field to meet ACGME goals through advancing novel learning opportunities, engaging in clinically relevant scenarios, and mastering technical and nontechnical skills.

RESULTS:
The sparse literature on SBE in pain medicine is highlighted, and we describe our pilot experience, which exemplifies a nascent effort that encountered early difficulties in implementing and refining an SBE program.

CONCLUSIONS:
The many complexities in offering a sophisticated simulated pain curriculum that is valid, reliable, feasible, and acceptable to learners and teachers may only be overcome with coordinated and collaborative efforts among pain medicine training programs and governing institutions.
Trapped as a Group, Escape as a Team: Applying Gamification to Incorporate Teambuilding Skills through an 'Escape Room' Experience


ABSTRACT
Teamwork, a skill critical for quality patient care, is recognized as a core competency by the Accreditation Council for Graduate Medical Education (ACGME). To date, there is no consensus on how to effectively teach these skills in a forum that engages learners, immerses members in life-like activities, and builds both trust and rapport. Recreational 'Escape Rooms' have gained popularity in creating a life-like environment that rewards players for working together, solving puzzles, and completing successions of mindbending tasks in order to effectively 'escape the room' in the time allotted. In this regard, escape rooms share many parallels with the multitasking and teamwork that is essential for a successful emergency department (ED) shift. A pilot group of nine emergency medicine (EM) residents and one senior EM faculty member underwent a commercial escape room as part of a teambuilding exercise in January 2018. The escape room required participants to practice teamwork, communication, task delegation, and critical thinking to tackle waves of increasingly complex puzzles, ranging from hidden objects, physical object assembly (i.e., jigsaw puzzles), and symbol matching. Activities required members to recognize and utilize the collective experiences, skills, knowledge base, and physical abilities of the group. After the game, players underwent a structured 'game-master' debriefing facilitated by an employee of the commercial escape room; this was followed by a postevent survey facilitated by a faculty member, which focused on participants' feelings, experiences, and problem-solving techniques. Escape rooms afford learners the opportunity to engage in an activity that rewards teamwork and effective leadership through experiences that directly link to specific ACGME milestones and educational learning theories. EM participants were engaged in the activity and felt that the escape room reproduced an environment analogous to the ED. The debriefing that followed the activity provided a satisfactory conclusion to the experience; but learners preferred a more organized debriefing format that provided them with constructive and specific feedback on their performance.
A Competency-Based Simulation Curriculum for Surgical Resident Trauma Resuscitation Skills


BACKGROUND:
Evidence-based curricula for nonprocedural simulation training in general surgery are lacking. Residency programs are required to implement simulation training despite this shortcoming. The goal of this project was the development of a simulation curriculum that measurably improves milestone performance and replaces traditional experienced-based training with a competency-based model.

MATERIALS AND METHODS:
SimMan 3G® (Laerdal Medical, Wappingers Falls, NY, USA) was utilized for simulation. Needs assessment targeted trauma and shock resuscitation. Scenario design applied deliberate practice methodology. Learner performance data included items such as identification of shock physiology, resuscitation products used, volume delivered, use of resuscitation end-points, and knowledge of massive transfusion. Characteristics essential for a successful program were tabulated.

RESULTS:
Forty-eight residents in postgraduate year (PGY) 2-5 participated representing 100% of the 48 eligible for the training. Senior residents (PGY 4 and 5) demonstrated near universal improvement. Junior residents (PGY 2 and 3) improved in some areas but showed more skill decay between sessions. Overall, milestone performance improved with each training session, and resident feedback was universally positive.

CONCLUSIONS:
This prototype curriculum improved surgical resident competency in shock resuscitation in a simulated patient care environment. It can be modified to accommodate centers with fewer resources and can be implemented by clinical faculty. The essential characteristics of a successful program are identified.
Creation and Evaluation of a Laboratory Administration Curriculum for Pathology Residents


OBJECTIVES:
A clinical laboratory management (CLM) curriculum that can objectively assess the Accreditation Council for Graduate Medical Education pathology systems-based practice milestones and can provide consistent resident training across institutions is needed.

METHODS:
Faculty at Emory University created a curriculum that consists of assay verification exercises and interactive, case-based online modules. Beta testing was done at Emory University and Johns Hopkins. Residents were required to obtain a score of more than 80% in the online modules to achieve levels 3 to 4 in the milestones. In addition, residents shadowed a laboratory director, performed an inspection of a laboratory section, and completed training in human subjects research and test utilization.

RESULTS:
Fourteen residents took and evaluated the laboratory administration curriculum. The printed certificates from the modules were used for objective faculty evaluation of mastery of concepts. Of all the activities the residents performed during the rotation, the online modules were ranked most helpful by all residents. A 25-question knowledge assessment was performed before and after the rotation and showed an average increase of 8 points (P = .0001).

CONCLUSIONS:
The multimodal CLM training described here is an easily adoptable, objective system for teaching CLM. It was well liked by residents and provided an objective measurement of mastery of concepts for faculty.
Development of a Global Health Milestones Tool for Learners in Emergency Medicine: A Pilot Project


OBJECTIVES:
In medical education and training, increasing numbers of institutions and learners are participating in global health experiences. Within the context of competency-based education and assessment methodologies, a standardized assessment tool may prove valuable to all of the aforementioned stakeholders. Milestones are now used as the standard for trainee assessment in graduate medical education. Thus, the development of a similar, milestone-based tool was undertaken, with learners in emergency medicine (EM) and global health in mind.

METHODS:
The Global Emergency Medicine Think Tank Education Working Group convened at the 2016 Society for Academic Medicine Annual Meeting in New Orleans, Louisiana. Using the Interprofessional Global Health Competencies published by the Consortium of Universities for Global Health's Education Committee as a foundation, the working group developed individual milestones based on the 11 stated domains. An iterative review process was implemented by teams focused on each domain to develop a final product.

RESULTS:
Milestones were developed in each of the 11 domains, with five competency levels for each domain. Specific learning resources were identified for each competency level and assessment methodologies were aligned with the milestones framework. The Global Health Milestones Tool for learners in EM is designed for continuous usage by learners and mentors across a career.

CONCLUSIONS:
This Global Health Milestones Tool for learners in EM may prove valuable to numerous stakeholders. The next steps include a formalized pilot program for testing the tool's validity and usability across training programs, as well as an assessment of perceived utility and applicability by collaborating colleagues working in training sites abroad.
A Taxonomy of Perioperative Surgical Learning: Trending Resident Skill Acquisition


BACKGROUND:
Resident and curriculum evaluation require tracking surgical resident operative performance, yet what and when to measure remains unclear.

METHODS:
From a multi-institutional database, we reviewed 611 resident/surgeon-paired assessments of ACGME Milestones and modified OPRS ratings for different cases and postgraduate years.

RESULTS:
Faculty Milestone ratings increased with each PGY (p=<0.001) and correlated with resident self- ratings (ICC = 0.83). Mean OPRS scores increased in small increments with substantial intra-year variability. Progression among individual OPRS subcategories was not apparent from more global analyses. Interestingly, male faculty offered lower ratings than female faculty.

CONCLUSIONS:
Milestones and modified mean OPRS ratings suggest residents are learning, yet lack sufficient discrimination for promotion or curricular analysis. Differential progression through OPRS subcategories suggests a taxonomy of surgical learning that can be tailored to focus on different skills at each point in the training continuum. The effect of faculty gender on resident ratings awaits further study.
Milestone Assessment of Minimally Invasive Surgery in Pediatric Urology Fellowship Programs


INTRODUCTION:
Minimally invasive surgery has become an important aspect of Pediatric Urology fellowship training. In 2014, the Accreditation Council for Graduate Medical Education published the Pediatric Urology Milestone Project as a metric of fellow proficiency in multiple facets of training, including laparoscopic/robotic procedures.

OBJECTIVE:
The present study assessed trends in minimally invasive surgery training and utilization of the Milestones among recent Pediatric Urology fellows.

STUDY DESIGN:
Using an electronic survey instrument, Pediatric Urology fellowship program directors and fellows who completed their clinical year in 2015 were surveyed. Participants were queried regarding familiarity with the Milestone Project, utilization of the Milestones, robotic/laparoscopic case volume and training experience, and perceived competency with robotic/laparoscopic surgery at the start and end of the fellowship clinical year according to Milestone criteria. Responses were accepted between August and November 2015.

RESULTS:
Surveys were distributed via e-mail to 35 fellows and 30 program directors. Sixteen fellows (46%) and 14 (47%) program directors responded. All fellows reported some robotic experience prior to fellowship, and 69% performed >50 robotic/laparoscopic surgeries during residency. Fellow robotic/laparoscopic case volume varied: three had 1-10 cases (19%), four had 11-20 cases (25%), and nine had >20 cases (56%). Supplementary or robotic training modalities included simulation (9), animal models (6), surgical videos (7), and courses (2). Comparison of beginning and end of fellowship robotic/laparoscopic Milestone assessment (Summary Fig.) revealed scores of <3 in (10) 62% of fellow self-assessments and 10 (75%) of program director assessments. End of training Milestone scores >4 were seen in 12 (75%) of fellow self-assessment and eight (57%) of program director assessments.

DISCUSSION:
An improvement in robotic/laparoscopic Milestone scores by both fellow self-assessment and program director assessment was observed during the course of training; however, 43% of program directors rated their fellow below the graduation target of a Milestone score of 4.

CONCLUSION:
The best ways to teach minimally invasive surgery in fellowship training must be critically considered.
Practical Implications for an Effective Radiology Residency Quality Improvement Program for Milestone Assessment


ABSTRACT
Utilization of a radiology resident-specific quality improvement (QI) program and curriculum based on the Accreditation Council for Graduate Medical Education (ACGME) milestones can enable a program's assessment of the systems-based practice component and prepare residents for QI implementation post graduation. This article outlines the development process, curriculum, QI committee formation, and resident QI project requirements of one institution's designated radiology resident QI program. A method of mapping the curriculum to the ACGME milestones and assessment of resident competence by postgraduate year level is provided. Sample projects, challenges to success, and lessons learned are also described. Survey data of current trainees and alumni about the program reveal that the majority of residents and alumni responders valued the QI curriculum and felt comfortable with principles and understanding of QI. The most highly valued aspect of the program was the utilization of a resident education committee. The majority of alumni responders felt the residency quality curriculum improved understanding of QI, assisted with preparation for the American Board of Radiology examination, and prepared them for QI in their careers. In addition to the survey results, outcomes of resident project completion and resident scholarly activity in QI are evidence of the success of this program. It is hoped that this description of our experiences with a radiology resident QI program, in accordance with the ACGME milestones, may facilitate the development of successful QI programs in other diagnostic radiology residencies.
New Roadmap for the Journey from Internist to Rheumatologist


**OBJECTIVE:**
Measurement is necessary to gauge improvement. US training programs have not previously used shared standards to assess trainees’ mastery of the knowledge, skills, and attitudes necessary to practice rheumatology competently. In 2014, the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System began requiring semiannual evaluation of all medicine subspecialty fellows on 23 internal medicine subspecialty reporting milestones. Since these reporting milestones are not subspecialty specific, rheumatology curricular milestones were needed to guide rheumatology fellowship training programs and fellows on the training journey from internist to rheumatologist.

**METHODS:**
Rheumatology curricular milestones were collaboratively composed by expanding the internal medicine reporting milestones to delineate the specific targets of rheumatology fellowship training within 6 ACGME core competencies. The 2006 American College of Rheumatology core curriculum for rheumatology training programs was updated.

**RESULTS:**
A total of 80 rheumatology curricular milestones were created, defining progressive learning through training; most focus on patient care and medical knowledge. The core curriculum update incorporates the new curricular milestones and rheumatology entrustable professional activities.

**CONCLUSION:**
Rheumatology curricular milestones are now available for implementation by rheumatology fellowship training programs, providing a clear roadmap for specific training goals and a guide to track each fellow’s achievement over a 2-year training period. The comprehensive core curriculum delineates the essential breadth of knowledge, skills, and attitudes that define rheumatology, and provides a guide for educational activities during fellowship training. These guiding documents are now used to train and assess fellows as they prepare for independent rheumatology practice as the next generation of rheumatologists.
Making Milestones: Development and Implementation of a Formal Socioeconomic Curriculum in a Neurosurgical Residency Training Program


BACKGROUND:
Improved training in the socioeconomic aspects of medicine is a priority of the Accreditation Council for Graduate Medical Education and the American Board of Neurological Surgeons. There is evidence that young neurosurgeons feel ill equipped in these areas and that additional education would improve patient care.

OBJECTIVE:
To present our experience with the introduction of a succinct but formal socioeconomic training course to the residency curriculum at our institution.

METHODS:
A monthly series of twelve 1-hour interactive modules was designed to address the pertinent Accreditation Council for Graduate Medical Education-American Board of Neurological Surgeons outcomes-based educational milestones. Slide-based lectures provided a comprehensive overview of social, legal, and business issues, and a monthly forum for open discussion allowed residents to draw on their applied experience. Residents took a 20-question pre- and postcourse knowledge assessment, as well as feedback surveys at 6 and 12 months.

RESULTS:
Residents were able to participate in the lectures, with an overall attendance rate of 91%. Residents felt that the course goals and objectives were well defined and communicated (4.88/5) and rated highly the content, quality, and relevance of the lectures (4.94/5). Performance on the knowledge assessment improved from 58% to 66%.

CONCLUSION:
Our experience demonstrates the feasibility of including a formal socioeconomic course in neurosurgical residency training with positive resident feedback and achievement of outcomes-based milestones. Extension to a 2-year curriculum cycle may allow the course to cover more material without compromising other residency training goals. Online modules should also be explored to allow for wider and more flexible participation.
What's New in 10 Years? A Revised Cardiothoracic Curriculum for Diagnostic Radiology Residency with Goals and Objectives Related to General Competencies


ABSTRACT
This is a cardiothoracic curriculum document for radiology residents meant to serve not only as a study guide for radiology residents but also as a teaching and curriculum reference for radiology educators and radiology residency program directors. This document represents a revision of a cardiothoracic radiology resident curriculum that was published 10 years ago in Academic Radiology. The sections that have been significantly revised, expanded, or added are (1) lung cancer screening, (2) lung cancer genomic profiling, (3) lung adenocarcinoma revised nomenclature, (4) lung biopsy technique, (5) nonvascular thoracic magnetic resonance, (6) updates to the idiopathic interstitial pneumonias, (7) cardiac computed tomography updates, (8) cardiac magnetic resonance updates, and (9) new and emerging techniques in cardiothoracic imaging. This curriculum was written and endorsed by the Education Committee of the Society of Thoracic Radiology. This curriculum operates in conjunction with the Accreditation Council for Graduate Medical Education (ACGME) milestones project that serves as a framework for semiannual evaluation of resident physicians as they progress through their training in an ACGME-accredited residency or fellowship programs. This cardiothoracic curriculum document is meant to serve not only as a more detailed guide for radiology trainees, educators, and program directors but also complementary to and guided by the ACGME milestones.
Have First-Year Emergency Medicine Residents Achieved Level 1 on Care-Based Milestones?


BACKGROUND:
Residents in Accreditation Council for Graduate Medical Education accredited emergency medicine (EM) residencies were assessed on 23 educational milestones to capture their progression from medical student level (Level 1) to that of an EM attending physician (Level 5). Level 1 was conceptualized to be at the level of an incoming postgraduate year (PGY)-1 resident; however, this has not been confirmed.

OBJECTIVES:
Our primary objective in this study was to assess incoming PGY-1 residents to determine what percentage achieved Level 1 for the 8 emergency department (ED) patient care-based milestones (PC 1-8), as assessed by faculty. Secondary objectives involved assessing what percentage of residents had achieved Level 1 as assessed by themselves, and finally, we calculated the absolute differences between self- and faculty assessments.

METHODS:
Incoming PGY-1 residents at 4 EM residencies were assessed by faculty and themselves during their first month of residency. Performance anchors were adapted from ACGME milestones.

RESULTS:
Forty-one residents from 4 programs were included. The percentage of residents who achieved Level 1 for each subcompetency on faculty assessment ranged from 20% to 73%, and on self-assessment from 34% to 92%. The majority did not achieve Level 1 on faculty assessment of milestones PC-2, PC-3, PC-5a, and PC-6, and on self-assessment of PC-3 and PC-5a. Self-assessment was higher than faculty assessment for PC-2, PC-5b, and PC-6.

CONCLUSIONS:
Less than 75% of PGY-1 residents achieved Level 1 for ED care-based milestones. The majority did not achieve Level 1 on 4 milestones. Self-assessments were higher than faculty assessments for several milestones.
Gearing Up for Milestones in Surgery: Will Simulation Play a Role?


BACKGROUND:
The Consortium of American College of Surgeons-Accredited Education Institutes was created to promote patient safety through the use of simulation, develop new education and technologies, identify best practices, and encourage research and collaboration.

METHODS:
During the 7th Annual Meeting of the Consortium, leaders from a variety of specialties discussed how simulation is playing a role in the assessment of resident performance within the context of the Milestones of the Accreditation Council for Graduate Medical Education as part of the Next Accreditation System.

CONCLUSION:
This report presents experiences from several viewpoints and supports the utility of simulation for this purpose.
INTRODUCTION:
Incorporation of the 6 ACGME core competencies into surgical training has proven a considerable challenge particularly for the two primarily behavioral competencies, professionalism and interpersonal and communication skills. We report on experience with two specific interventions to foster the teaching and continuous evaluation of these competencies for neurosurgery residents.

MATERIAL AND METHODS:
In 2010, the Society of Neurological Surgeons (SNS) organized the first comprehensive Neurosurgery Boot Camp courses, held at six locations throughout the US and designed to assess and teach not only psychomotor skills but also components of all six Accreditation Council for Graduate Medical Education (ACGME) core competencies. These courses are comprised of various educational methodologies, including online material, faculty lectures, clinical scenario and group discussions, manual skills stations, and pre- and post-course assessments. Resident progress in each of the 6 ACGME competencies is now tracked using the neurosurgical Milestones, developed by the ACGME in collaboration with the SNS. In addition, the Milestones drafting group for neurosurgery has formulated a milestone-compatible evaluation system to directly populate Milestone reports. These evaluations utilize formative, summative, and 360-degree evaluations that are considered by a faculty core competency committee in finalizing milestones levels for each resident.

RESULTS:
Initial attendance at the 2010 Boot Camp course was 94% of the incoming resident class and in subsequent years, 100%. Pre- and post-course surveys demonstrated a significant and sustained increase in knowledge. The value of these courses has been recognized by the ACGME, which requires Boot Camp or equivalent participation prior to acting with indirect supervision during clinical activities. Neurosurgery was one of 7 early Milestone adopter specialties, beginning use in July, 2013. Early milestone data will establish benchmarks prior to utilization for "high stake" decisions such as promotion, graduation, and termination.

CONCLUSIONS:
The full impact of the neurosurgical Boot Camps and Milestones on residency education remains to be measured, although published data from the first years of the Boot Camp Courses demonstrate broad acceptance and early effectiveness. A complementary junior resident course has now been introduced for rising second-year residents. The Milestones compatible evaluation system now provides for multi-source formative and summative evaluation of neurosurgical residents within the new ACGME reporting rubric. Combined with consensus milestone assignments, this system provides new specificity and objectivity to resident evaluations. The correlation of milestone level assignments with other measurements of educational outcome awaits further study.
Entrustable Professional Activities and Curricular Milestones for Fellowship Training in Pulmonary and Critical Care Medicine: Report of a Multisociety Working Group


ABSTRACT
This article describes the curricular milestones and entrustable professional activities for trainees in pulmonary, critical care, or combined fellowship programs. Under the Next Accreditation System of the Accreditation Council for Graduate Medical Education (ACGME), curricular milestones compose the curriculum or learning objectives for training in these fields. Entrustable professional activities represent the outcomes of training, the activities that society and professional peers can expect fellowship graduates to be able to perform unsupervised. These curricular milestones and entrustable professional activities are the products of a consensus process from a multidisciplinary committee of medical educators representing the American College of Chest Physicians (CHEST), the American Thoracic Society, the Society of Critical Care Medicine, and the Association of Pulmonary and Critical Care Medicine Program Directors. After consensus was achieved using the Delphi process, the document was revised with input from the sponsoring societies and program directors. The resulting lists can serve as a roadmap and destination for trainees, program directors, and educators. Together with the reporting milestones, they will help mark trainees' progress in the mastery of the six ACGME core competencies of graduate medical education.
Teaching Ultrasound Professionalism


ABSTRACT
Professionalism is part of the milestone program instituted by the Accreditation Council for Graduate Medical Education and the American Board of Radiology. A unique feature of ultrasound professionalism is the relationship between the radiologist and the sonographer. Because this relationship is important for sonographic quality and ultimately patient outcome, residents should be trained to achieve an optimal professional relationship with sonographers. This article describes milestones for ultrasound professionalism and suggests methods of implementation.
Playing with Curricular Milestones in the Educational Sandbox: Q-Sort Results from an Internal Medicine Educational Collaborative


PURPOSE:
In competency-based medical education, the focus of assessment is on learner demonstration of predefined outcomes or competencies. One strategy being used in internal medicine (IM) is applying curricular milestones to assessment and reporting milestones to competence determination. The authors report a practical method for identifying sets of curricular milestones for assessment of a landmark, or a point where a resident can be entrusted with increased responsibility.

METHOD:
Thirteen IM residency programs joined in an educational collaborative to apply curricular milestones to training. The authors developed a game using Q-sort methodology to identify high-priority milestones for the landmark "Ready for indirect supervision in essential ambulatory care" (EsAMB). During May to December 2010, the programs' ambulatory faculty participated in the Q-sort game to prioritize 22 milestones for EsAMB. The authors analyzed the data to identify the top 8 milestones.

RESULTS:
In total, 149 faculty units (1-4 faculty each) participated. There was strong agreement on the top eight milestones; six had more than 92% agreement across programs, and five had 75% agreement across all faculty units. During the Q-sort game, faculty engaged in dynamic discussion about milestones and expressed interest in applying the game to other milestones and educational settings.

CONCLUSIONS:
The Q-sort game enabled diverse programs to prioritize curricular milestones with interprogram and interparticipant consistency. A Q-sort exercise is an engaging and playful way to address milestones in medical education and may provide a practical first step toward using milestones in the real-world educational setting.
Applying the Milestones in an Internal Medicine Residency Program Curriculum: A Foundation for Outcomes-Based Learner Assessment under the Next Accreditation System


ABSTRACT
In 2010, University of Kansas Medical Center internal medicine residency program leaders concluded that their competency-based curriculum and evaluation system was not sufficient to promote accurate assessment of learners' performance and needed revision to meet the requirements of the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System (NAS). Evaluations of learners seldom referenced existing curricular goals and objectives and reflected an "everyone is exceptional, no one is satisfactory" view. The authors identified the American Board of Internal Medicine and ACGME's Developmental Milestones for Internal Medicine Residency Training as a published standard for resident development. They incorporated the milestones into templates, a format that could be modified for individual rotations. A milestones-based curriculum for each postgraduate year of training and every rotation was then created, with input from educational leaders within each division in the Department of Internal Medicine and with the support of the graduate medical education office. In this article, the authors share their implementation process, which took approximately one year, and discuss their current work to create a documentation system for direct observation of entrustable professional activities, with the aim of providing guidance to other programs challenged with developing an outcomes-based curriculum and assessment system within the time frame of the NAS.
The Pediatrics Milestones: Initial Evidence for their Use as Learning Road Maps for Residents


OBJECTIVE:
As the next step in competency-based medical education, the Pediatrics Milestone Project seeks to provide a learner-centered approach to training and assessment. To help accomplish this goal, this study sought to determine how pediatric residents understand, interpret, and respond to the Pediatrics Milestones.

METHODS:
Cognitive interviews with 48 pediatric residents from all training levels at 2 training programs were conducted. Each participant reviewed one Pediatrics Milestone document (PMD). Eight total Pediatrics Milestones, chosen for their range of complexity, length, competency domain, and primary author, were included in this study. Six residents, 2 from each year of residency training, reviewed each PMD. Interviews were transcribed and coded using inductive methods, and codes were grouped into themes that emerged.

RESULTS:
Four major themes emerged through coding and analysis: 1) the participants' degree of understanding of the PMDs is sufficient, often deep; 2) the etiology of participants' understanding is rooted in their experiences; 3) there are qualities of the PMD that may contribute to or detract from understanding; and 4) participants apply their understanding by noting the PMD describes a developmental progression that can provide a road map for learning. Additionally, we learned that residents are generally comfortable being placed in the middle of a series of developmental milestones. Two minor themes focusing on interest and practicality were also identified.

CONCLUSIONS:
This study provides initial evidence for the Pediatrics Milestones as learner-centered documents that can be used for orientation, education, formative feedback, and, ultimately, assessment.
Rationale for Milestones
ABSTRACT
W. Edwards Deming, in his System of Profound Knowledge, asserts that leaders who wish to transform a system should understand four essential elements: appreciation for a system, theory of knowledge, knowledge about variation, and psychology. The Accreditation Council for Graduate Medical Education (ACGME) introduced the milestones program as a part of the Next Accreditation System to create developmental language for the six core competencies and facilitate programmatic assessment within graduate medical education systems. Viewed through Deming's lens, the ACGME can be seen as the steward of a large system, with everyone who provides assessment data as workers in that system. The authors use Deming's framework to illustrate the working components of the assessment system of the University of Cincinnati College of Medicine's internal medicine residency program and draw parallels to the macrocosm of graduate medical education. Successes and failures in transforming resident assessment can be understood and predicted by identifying the system and its aims, turning information into knowledge, developing an understanding of variation, and appreciating the psychology of motivation of participants. The authors offer insights from their experience for educational leaders who wish to apply Deming's elements to their own assessment systems, with questions to explore, pitfalls to avoid, and practical approaches in doing this type of work.
Competencies, Milestones, and a Level of Supervision Scale for Entrustable Professional Activities for Scholarship


PROBLEM:
Scholarship is an important element of both undergraduate and graduate medical education, and scholarly activity is required for all pediatric fellows. However, despite the creation of entrustable professional activities (EPAs) for scholarship, the specific progressive levels of performance and the appropriate level of supervision for a given performance level have not been defined. The authors developed competencies and milestones for the scholarship EPA to provide a framework for assessment across the continuum; a level of supervision scale was also developed.

APPROACH:
The Vitae Researcher Development Framework served as a template to create the competencies and milestones for the scholarship EPA. Beginning in September 2015 and using a modified Delphi approach, three distinct drafts were circulated to individuals with expertise in various types of scholarship until broad agreement was achieved. Then, in October 2016, the Steering Committee of the Subspecialty Pediatrics Investigator Network created a level of supervision scale, modeled after one it had previously developed.

OUTCOMES:
Eight competencies were identified as important in making entrustment decisions related to scholarship. For each competency, five milestone levels that span the continuum from novice to expert were created. A supervision scale with five progressive levels of entrustment was also created.

NEXT STEPS:
Next steps include a study to obtain validity evidence for the supervision scale and determine the correlation between milestone and supervision levels. These competencies, milestones, and supervision levels can potentially serve as a roadmap for trainees and junior faculty and also play a role in the assessment of physician-scientists.
Evaluating Surgical Residents' Patient-Centered Communication Skills: Practical Alternatives to the "Apprenticeship Model"


OBJECTIVES:
The Accreditation Council for Graduate Medical Education (ACGME) requires residency programs to assess communication skills and provide feedback to residents. We aimed to develop a feasible data collection process that generates objective clinical performance information to guide training activities, inform ACGME milestone evaluations, and validate assessment instruments.

DESIGN:
Residents care for patients in the surgical clinic and in the hospital, and participate in a communication curriculum providing practice with standardized patients (SPs). We measured perception of resident communication using the 14-item Communication Assessment Tool (CAT), collecting data from patients at the surgery clinic and surgical wards in the hospital, and from SP encounters during simulated training scenarios. We developed a handout of CAT example behaviors to guide patients completing the communication assessment.

SETTING: Independent academic medical center.

PARTICIPANTS: General surgery residents.

RESULTS:
The primary outcome is the percentage of total items patients rated "excellent;" we collected data on 24 of 25 residents. Outpatient evaluations resulted in significantly higher scores (mean 84.5% vs. 68.6%, p < 0.001), and female patients provided nearly statistically significantly higher ratings (mean 85.2% vs. 76.7%, p = 0.084). In multivariate analysis, after controlling for patient gender, visit reason, and race, (1) residents' CAT scores from SPs in simulation were independently associated with communication assessments in their concurrent patient population (p = 0.017), and (2) receiving CAT example instructions was associated with a lower percentage of excellent ratings by 9.3% (p = 0.047).

CONCLUSIONS:
Our data collection process provides a model for obtaining meaningful information about resident communication proficiency. CAT evaluations of surgical residents by the inpatient population had not previously been described in the literature; our results provide important insight into relationships between the evaluations provided by inpatients, clinic patients, and SPs in simulation. Our example behaviors guide shows promise for addressing a common concern, minimizing ceiling effects when measuring physician-patient communication.
Teaching and Assessing Professionalism in Radiology: Resources and Scholarly Opportunities to Contribute to Required Expectations


ABSTRACT
Teaching and assessing trainees' professionalism now represents an explicit expectation for Accreditation Council Graduate Medical Education-accredited radiology programs. Challenges to meeting this expectation include variability in defining the construct of professionalism; limits of traditional teaching and assessment methods, used for competencies historically more prominent in medical education, for professionalism; and emerging expectations for credible and feasible professionalism teaching and assessment practices in the current context of health-care training and practice. This article identifies promising teaching resources and methods that can be used strategically to augment traditional teaching of the cognitive basis for professionalism, including role modeling, case-based scenarios, debriefing, simulations, narrative medicine (storytelling), guided discussions, peer-assisted learning, and reflective practice. This article also summarizes assessment practices intended to promote learning, as well as to inform how and when to assess trainees as their professional identities develop over time, settings, and autonomous practice, particularly in terms of measurable behaviors. This includes assessment tools (including mini observations, critical incident reports, and appreciative inquiry) for authentic assessment in the workplace; engaging multiple sources (self-peer, other health professionals, and patients) in assessment; and intentional practices for trainees to take responsibility for seeking our actionable feedback and reflection. This article examines the emerging evidence of the feasibility and value added of assessment of medical competency milestones, including professionalism, coordinated by the Accreditation Council Graduate Medical Education in radiology and other medical specialties. Radiology has a strategic opportunity to contribute to scholarship and inform policies in professionalism teaching and assessment practices.
The American Society of Hematology and ASCO Curricular Milestones for Assessment of Fellows in Hematology/Oncology: Development, Reflection, and Next Steps


ABSTRACT
The American Society of Hematology (ASH)/ASCO Curricular Milestones is a tool for assessment and teaching for fellows in hematology/oncology. The expectations of the Next Accreditation System of the Accreditation Council of Graduate Medical Education (ACGME) was developed over years from the creation of the six core competencies in 1999 to the current data-driven outcomes-based system. The current internal medicine subspecialty milestones (ACGME reporting milestones) follow the general rubric of the general internal medicine milestones. The ASH/ASCO curricular milestones were developed from the foundational elements of the specialty, and they are interwoven with the ACGME reporting milestones. The 2017 ACGME Milestones Report shows that the milestones display progression in performance through clear anchors. Educational outcomes are available in many specialties. The internal medicine subspecialties have been given the opportunity to update the ACGME reporting milestones. The ACGME has acknowledged that these milestones may be different for each of the specialties. The program committees of ASH and ASCO agree that revision of the ACGME reporting milestones would decrease the overlap of domains, lack of clarity, and negative language that is present in version 1.0. ASH and ASCO are working with the ACGME and American Board of Internal Medicine (ABIM) to develop Curricular Milestones, version 2.0.
Using an Alumni Survey to Target Improvements in an Emergency Medicine Training Programme


INTRODUCTION:
The Accreditation Council for Graduate Medical Education (ACGME) is the governing body responsible for accrediting graduate medical training programme in the USA. The Emergency Medicine Milestones (EM-Milestones) were developed by the ACGME and American Board of Emergency Medicine as a guide and monitoring tool for the knowledge, skills, abilities and experiences to be acquired during training. Alumni surveys have been reported as a valuable resource for training programme to identify areas for improvement; however, there are few studies regarding programme improvement in emergency medicine. We aimed to use the EM-Milestones, adapted as an alumni self-assessment survey, to identify areas for training programme improvement.

METHODS:
This study was conducted at an urban, academic affiliated, community hospital in New York city with an emergency medicine training programme consisting of 30 residents over 3 years. Alumni of our emergency medicine training programme were sent an EM-Milestones-based self-assessment survey. Participants evaluated their ability in each EM-Milestones subcompetency on a Likert scale. Data were analysed using descriptive statistics.

RESULTS:
Response rate was 74% (69/93). Alumni reported achieving the target performance in 5/6 general competencies, with Systems-Based Practice falling below the target performance. The survey further identified 6/23 subcompetencies (Pharmacotherapy, Ultrasound, Wound Management, Patient Safety, Systems-Based Management and Technology) falling below the target performance level.

DISCUSSION:
Alumni self-evaluation of competence using the EM-Milestones provides valuable information concerning confidence to practice independently; these data, coupled with regular milestone evaluation of existing trainees, can identify problem areas and provide a blueprint for targeted programme improvement.
Are Surgical Milestone Assessments Predictive of In-Training Examination Scores?


OBJECTIVES:
With the recent utilization of Accreditation Council for Graduate Medical Education developmental milestones as part of resident evaluation, we sought to explore whether milestone-based ratings were associated with American Board of Surgery In-Training Examination (ABSITE) scores.

METHODS:
Mid-year milestone ratings were obtained from the Accreditation Council for Graduate Medical Education Accreditation Database System for years 2014, 2015, and 2016 for all postgraduate years 1-5 general surgery residents in our program and paired with ABSITE scores (n = 69) from January of the following year. Linear regression was used to assess predictor importance of milestones on both ABSITE percentage correct scores and ABSITE percentile scores.

RESULTS:
Minimal to small positive correlations were observed between milestones and ABSITE percentile scores (r = 0.09-0.25), while moderately large correlations were observed between milestones and percentage correct scores (r = 0.65-0.76). The Medical Knowledge 1 (MK1) milestone significantly predicted ABSITE percentage correct scores, and explained 60% of the variance (adjusted R² = 0.603). MK1 also significantly predicted ABSITE percentile scores, although weaker in magnitude, with MK1 explaining 20% of the variance (adjusted R² = 0.197). Postgraduate year level and other milestones were not influential predictors of ABSITE scores.

CONCLUSIONS:
The mid-year MK1 milestone rating was predictive of ABSITE scores and may serve as a useful marker for Clinical Competency Committees to identify residents who could benefit from additional support to prepare for the ABSITE, although given the small exploratory nature of this study, additional research is still needed.
Examination to Assess the Clinical Examination and Documentation of Spine Pathology among Orthopedic Residents


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) guidelines requires residency programs to teach and evaluate residents in six overarching "core competencies" and document progress through educational milestones. To assess the progress of orthopedic interns' skills in performing a history, physical examination, and documentation of the encounter for a standardized patient with spinal stenosis, an objective structured clinical examination (OSCE) was conducted for 13 orthopedic intern residents, following a 1-month boot camp that included communications skills and curriculum in history and physical examination. Interns were objectively scored based on their performance of the physical examination, communication skills, completeness and accuracy of their electronic medical record (EMR), and their diagnostic conclusions gleaned from the patient encounter.

PURPOSE:
The purpose of this study was to meaningfully assess the clinical skills of orthopedic post-graduate year (PGY)-1 interns. The findings can be used to develop a standardized curriculum for documenting patient encounters and highlight common areas of weakness among orthopedic interns with regard to the spine history and physical examination and conducting complete and accurate clinical documentation.

STUDY SETTING:
A major orthopedic specialty hospital and academic medical center.

METHODS:
Thirteen PGY-1 orthopedic residents participated in the OSCE with the same standardized patient presenting with symptoms and radiographs consistent with spinal stenosis. Videos of the encounters were independently viewed and objectively evaluated by one investigator in the study. This evaluation focused on the completeness of the history and the performance and completion of the physical examination. The standardized patient evaluated the communication skills of each intern with a separate objective evaluation. Interns completed these same scoring guides to evaluate their own performance in history, physical examination, and communications skills. The interns' documentation in the EMR was then scored for completeness, internal consistency, and inaccuracies.

RESULTS:
The independent review revealed objective deficits in both the orthopedic interns' history and the physical examination, as well as highlighted trends of inaccurate and incomplete documentation in the corresponding medical record. Communication skills with the patient did not meet expectations. Further, interns tended to overscore themselves, especially with regard to their performance on the physical examination (p<.0005). Inconsistencies, omissions, and inaccuracies were common in the corresponding medical notes when compared with the events of the patient encounter. Nine of the 13 interns (69.2%) documented at least one finding that was not assessed or tested in the clinical encounter, and four of the 13 interns (30.8%) included inaccuracies in the medical record, which contradicted the information collected at the time of the encounter.

CONCLUSIONS:
The results of this study highlighted significant shortcomings in the completeness of the interns' spine history and physical examination, and the accuracy and completeness of their EMR note. The study
provides a valuable exercise for evaluating residents in a multifaceted, multi-milestone manner that more accurately documents residents' clinical strengths and weaknesses. The study demonstrates that orthopedic residents require further instruction on the complexities of the spinal examination. It validates a need for increased systemic support for improving resident documentation through comprehensive education and evaluation modules.
Forensic Psychiatry Milestones: Experience after 1 Year


OBJECTIVE:
In 1999, the Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties identified six core competencies for medical practice. In 2013, the milestones were introduced to demonstrate these educational outcomes across each specialty. This study represents the first examination of the sub-specialty Forensic Psychiatry Milestones.

METHODS:
Members of the Association of Directors of Forensic Psychiatry Fellowships were surveyed. Areas of inquiry included whether milestones assisted in identifying areas of deficiency in fellows or programs, whether the graduation milestones matched the goals of training, and what changes were planned, or had been made, based on their implementation.

RESULTS:
Twenty-six of 35 programs responded, for a response rate of 74%. The majority found the milestones somewhat or very useful, half found the graduation-level milestones matched the program's graduation goals, and a significant majority reported that the milestones assisted in identifying improvements, change, or intended change. In choosing terms to describe the milestones, however respondents chose a variety of negative or neutral terms, rather than positive ones.

CONCLUSIONS:
The milestones provided a standard mechanism for identifying areas for improvement and a common language to standardize practice. However, due to the variability across fellowship programs and the limitations of educational resources and time, implementation of the new ACGME requirement was characterized in largely negative terms. Recommendations for improvement included modification of the milestones themselves, flexibility in their implementation, and evidentiary support for their use.
Purpose: To identify common and overlapping themes among the interpersonal and communication skills (ICS), practice-based learning and improvement (PBLI), professionalism (PROF), and systems-based practice (SBP) milestones of the transitional year and 26 specialties.

Method: In May 2017, milestones were accessed from the Accreditation Council for Graduate Medical Education specialties website. A thematic analysis of the ICS, PBLI, PROF, and SBP milestones was performed to determine unique and common themes across these competencies and across specialties. Keywords from the common program requirements were initially applied as codes to the milestones. Codes were then grouped into common themes.

Results: Twenty-two themes were identified: 15 (68%) were unique to a given competency (3 related to ICS, 4 related to PBLI, 5 related to PROF, and 3 related to SBP), and 7 (32%) appeared in the milestones of more than one core competency. Eleven themes (50%) were used by 20 or more specialties, and 6 themes (27%) by 10 or fewer specialties. No theme was present across all specialties.

Conclusions: The ICS, PBLI, PROF, and SBP milestones contain multiple themes with areas of overlap among these four competencies and substantial variability across specialties. This variability may create differential expectations of residents across specialties, complicate faculty development, and make sharing assessment tools difficult. The thematic analysis provides important insights into how individual specialties interpret and operationalize the ICS, PBLI, PROF, and SBP competency domains and can inform future revisions of milestones to enable harmonization and shared understanding of these competencies across specialties where appropriate.
Impact of an Event Reporting System on Resident Complication Reporting in Plastic Surgery Training: Addressing an ACGME and Plastic Surgery Milestone Project Core Competency


BACKGROUND:
The Accreditation Council for Graduate Medical Education and Plastic Surgery Milestone Project has identified practice-based learning and improvement, which involves systematically analyzing current practices and implementing changes, as a core competency in residency education. In surgical care, complication reporting is an essential component of practice-based learning and improvement as complications are analyzed in morbidity and mortality conference for quality improvement. Unfortunately, current methods for capturing a comprehensive profile of complications may significantly underestimate the true occurrence of complications. Therefore, the objectives of this study are to evaluate an intervention for complication reporting and compare this to current practice, in a plastic surgery training program.

METHODS:
This is a pre-intervention and post-intervention study evaluating resident reporting of complications on a plastic surgery service. The intervention was an online event reporting system developed by department leadership and patient safety experts. The cohorts consisted of all patients undergoing surgery during two separate 3-month blocks bridged by an implementation period. A trained reviewer recorded complications, and this served as the reference standard. Fisher's exact test was used for binary comparisons.

RESULTS:
There were 32 complications detected in 219 patients from June to August of 2015 and 35 complications in 202 patients from October to December of 2015. The proportion of complications reported in the pre-intervention group was nine of 32 (28.1 percent). After the intervention, this significantly increased to 32 of 35 (91.4 percent) (p < 0.001).

CONCLUSION:
An intervention using an event reporting system, supported by departmental leadership, led to significant improvements in complication reporting by plastic surgery residents.
Competency-Based Medical Education and the Ghost of Kuhn: Reflections on the Messy and Meaningful Work of Transformation.


ABSTRACT
The transition, if not transformation, to outcomes-based medical education likely represents a paradigm shift struggling to be realized. Paradigm shifts are messy and difficult but ultimately meaningful if done successfully. This struggle has engendered tension and disagreements, with many of these disagreements cast as either-or polarities. There is little disagreement, however, that the health care system is not effectively achieving the triple aim for all patients. Much of the tension and polarity revolve around how more effectively to prepare students and residents to work in and help change a complex health care system. Competencies were an initial attempt to facilitate this shift by creating frameworks of essential abilities needed by physicians. However, implementation of competencies has proven to be difficult. Entrustable professional activities (EPAs) in undergraduate and graduate medical education and Milestones in graduate medical education are recent concepts being tried and studied as approaches to guide the shift to outcomes. Their primary purpose is to help facilitate implementation of an outcomes-based approach by creating shared mental models of the competencies, which in turn can help to improve curricula and assessment. Understanding whether and how EPAs and Milestones effectively facilitate the shift to outcomes has been and will continue to be an iterative and ongoing reflective process across the entire medical education community using lessons from implementation and complexity science. In this Invited Commentary, the author reflects on what got the community to this point and some sources of tension involved in the struggle to move to outcomes-based education.
Objective:
As a result of the Milestones Project, all Accreditation Council for Graduate Medical Education accredited training programs now use an evaluation framework based on outcomes in 6 core competencies. Despite their widespread use, the Milestones have not been broadly evaluated. This study sought to examine program director (PD) perceptions of the Milestones Project.

Design, Setting, and Participants:
A national survey of general surgery PDs distributed between January and March of 2016.

Results:
A total of 132 surgical PDs responded to the survey (60% response rate). Positive perceptions included value for education (55%) and evaluation of resident performance (58%), as well as ability of Milestones to provide unbiased feedback (55%) and to identify areas of resident deficiency (58%). Meanwhile, time input and the ability of Milestones to discriminate underperforming programs were less likely to be rated positively (25% and 21%, respectively). Half of PDs felt that the Milestones were an improvement over their previous evaluation system (55%).

Conclusions:
Using the Milestones as competency-based, developmental outcomes measures, surgical PDs reported perceived benefits for education and objectivity in the evaluation of resident performance. The overall response to the Milestones was generally favorable, and most PDs would not return to their previous evaluation systems. To improve future iterations of the Milestones, many PDs expressed a desire for customization of the Milestones’ content and structure to allow for programmatic differences.
Building a Framework of Entrustable Professional Activities, Supported by Competencies and Milestones, to Bridge the Educational Continuum


ABSTRACT
The transition to competency-based medical education (CBME) and adoption of the foundational domains of competence by the Accreditation Council for Graduate Medical Education, Association of American Medical Colleges (AAMC), and American Board of Medical Specialties' certification and maintenance of certification (MOC) programs provided an unprecedented opportunity for the pediatrics community to create a model of learning and assessment across the continuum. Two frameworks for assessment in CBME have been promoted: (1) entrustable professional activities (EPAs) and (2) milestones that define a developmental trajectory for individual competencies. EPAs are observable and measureable units of work that can be mapped to competencies and milestones critical to performing them safely and effectively. The pediatrics community integrated the two frameworks to create a potential pathway of learning and assessment across the continuum from undergraduate medical education (UME) to graduate medical education (GME) and from GME to practice. The authors briefly describe the evolution of the Pediatrics Milestone Project and the process for identifying EPAs for the specialty and subspecialties of pediatrics. The method of integrating EPAs with competencies and milestones through a mapping process is discussed, and an example is provided. The authors illustrate the alignment of the AAMC’s Core EPAs for Entering Residency with the general pediatrics EPAs and, in turn, the alignment of the latter with the subspecialty EPAs, thus helping build the bridge between UME and GME. The authors propose how assessment in GME, based on EPAs and milestones, can guide MOC to complete the bridge across the education continuum.
Initial Comparison of Resident and Attending Milestones Evaluations in Plastic Surgery.


BACKGROUND:
Graduate medical education has recently undergone a major archetypal shift toward competency-based evaluations of residents' performance. The implementation of the Milestones program by the Accreditation Council for Graduate Medical Education (ACGME) is a core component of the shift, designed to ensure uniformity in measuring residency knowledge using a series of specialty-specific achievements. This study evaluates the correlation between residents' self-evaluations and program directors' assessments of their performance.

METHODS:
The study population comprised 12 plastic surgery residents, ranging from postgraduate year 1 to postgraduate year 6, enrolled in an integrated residency program at a single institution.

RESULTS:
Overall, average attending scores were lower than average resident scores at all levels except postgraduate year 6. Correlation between resident and attending evaluations ranged from 0.417 to 0.957, with the correlation of average scores of Patient Care (0.854) and Medical Knowledge (0.816) Milestones significantly higher than those of professional skillsets (0.581). "Patient care, facial esthetics" was the Milestone with the lowest average scores from both groups. Residents scored themselves notably higher than their attendings' evaluations in Practice-based Learning and Improvement categories (+0.958) and notably lower in Medical Knowledge categories such as "Cosmetic Surgery, Trunk and Lower Extremities" (-0.375) and "Non-trauma hand" (-0.208). The total possible number of participants in this study was 12. The actual number of participants was 12 (100%).

CONCLUSIONS:
The remarkable range of correlations suggests that expectations for performance standards may vary widely between residents and program directors. Understanding gaps between expectations and performance is vital to inform current and future residents as the restructuring of the accreditation process continues.
New Roadmap for the Journey from Internist to Rheumatologist


OBJECTIVE:
Measurement is necessary to gauge improvement. US training programs have not previously used shared standards to assess trainees’ mastery of the knowledge, skills, and attitudes necessary to practice rheumatology competently. In 2014, the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System began requiring semiannual evaluation of all medicine subspecialty fellows on 23 internal medicine subspecialty reporting milestones. Since these reporting milestones are not subspecialty specific, rheumatology curricular milestones were needed to guide rheumatology fellowship training programs and fellows on the training journey from internist to rheumatologist.

METHODS:
Rheumatology curricular milestones were collaboratively composed by expanding the internal medicine reporting milestones to delineate the specific targets of rheumatology fellowship training within 6 ACGME core competencies. The 2006 American College of Rheumatology core curriculum for rheumatology training programs was updated.

RESULTS:
A total of 80 rheumatology curricular milestones were created, defining progressive learning through training; most focus on patient care and medical knowledge. The core curriculum update incorporates the new curricular milestones and rheumatology entrustable professional activities.

CONCLUSION:
Rheumatology curricular milestones are now available for implementation by rheumatology fellowship training programs, providing a clear roadmap for specific training goals and a guide to track each fellow’s achievement over a 2-year training period. The comprehensive core curriculum delineates the essential breadth of knowledge, skills, and attitudes that define rheumatology, and provides a guide for educational activities during fellowship training. These guiding documents are now used to train and assess fellows as they prepare for independent rheumatology practice as the next generation of rheumatologists.
ASDS Cosmetic Dermatologic Surgery Fellowship Milestones


BACKGROUND:
The American Council of Graduate Medical Education, which oversees much of postgraduate medical education in the United States, has championed the concept of "milestones," standard levels of achievement keyed to particular time points, to assess trainee performance during residency.

OBJECTIVE:
To develop a milestones document for the American Society for Dermatologic Surgery (ASDS) Cosmetic Dermatologic Surgery (CDS) fellowship program.

METHODS:
An ad hoc milestone drafting committee was convened that included members of the ASDS Accreditation Work Group and program directors of ASDS-approved Cosmetic Dermatologic Surgery (CDC) fellowship training programs. Draft milestones were circulated through email in multiple rounds until consensus was achieved.

RESULTS:
Thirteen milestones were developed in the 6 Accreditation Council for Graduate Medical Education (ACGME) competency areas, with 8 of these being patient-care milestones. Additional instructions for milestone administration more specific to the CDS fellowship than general ACGME instructions were also approved. Implementation of semiannual milestones was scheduled for the fellowship class entering in July 2018.

CONCLUSION:
Milestones are now available for CDS fellowship directors to implement in combination with other tools for fellow evaluation.
The Road Ahead in Education: Milestones for Geriatric Psychiatry Subspecialty Training


OBJECTIVE:
The Accreditation Council of Graduate Medical Education (ACGME) Milestone Project is the next step in a series of changes revamping the system of graduate medical education. In 2013 the ACGME completed the general psychiatry milestones. The ACGME then pursued creation of milestones for accredited psychiatric subspecialty fellowships. This article documents the work of the geriatric psychiatry subspecialty milestones work group. It reports the history and rationale supporting the milestones, the milestone development process, and the implications for geriatric psychiatry fellowship training.

METHODS:
In consultation with the American Association for Geriatric Psychiatry, the American Board of Psychiatry and Neurology, and the ACGME Psychiatry Residency Review Committee, the ACGME appointed a working group to create the geriatric psychiatry milestones using the general psychiatry milestones as a guide.

CONCLUSION:
The geriatric psychiatry milestones are the result of an iterative process resulting in the definition of the characteristics vital to a fellowship-trained geriatric psychiatrist. It is premature to assess their effect on psychiatric training. The true impact of the milestones will be determined as each training director uses the milestones to re-evaluate their program curriculum and the educational and clinical learning environment. The ACGME is currently collecting the information about the milestone performance of residents and fellows to further refine and determine how the milestones can best be used to assist programs in improving training.
Defining, Achieving, and Maintaining Competence in Cardiovascular Training and Practice


ABSTRACT

Patients, hospitals, insurers, and the public rely on competent physicians. The definition and documentation of competency in cardiovascular training and practice continues to evolve. New tools, such as the American College of Cardiology's in-training examination, restructured Core Cardiovascular Training Statement, curricular and lifelong learning competencies, and the Accreditation Council for Graduate Medical Education Milestones help define competent trainees and practitioners, and level the playing field. The American Board of Internal Medicine's Maintenance of Certification program is undergoing critical review, and a common vision of its future form and role are not yet clear. This paper explores present-day cardiovascular competency components, assessment tools, and strategies, and identifies challenges for the future.
What Is a Rheumatologist and How Do We Make One?


OBJECTIVE:
Graduate medical education is a critical time in the training of a rheumatologist, and purposeful evaluation of abilities during this time is essential for long-term success as an independent practitioner. The internal medicine subspecialties collectively developed a uniform set of reporting milestones by which trainees can be assessed and receive formative feedback, providing clarity of accomplishment as well as areas for improvement in training. Furthermore, the reporting milestones provide a schema for assessment and evaluation of fellows by supervisors. The internal medicine subspecialties were also tasked with considering entrustable professional activities (EPAs), which define the abilities of a subspecialty physician who has attained sufficient mastery of the field to be accountable to stakeholders and participate in independent practice. Although EPAs have been established for a few specialties, they had not yet been described for rheumatology. EPAs have value as descriptors of the comprehensive abilities, knowledge, and skills of a practicing rheumatologist. The rheumatology EPAs have a role in defining a specialist in rheumatology upon completion of training, and also represent the ways our specialty defines our abilities that are enduring throughout practice.

METHODS:
We describe the collaborative process of the development of both the subspecialty reporting milestones and the rheumatology EPAs. The reporting milestones evolved through discussions and collaborations among representatives from the Association of Specialty Professors, the Alliance for Academic Internal Medicine, the American Board of Internal Medicine, and the Accreditation Council for Graduate Medical Education. The EPAs were a product of deliberations by the Next Accreditation System (NAS) working group of the American College of Rheumatology (ACR) Committee on Rheumatology Training and Workforce Issues.

RESULTS:
Twenty-three subspecialty reporting milestones and 14 rheumatology EPAs were advanced and refined over the course of 3 subspecialty reporting milestone development summits and 3 ACR NAS working group meetings, respectively.

CONCLUSION:
The subspecialty reporting milestones and rheumatology EPAs presented here stipulate reasonable and measurable expectations for rheumatologists-in-training. Together, these tools aim to promote enrichment and greater accountability in the training of fellows. Additionally, the EPAs define, for all stakeholders, the expertise of a rheumatologist in practice.
"What Program Directors Think" III: Results of the 2014/2015 Annual Surveys of the Association of Program Directors in Radiology (APDR)


RATIONALE AND OBJECTIVES:
The Association of Program Directors in Radiology regularly surveys its members regarding issues of importance to support radiology residency programs and their directors.

MATERIALS AND METHODS:
This is an observational cross-sectional study using two Web-based surveys posed to the Association of Program Directors in Radiology membership in the fall of 2014 (49 items) and the spring of 2015 (46 items) on the subjects of importance to the members, including the Accreditation Council on Graduate Medical Education Milestones, the Non-Interpretative Skills Curriculum, the American Board of Radiology Core Examination, the effect of the new resident testing and program accreditation paradigms on training outcomes, the 2015 Residency Match, the Interventional Radiology/Diagnostic Radiology (IR/DR) Residency, and Program Director (PD)/Program Coordinator resources.

RESULTS:
Responses were collected electronically, results were tallied using Survey Monkey software, and qualitative responses were tabulated or summarized as comments. Findings were reported during the 63rd annual meeting of the Association of University Radiologists. The maximal response rate was 33% in the fall of 2014 and 36% in the spring of 2015.

CONCLUSIONS:
PDs believed that the radiology Milestones, now largely implemented, did not affect overall resident evaluation, was not reflective of resident experience, and actually made evaluation of residents more difficult. PDs also felt that although the American Board of Radiology oral examination had been a better test for clinical practice preparedness, their new residents knew at least as much as before. There was little evidence of recall reemergence. The radiology training community saw a drop in residency applicant quality as demonstrated by the United States Medical Licensing Examination scores and clinical rotation grades. Because the new IR/DR Residency positions were to be funded at the expense of the traditional DR positions, the majority of PDs expected a negative effect of the impending IR/DR match on their DR recruitment. PDs were in favor of a unified clinical radiology curriculum similar to the Radiological Society of North America online physics modules.
Recognition of the importance of informatics to the practice of pathology has surged. Training residents in pathology informatics has been a daunting task for most residency programs in the United States because faculty often lacks experience and training resources. Nevertheless, developing resident competence in informatics is essential for the future of pathology as a specialty.

OBJECTIVE:
To develop and deliver a pathology informatics curriculum and instructional framework that guides pathology residency programs in training residents in critical pathology informatics knowledge and skills, and meets Accreditation Council for Graduate Medical Education Informatics Milestones.

DESIGN:
The College of American Pathologists, Association of Pathology Chairs, and Association for Pathology Informatics formed a partnership and expert work group to identify critical pathology informatics training outcomes and to create a highly adaptable curriculum and instructional approach, supported by a multiyear change management strategy.

RESULTS:
Pathology Informatics Essentials for Residents (PIER) is a rigorous approach for educating all pathology residents in important pathology informatics knowledge and skills. PIER includes an instructional resource guide and toolkit for incorporating informatics training into residency programs that vary in needs, size, settings, and resources. PIER is available at http://www.apcprods.org/PIER (accessed April 6, 2016).

CONCLUSIONS:
PIER is an important contribution to informatics training in pathology residency programs. PIER introduces pathology trainees to broadly useful informatics concepts and tools that are relevant to practice. PIER provides residency program directors with a means to implement a standardized informatics training curriculum, to adapt the approach to local program needs, and to evaluate resident performance and progress over time.
The Problem Resident Behavior Guide: Strategies for Remediation


ABSTRACT

In 2012, the ACGME supplemented the core competencies with outcomes-based milestones for resident performance within the six competency domains. These milestones address the knowledge, skills, abilities, attitudes, and experiences that a resident is expected to progress through during the course of training. Even prior to the initiation of the milestones, there was a paucity of EM literature addressing the remediation of problem resident behaviors and there remain few readily accessible tools to aid in the implementation of a remediation plan. The goal of the "Problem Resident Behavior Guide" is to provide specific strategies for resident remediation based on deficiencies identified within the framework of the EM milestones. The "Problem Resident Behavior Guide" is a written instructional manual that provides concrete examples of remediation strategies to address specific milestone deficiencies. The more than 200 strategies stem from the experiences of the authors who have professional experience at three different academic hospitals and emergency medicine residency programs, supplemented by recommendations from educational leaders as well as utilization of valuable education adjuncts, such as focused simulation exercises, lecture preparation, and themed ED shifts. Most recommendations require active participation by the resident with guidance by faculty to achieve the remediation expectations. The ACGME outcomes-based milestones aid in the identification of deficiencies with regards to resident performance without providing recommendations on remediation. The Problem Resident Behavior Guide can therefore have a significant impact by filling in this gap.
Implementation of Competency-Based Medical Education: Are We Addressing the Concerns and Challenges?


CONTEXT:
Competency-based medical education (CBME) has emerged as a core strategy to educate and assess the next generation of physicians. Advantages of CBME include: a focus on outcomes and learner achievement; requirements for multifaceted assessment that embraces formative and summative approaches; support of a flexible, time-independent trajectory through the curriculum; and increased accountability to stakeholders with a shared set of expectations and a common language for education, assessment and regulation.

OBJECTIVES:
Despite the advantages of CBME, numerous concerns and challenges to the implementation of CBME frameworks have been described, including: increased administrative requirements; the need for faculty development; the lack of models for flexible curricula, and inconsistencies in terms and definitions. Additionally, there are concerns about reductionist approaches to assessment in CBME, lack of good assessments for some competencies, and whether CBME frameworks include domains of current importance. This study will outline these issues and discuss the responses of the medical education community.

METHODS:
The concerns and challenges expressed are primarily categorised as: (i) those related to practical, administrative and logistical challenges in implementing CBME frameworks, and (ii) those with more conceptual or theoretical bases. The responses of the education community to these issues are then summarised.

CONCLUSIONS:
The education community has begun to address the challenges involved in implementing CBME. Models and guidance exist to inform implementation strategies across the continuum of education, and focus on the more efficient use of resources and technology, and the use of milestones and entrustable professional activities-based frameworks. Inconsistencies in CBME definitions and frameworks remain a significant obstacle. Evolution in assessment approaches from in vitro task-based methods to in vivo integrated approaches is responsive to many of the theoretical and conceptual concerns about CBME, but much work remains to be done to bring rigour and quality to work-based assessment.
Interprofessional Collaboration Milestones: Advocating for Common Assessment Criteria in Graduate Medical Education


BACKGROUND: Milestone-based assessments of resident physicians inform critical decisions regarding resident competence and advancement. Thus, it is essential that milestone evaluations are based upon strong validity evidence and that consistent evaluation criteria are used across residency programs. A common approach to assessment of interprofessional collaboration milestones is particularly important since standardized measures of individual resident competence in interprofessional collaboration have not been established.

DISCUSSION: We propose that assessments of interprofessional collaboration in graduate medical education meet common criteria, namely, these assessments should: 1) measure competency of an individual resident, 2) occur in the context of an interprofessional team, 3) be ascertained via direct observation of the resident, 4) be performed in a real-world clinical practice setting (such as a hospital ward, outpatient clinic, or operating room). We present the evidence-based rationale for these criteria and cite examples of published assessment instruments that fulfill one or more of the criteria, however further research is needed to ensure fidelity of assessments. The proposed criteria may assist residency educators as they endeavor to provide robust and consistent assessments of interprofessional collaboration milestones.
Achieving the Desired Transformation: Thoughts on Next Steps for Outcomes-Based Medical Education


ABSTRACT
Since the introduction of the outcomes-based medical education (OBME) movement, progress toward implementation has been active but challenging. Much of the angst and criticism has been directed at the approaches to assessment that are associated with outcomes-based or competency frameworks, particularly defining the outcomes. In addition, these changes to graduate medical education (GME) are concomitant with major change in health care systems—specifically, changes to increase quality and safety while reducing cost. Every sector, from medical education to health care delivery and financing, is in the midst of substantial change and disruption. The recent release of the Institute of Medicine’s report on the financing and governance of GME highlights the urgent need to accelerate the transformation of medical education. One source of continued tension within the medical education community arises from the assumption that the much-needed increases in value and improvement in health care can be achieved by holding the current educational structures and architecture of learning in place while concomitantly withdrawing resources. The authors of this Perspective seek to reframe the important and necessary debate surrounding the current challenges to implementing OBME. Building on recent change and service theories (e.g., Theory U and coproduction), they propose several areas of redirection, including reexamination of curricular models and greater involvement of learners, teachers, and regulators in cocreating new training models, to help facilitate the desired transformation in medical education.
Reflections on the First 2 Years of Milestone Implementation


BACKGROUND
The Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties (ABMS) collectively constitute the foundation of professional self-regulation in the United States. In February 1999, the 2 organizations approved 6 general competencies broadly relevant for all medical practice, followed by the official launch of the Outcomes Project in 2001. It was expected that the competencies would be an antidote to overspecification of accreditation standards, and that they would empower programs to create training programs grounded in meaningful outcomes in a developmental approach.

As many programs can attest, the implementation of outcomes-based (eg, competency-based) medical education has been challenging. One reason has been the difficulty in implementing the competencies in both curriculum and assessment. Program leaders lacked shared mental models within their own training programs, accompanied by a lack of shared understanding nationally within disciplines. It is important to remember that one of the thorny problems the milestones were intended to address was the sources of unwanted and unwarranted variability in educational and, by extension, clinical outcomes. In addition, the community cannot improve at scale what cannot be measured, and prior frames and approaches to measurement were insufficient and ineffective. A key goal for milestones thus is to help improve the state and quality of measurement through better assessment in graduate medical education to facilitate the improved outcomes everyone desires.

Approximately 10 years ago, conversations began on how to more effectively and meaningfully operationalize the competencies to help improve the design of residency and fellowship programs through the use of a developmental framework. In parallel, the ACGME began to explore mechanisms to move the accreditation system to a focus on outcomes using a continuous quality improvement philosophy. Developmental milestones, using narratives to describe in more descriptive terms the professional trajectories of residents, were seen as a way to move the outcomes project forward. Starting in 2007, the disciplines of internal medicine, pediatrics, and surgery began to create developmental milestones for the 6 competencies.

Surgery would subsequently delay the development of their milestones focusing first on the SCORE curriculum. The ACGME began to restructure its accreditation processes in 2009, and soon after, milestone groups were constituted for all specialties. Milestone writing groups were cosponsored by the ACGME and the ABMS member certification boards. Early groups had significant latitude in developing their subcompetencies and milestones; specialties that started the process after 2010 used a standard template. Each milestone set was subjected to review by the educational community in the specialty. Box 1 provides an overview of the purposes of the milestones across key stakeholders, and figure 1 provides an example of a key driver diagram of milestones as an educational and clinical intervention. As figure 1 highlights, milestones can potentially trigger a number of drivers, or mechanisms, to help enable changes in residency and fellowship education.
Developing a Comprehensive Resident Education Evaluation System in the Era of Milestone Assessment


OBJECTIVES:
In an effort to move training programs toward competency-based education, the Accreditation Council for Graduate Medical Education (ACGME) introduced the Next Accreditation System (NAS), which organizes specific milestones regarding resident skills, knowledge, and abilities along a continuum. In order to foster innovation and creativity, the ACGME has provided programs with minimal guidelines regarding the optimal way to approach these milestones.

METHODS:
The education team at UT Southwestern embraced the milestones and developed a process in which performance assessment methods were critically evaluated, mapped onto an extrapolated performance list corresponding to the areas required by the ACGME milestones, and filled gaps in the previous system by modifying evaluation tools and creating new program components.

RESULTS:
Although the authors are early in the evolution of applying the new milestones system, this approach has thus far allowed them to comprehensively evaluate the residents and the program in an efficient and effective fashion, with notable improvements compared to the prior approach.

CONCLUSIONS:
The authors hope that these experiences can inform others embarking upon similar journeys with the milestones.


BACKGROUND:
The Accreditation Council for Graduate Medical Education Next Accreditation System milestones were implemented for plastic surgery programs in July of 2014. Forward progress through the milestones is an indicator of trainee-appropriate development, whereas regression or stalling may indicate the need for concentrated, targeted training.

METHODS:
Online software at www.surveymonkey.com was used to create a survey about the program's approaches to milestones and was distributed to program directors and administrators of 96 Accreditation Council for Graduate Medical Education-approved plastic surgery programs.

RESULTS:
The authors had a 63.5 percent response rate (61 of 96 plastic surgery programs). Most programs report some level of readiness, only 22 percent feel completely prepared for the Next Accreditation System milestones, and only 23 percent are completely satisfied with their planned approach for compliance. Seventy-five percent of programs claim to be using some form of electronic tracking system. Programs plan to use multiple tools to capture and report milestone data. Most programs (44.4 percent) plan to administer evaluations at the end of each rotation. Over 70 percent of respondents believe that the milestones approach would improve the quality of resident training. However, programs were less than confident that their current compliance systems would live up to their full potential.

CONCLUSIONS:
The Next Accreditation System has been implemented nationwide for plastic surgery training programs. Milestone-based resident training is a new paradigm for residency training evaluation; programs are in the process of making this transition to find ways to make milestone data meaningful for faculty and residents.
Implementation of Nephrology Subspecialty Curricular Milestones


ABSTRACT
Beginning in the 2014-2015 training year, the US Accreditation Council for Graduate Medical Education (ACGME) required that nephrology Clinical Competency Committees assess fellows' progress toward 23 subcompetency "context nonspecific" internal medicine subspecialty milestones. Fellows' advancement toward the "ready for unsupervised practice" target milestone now is tracked in each of the 6 competencies: Patient Care, Medical Knowledge, Professionalism, Interpersonal Communication Skills, Practice-Based Learning and Improvement, and Systems-Based Practice. Nephrology program directors and subspecialty societies must define nephrology-specific "curricular milestones," mapped to the nonspecific ACGME milestones. Although the ACGME goal is to produce data that can discriminate between successful and underperforming training programs, the approach is at risk to produce biased, inaccurate, and unhelpful information. We map the ACGME internal medicine subspecialty milestones to our previously published nephrology-specific milestone schema and describe entrustable professional activities and other objective assessment tools that inform milestone decisions. Mapping our schema onto the ACGME subspecialty milestone reporting form allows comparison with the ACGME subspecialty milestones and the curricular milestones developed by the American Society of Nephrology Program Directors. Clinical Competency Committees may easily adapt and directly translate milestone decisions reached using our schema onto the ACGME internal medicine subspecialty competency milestone-reporting format.
Initial Validity Analysis of the Emergency Medicine Milestones


OBJECTIVES:
The Accreditation Council for Graduate Medical Education (ACGME) Milestones describe behavioral markers for the progressive acquisition of competencies during residency. As a key component of the Next Accreditation System, all residents are evaluated for the acquisition of specialty-specific Milestones. The objective was to determine the validity and reliability of the emergency medicine (EM) Milestones.

METHODS:
The ACGME and the American Board of Emergency Medicine performed this single-event observational study. The data included the initial EM Milestones performance ratings of all categorical EM residents submitted to the ACGME from October 31, 2013, to January 6, 2014. Mean performance ratings were determined for all 23 subcompetencies for every year of residency training. The internal consistency (reliability) of the Milestones was determined using a standardized Cronbach’s alpha coefficient.

Exploratory factor analysis was conducted to determine how the subcompetencies were interrelated.

RESULTS:
EM Milestone performance ratings were obtained on 100% of EM residents (n = 5,805) from 162 residency programs. The mean performance ratings of the aggregate and individual subcompetency scores showed discrimination between residency years, and the factor structure further supported the validity of the EM Milestones. The reliability was $a = 0.96$ within each year of training.

CONCLUSIONS:
The EM Milestones demonstrated validity and reliability as an assessment instrument for competency acquisition. EM residents can be assured that this evaluation process has demonstrated validity and reliability; faculty can be confident that the Milestones are psychometrically sound; and stakeholders can know that the Milestones are a nationally standardized, objective measure of specialty-specific competency acquisition.
Internal Medicine Residents’ Perspectives on Receiving Feedback in Milestone Format


BACKGROUND:
In contrast to historical feedback, which was vague or provided residents' numerical scores without clear meaning, milestone-based feedback is focused on specific knowledge, skills, and behaviors that define developmental trajectory. It was anticipated that residents would welcome the more specific and actionable feedback provided by the milestone framework, but this has not been studied.

OBJECTIVE:
We assessed internal medicine (IM) residents' perceptions of receiving feedback in the milestone framework, particularly assessing perception of the utility of milestone-based feedback compared to non-milestone-based feedback.

METHODS:
We surveyed a total of 510 IM residents from 7 institutions. Survey questions assessed resident perception of milestone feedback in identifying strengths, weaknesses, and trajectory of professional development. Postgraduate years 2 and 3 (PGY-2 and PGY-3) residents were asked to compare milestones with prior methods of feedback.

RESULTS:
Of 510 residents, 356 (69.8%) responded. Slightly less than half of the residents found milestone-based feedback "extremely useful" or "very useful" in identifying strengths (44%), weaknesses (43%), specific areas for improvement (45%), and appropriate education progress (48%). Few residents found such feedback "not very useful" or "not at all useful" in these domains. A total of 51% of PGY-2 and PGY-3 residents agreed that receiving milestone-based feedback was more helpful than previous forms of feedback.

CONCLUSIONS:
IM residents are aware of the concepts of milestones, and half of the residents surveyed found milestone feedback more helpful than previous forms of feedback. More work needs to be done to understand how milestone-based feedback could be delivered more effectively to enhance resident development.
Realizing the Promise of Competency-Based Medical Education


ABSTRACT
Competency-based medical education (CBME) places a premium on both educational and clinical outcomes. The Milestones component of the Next Accreditation System represents a fundamental change in medical education in the United States and is part of the drive to realize the full promise of CBME. The Milestones framework provides a descriptive blueprint in each specialty to guide curriculum development and assessment practices. From the beginning of the Outcomes project in 1999, the Accreditation Council for Graduate Medical Education and the larger medical education community recognized the importance of improving their approach to assessment. Work-based assessments, which rely heavily on the observations and judgments of clinical faculty, are central to a competency-based approach. The direct observation of learners and the provision of robust feedback have always been recognized as critical components of medical education, but CBME systems further elevate their importance. Without effective and frequent direct observation, coaching, and feedback, the full potential of CBME and the Milestones cannot be achieved.

Furthermore, simply using the Milestones as end-of-rotation evaluations to "check the box" to meet requirements undermines the intent of an outcomes-based accreditation system. In this Commentary, the author explores these challenges, addressing the concerns raised by Williams and colleagues in their Commentary. Meeting the assessment challenges of the Milestones will require a renewed commitment from institutions to meet the profession's "special obligations" to patients and learners. All stakeholders in graduate medical education must commit to a professional system of self-regulation to prepare highly competent physicians to fulfill this social contract.
The New Milestones: Do We Need to Take a Step Back to Go a Mile Forward?


ABSTRACT
The Milestones Project, like all previous systems and changes in graduate psychiatric education, for example, moving from 3 to 4 years of training or adopting six competency domains, has been devised without any supporting data and does not assess meaningful outcomes, such as improved patient outcomes. No evidence is presented that Milestones-based training will produce better psychiatrists. There is a path forward. First, replace unproven expert consensus with scientific and evidence-based approaches. Second, exchange endpoints that are easy to assess but uncorrelated with real world functioning (e.g., multiple-choice examinations) for outcomes that are meaningful and external to the training program (e.g., patient outcomes). Finally, to prevent possible waste, excess burden, or harm, no changes should be mandated until proven in prospective studies.
The Milestones for Psychosomatic Medicine Subspecialty Training


BACKGROUND:
The Accreditation Council of Graduate Medical Education Milestones project is a key element in the Next Accreditation System for graduate medical education. On completing the general psychiatry milestones in 2013, the Accreditation Council of Graduate Medical Education began the process of creating milestones for the accredited psychiatric subspecialties.

METHODS:
With consultation from the Academy of Psychosomatic Medicine, the Accreditation Council of Graduate Medical Education appointed a working group to create the psychosomatic medicine milestones, using the general psychiatry milestones as a starting point.

RESULTS:
This article represents a record of the work of this committee. It describes the history and rationale behind the milestones, the development process used by the working group, and the implications of these milestones on psychosomatic medicine fellowship training.

CONCLUSIONS:
The milestones, as presented in this article, will have an important influence on psychosomatic medicine training programs. The implications of these include changes in how fellowship programs will be reviewed and accredited by the Accreditation Council of Graduate Medical Education and changes in the process of assessment and feedback for fellows.
Two Cheers for Milestones


This editorial will explore the implementation of milestones across graduate medical education (GME) from 2 perspectives. The first is my perspective as a clinician, who often asks, “How do I make decisions with a patient when there isn't evidence to use as a guideline?” The second is my perspective as a department chair who asks a different question: “What resources are needed for milestone implementation?”
PURPOSE
Milestones for the assessment of residents in graduate medical education mark a change in our evaluation paradigms. The Accreditation Council for Graduate Medical Education has created milestones and defined them as significant points in development of a resident based on the 6 competencies. We propose that a similar approach be taken for resident assessment of teaching faculty. We believe this will establish parity and objectivity for faculty evaluation, provide improved data about attending surgeons’ teaching, and standardize faculty evaluations by residents.

METHODS
A small group of advanced surgery educators determined appropriate educational characteristics, resulting in creation of 11 milestones (Fig. 2) that were reviewed by faculty and residents. The residents have historically answered 16 questions, developed by our surgical education committee (Fig. 3), on a 5-point Likert score (never to very often). Three weeks after completing this Likert-type evaluation, the residents were asked to again evaluate attending faculty using the Faculty Milestones evaluation. The residents then completed a survey of 7 questions (scale of 1-9—disagree to strongly agree, neutral = 5), assessing the new milestones and compared with the previous Likert evaluation system.

RESULTS
Of 32 surgery residents, 13 completed the Likert evaluations (3760 data points) and 13 completed the milestones evaluations (1800 data points). The number completing both or neither is not known, as the responses are anonymous when used for faculty feedback. The Faculty Milestones attending physicians’ scores have far fewer top of range scores (21% vs 42%) and have a wider spread of data giving better indication of areas for improvement in teaching skills. The residents completed 17 surveys (116 responses) to evaluate the new milestones system. Surveys indicated that milestones were easier to use (average rating 6.13 ± 0.42 Standard Error (SE)), effective (6.82 ± 0.39) and efficient (6.11 ± 0.53), and more objective (6.69 ± 0.39/6.75 ± 0.38) than the Likert evaluations are. Average response was 6.47 ± 0.46 for overall satisfaction with the Faculty Milestones evaluation. More surveys were completed than evaluations, as all residents had an opportunity to review both evaluation systems.

CONCLUSIONS
Faculty Milestones are more objective in evaluating surgical faculty and mirror the new paradigm in resident evaluations. Residents found this was an easier, more effective, efficient, and objective evaluation of our faculty. Although our Faculty Milestones are designed for surgical educators, they are likely to be applicable with appropriate modifications to other medical educators as well.
Cytopathology Fellowship Milestones


ABSTRACT

The American Society of Cytopathology has provided guidelines for goals and objectives for cytopathology fellows. There are 90 Accreditation Council for Graduate Medical Education- accredited cytopathology fellowship training programs in the United States, each with its own unique curriculum designed to achieve these goals and objectives. The Accreditation Council for Graduate Medical Education cytopathology fellowship milestones were developed to ensure some uniformity in the outcomes of the various skill sets and competencies expected of a graduating cytopathology fellow. The rationale, development, and details of the cytopathology fellowship milestones are described herein.
Medical Specialty Boards Can Help Measure Graduate Medical Education Outcomes


ABSTRACT
U.S. graduate medical education (GME) training institutions are under increasing scrutiny to measure program outcomes as a demonstration of accountability for the sizeable funding they receive from the federal government. The Accreditation Council for Graduate Medical Education (ACGME) is a potential agent of measuring GME accountability but has no interaction with physicians after residency training is completed. American Board of Medical Specialty (ABMS) member boards interact with physicians throughout their careers through maintenance of certification (MOC) and are a potential source of valuable data on physician competency and quality of care, both of which could be used to measure GME accountability. The authors propose that ABMS boards and the ACGME deepen their existing relationship to better assess residency training outcomes. ABMS boards have a wealth of data on physicians collected as a by-product of MOC and business operations. Further, many ABMS boards collect practice demographics and scope-of-practice information through MOC enrollment surveys or recertification examination questionnaires. These data are potentially valuable in helping residencies know what their graduates are doing in practice. Part 4 of MOC generally involves assessment of the quality of care delivered in practice, and ABMS boards could share these deidentified data with the ACGME and residency programs to provide direct feedback on the practice outcomes of graduates. ABMS member boards and the ACGME should broaden their long-standing relationship to further develop shared roles and data-sharing mechanisms to better inform residencies and the public about GME training outcomes.
ABSTRACT
The Accreditation Council for Graduate Medical Education has partnered with member boards of the American Board of Medical Specialties to initiate the next steps in advancing competency-based assessment in residency programs. This initiative, known as the Milestone Project, is a paradigm shift from traditional assessment efforts and requires all pediatrics residency programs to report individual resident progression along a series of 4 to 5 developmental levels of performance, or milestones, for individual competencies every 6 months beginning in June 2014. The effort required to successfully make this shift is tremendous given the number of training programs, training institutions, and trainees. However, it holds great promise for achieving training outcomes that align with patient needs; developing a valid, reliable, and meaningful way to track residents' development; and providing trainees with a roadmap for learning. Recognizing the resources needed to implement this new system, the authors, all residency program leaders, provide their consensus view of the components necessary for implementing and sustaining this effort, including resource estimates for completing this work. The authors have identified 4 domains: (1) Program Review and Development of Stakeholders and Participants, (2) Assessment Methods and Validation, (3) Data and Assessment System Development, and (4) Summative Assessment and Feedback. This work can serve as a starting point and framework for collaboration with program, department, and institutional leaders to identify and garner necessary resources and plan for local and national efforts that will ensure successful transition to milestones-based assessment.
Progress toward Improved Leadership and Management Training in Pathology


CONTEXT:
Competency gaps in leadership and laboratory management skills continue to exist between what training programs deliver and what recent graduates and future employers expect. A number of recent surveys substantiate this. Interest in delivering content in these areas is challenged by time constraints, the presence of knowledgeable faculty role models, and the necessary importance placed on diagnostic skills development, which overshadows any priority trainees have toward developing these skills.

OBJECTIVE:
To describe the problem, the near-future horizon, the current solutions, and the recommendations for improving resident training in laboratory management.

DATA SOURCES:
The demands of new health care delivery models and the value being placed on these skills by the Pathology Milestones and Next Accreditation System initiative of the Accreditation Council for Graduate Medical Education for training programs emphasizes their importance. This initiative includes 6 milestone competencies in laboratory management. Organizations like the American Society for Clinical Pathology, the American Pathology Foundation, the College of American Pathologists, and the Association of Pathology Chairs Program Directors Section recognize these competencies and are working to create new tools for training programs to deploy.

CONCLUSIONS:
It is our recommendation that (1) every training program develop a formal educational strategy for management training, (2) greater opportunity and visibility be afforded for peer-reviewed publications on management topics in mainstream pathology literature, and (3) pathology milestones-oriented tools be developed to assist program directors and their trainees in developing this necessary knowledge and skills.
The Pathology Milestones and the Next Accreditation System


BACKGROUND:
In the late 1990s, the Accreditation Council for Graduate Medical Education developed the Outcomes Project and the 6 general competencies with the intent to improve the outcome of graduate medical education in the United States. The competencies were used as the basis for developing learning goals and objectives and tools to evaluate residents' performance. By the mid-2000s the stakeholders in resident education and the general public felt that the Outcomes Project had fallen short of expectations.

OBJECTIVE:
To develop a new evaluation method to track trainee progress throughout residency using benchmarks called milestones. A change in leadership at the Accreditation Council for Graduate Medical Education brought a new vision for the accreditation of training programs and a radically different approach to the evaluation of residents.

DATA SOURCES:
The Pathology Milestones Working Group reviewed examples of developing milestones in other specialties, the literature, and the Accreditation Council for Graduate Medical Education program requirements for pathology to develop pathology milestones. The pathology milestones are a set of objective descriptors for measuring progress in the development of competency in patient care, procedural skill sets, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice.

CONCLUSIONS:
The milestones provide a national standard for evaluation that will be used for the assessment of all residents in Accreditation Council for Graduate Medical Education-accredited pathology training programs.
ABSTRACT
Medical education shaped by the learning sciences can better serve medical students, residents, faculty, health care institutions, and patients. With increasing innovation in undergraduate and graduate medical education and more focused attention on educational principles and how people learn, this era of educational transformation offers promise. Principles manifest in "educational continuity" are informing changes in educational structures and venues and are enriching new discourse in educational pedagogy, assessment, and scholarship. The articles by Myhre and colleagues and Woloschuk and colleagues in this issue, along with mounting evidence preceding these works, should reassure that principle-driven innovation in medical education is not only possible but can be achieved safely. In this commentary, the authors draw from these works and the wider literature on longitudinal integrated educational design. They suggest that the confluences of movements for longitudinal integrated clerkships and entrustable professional activities open new possibilities for other educational and practice advancements in quality and safety. With the advent of competency-based education, explicit milestones, and improved assessment regimens, overseers will increasingly evaluate students, trainees, and other learners on their ability rather than relying solely on time spent in an activity. The authors suggest that, for such oversight to have the most value, assessors and learners need adequate oversight time, and redesign of educational models will serve this operational imperative. As education leaders are reassessing old medical school and training models, rotational blocks, and other barriers to progress, the authors explore the dynamic interplay between longitudinal integrated learning models and entrustment.
Early Feedback on the Use of the Internal Medicine Reporting Milestones in Assessment of Resident Performance


BACKGROUND:
The educational milestones were designed as a criterion-based framework for assessing resident progression on the 6 Accreditation Council for Graduate Medical Education competencies.

OBJECTIVE:
We obtained feedback on, and assessed the construct validity and perceived feasibility and utility of, draft Internal Medicine Milestones for Patient Care and Systems-Based Practice.

METHODS:
All participants in our mixed-methods study were members of competency committees in internal medicine residency programs. An initial survey assessed participant and program demographics; focus groups obtained feedback on the draft milestones and explored their perceived utility in resident assessment, and an exit survey elicited input on the value of the draft milestones in resident assessment. Surveys were tabulated using descriptive statistics. Conventional content analysis method was used to assess the focus group data.

RESULTS:
Thirty-four participants from 17 programs completed surveys and participated in 1 of 6 focus groups. Overall, the milestones were perceived as useful in formative and summative assessment of residents. Participants raised concerns about the length and complexity of some draft milestones and suggested specific changes. The focus groups also identified a need for faculty development. In the exit survey, most participants agreed that the Patient Care and Systems-Based Practice Milestones would help competency committees assess trainee progress toward independent practice.

CONCLUSIONS:
Draft reporting milestones for 2 competencies demonstrated significant construct validity in both the content and response process and the perceived utility for the assessment of resident performance. To ensure success, additional feedback from the internal medicine community and faculty development will be necessary.
The ACGME Milestone Project in Ophthalmology


ABSTRACT
The ACGME is moving towards the next generation of accreditation in the USA called the Next Accreditation System (NAS). The NAS is anticipated to reduce the burden on programs to comply with accreditation requirements; to produce meaningful, innovative, and continuous benchmark outcomes data; to use ongoing individual and programmatic milestones to judge performance; and ultimately to produce better trained residents, to improve the quality of care, to reduce health care costs and health care disparities, and to provide objective evidence to the public and other external stakeholders of the quality of graduate medical education across the specialties of medicine. We describe the ACGME milestone development process for ophthalmology. If successful, the NAS will benefit all programs by reducing the programmatic burden and paperwork; increasing accreditation cycle length; and improving all programs through formative and summative feedback.
ABSTRACT
The Accreditation Council for Graduate Medical Education (ACGME) has outlined its "Next Accreditation System" (NAS) that will focus on resident and residency outcome measurements. Emergency medicine (EM) is one of seven specialties that will implement the NAS beginning July 2013. All other specialties will follow in July 2014. A key component of the NAS is the development of assessable milestones, which are explicit accomplishments or behaviors that occur during the process of residency education. Milestones describe competencies more specifically and identify specialty-specific knowledge, skills, attitudes, and behaviors (KSABs) that can be used as outcome measures within the general competencies. The ACGME and the American Board of Emergency Medicine (ABEM) convened an EM milestone working group to develop the EM milestones. This article describes the development, use within the NAS, and challenges of the EM milestones.
Educational Milestone Development in the First 7 Specialties to Enter the Next Accreditation System


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) Outcome Project introduced six general competencies relevant to medical practice but fell short of its goal to create a robust assessment system that would allow program accreditation based on outcomes. In response, the ACGME, the specialty boards, and other stakeholders collaborated to develop educational milestones, observable steps in residents' professional development that describe progress from entry to graduation and beyond.

OBJECTIVES:
We summarize the development of the milestones, focusing on 7 specialties, moving to the next accreditation system in July 2013, and offer evidence of their validity.

METHODS:
Specialty work groups with broad representation used a 5-level developmental framework and incorporated information from literature reviews, specialty curricula, dialogue with constituents, and pilot testing.

RESULTS:
The work groups produced richly diverse sets of milestones that reflect the community's consideration of attributes of competence relevant to practice in the given specialty. Both their development process and the milestones themselves establish a validity argument, when contemporary views of validity for complex performance assessment are used.

CONCLUSIONS:
Initial evidence for validity emerges from the development processes and the resulting milestones. Further advancing a validity argument will require research on the use of milestone data in resident assessment and program accreditation.
Operationalizing the Internal Medicine Milestones – An Early Status Report


BACKGROUND:
The internal medicine milestones were developed to advance outcomes-based residency training and will play an important role in the next accreditation system.

INNOVATION:
As an element of our program’s participation in the internal medicine educational innovations project, we implemented a milestones-based evaluation process in our general medicine and pulmonary-critical care rotations on July 1, 2010.

MEASURES:
Outcomes assessed included survey-rated acceptability to participating faculty, residents, and clinical competency committee members.

RESULTS:
Faculty and residents agreed that the milestones promoted a common understanding of what knowledge, skills, and attitudes should be displayed at particular points in residents' professional development and enhanced evaluators' ability to provide specific performance feedback. Most residents and faculty members agreed that the milestones promoted fairness and uniformity in the evaluation process. Clinical competency committee members agreed the milestones improved the quality of information available for deliberations and resulted in more uniform promotion standards. Faculty rated the use of too many milestones per form/tool at a mean of 7.3 (where 1 was minimally problematic, and 10 was maximally problematic) and the potential for evaluator fatigue (mean, 8.2) as the most significant challenges to the use of milestones. Eight of 12 faculty members would recommend milestones in other programs; 4 were uncertain.

CONCLUSIONS:
Despite logistical challenges, educators and trainees found that milestones promoted a common understanding of what knowledge, skills and attitudes should be displayed at particular stages of training; permitted greater specificity in performance feedback; and enhanced uniformity and fairness in promotion decisions.
From Theory to Actual Practice: Creation and Application of Milestones in an Internal Medicine Residency Program, 2004-2010


BACKGROUND:
In the USA, the Accreditation Council of Graduate Medical Education, Educational Innovations Project is a partner in reshaping residency training to meet increasingly complex systems of health care delivery.

AIM:
We describe the creation and implementation of milestones as a vehicle for translating educational theory into practice in preparing residents to provide safe, autonomous patient care.

METHOD:
Six program faculty leaders, all with advanced medical education training, met in an iterative process of developing, implementing, and modifying milestones until a final set were vetted.

RESULTS:
We first formed the profile of a Master Internist. We then translated it into milestone language and implemented its integration across the program. Thirty-seven milestones were applied in all settings and rotations to reach explicit educational outcomes. We created three types of milestones: Progressive, build one on top of the other to mastery; additive, adding multiple behaviors together to culminate in mastery; and descriptive, using a proscribe set of complex, predetermined steps toward mastery.

CONCLUSIONS:
Using milestones, our program has enhanced an educational model into explicit, end of training goals. Milestone implementation has yielded positive results toward competency-based training and others may adapt our strategies in a similar effort.
Milestones for Apheresis Education


ABSTRACT
Milestones represent the essential knowledge, skills, and attitudes required for the practice of a medical discipline. Defining these milestones for each medical specialty has become a focus for the American Council of Graduate Medical Education (ACGME). Practitioners of Apheresis Medicine come from a variety of medical specialties making it challenging to establish the essential educational milestones for all. The American Society for Apheresis (ASFA) has an interest in promoting standards of excellence for Apheresis Medicine. ASFA's Physician's Curriculum Content Committee is a group of physician educators in the field of Apheresis Medicine, both donor and therapeutic, from across the United States, who have met regularly for several years to discuss the appropriate educational milestones in Apheresis training. The committee members teach residents and fellows from Pathology, Transfusion Medicine, Hematology/Oncology, Nephrology and other specialties. In this document, we have outlined the basic set of Apheresis milestones required in the ACGME defined competency areas of Patient Care and Medical Knowledge. We have also recommended methods of evaluation and estimated the time necessary for the acquisition of these cognitive and behavioral elements.
Trends in Milestones Data
Visuospatial Aptitude Testing Differentially Predicts Simulated Surgical Skill


OBJECTIVE:
To determine whether visuospatial perception (VSP) testing is correlated to simulated or intraoperative surgical performance as rated by the American College of Graduate Medical Education (ACGME) milestones.

DESIGN: (Canadian Task Force classification II-2).

SETTING: Two academic training institutions.

PARTICIPANTS:
Forty-one residents, including 19 from Brigham and Women's Hospital and 22 from the Mayo Clinic, from 3 different specialties: obstetrics and gynecology, general surgery, and urology.

INTERVENTION:
Participants underwent 3 different tests: visuospatial perception testing (VSP), Fundamentals of Laparoscopic Surgery (FLS) peg transfer, and da Vinci robotic simulation peg transfer. Surgical grading from the ACGME milestones tool was obtained for each participant. Demographic and background information was also collected, including specialty, year of training, previous experience with simulated skills, and surgical interest. Standard statistical analyses were performed using Student's t test, and correlations were determined using adjusted linear regression models.

MEASUREMENTS AND MAIN RESULTS:
In univariate analysis, Brigham and Women's Hospital and Mayo Clinic training programs differed in times and overall scores for both the FLS peg transfer and da Vinci robotic simulation peg transfer tests (p < .05 for all). In addition, type of residency training affected time and overall score on the robotic peg transfer test. Familiarity with tasks correlated with higher score and faster task completion (p = .05 for all except VSP score). There were no differences in VSP scores by program, specialty, or year of training. In adjusted linear regression modeling, VSP testing was correlated only to robotic peg transfer skills (average time, p = .006; overall score, p = .001). Milestones did not correlate to either VSP or surgical simulation testing.

CONCLUSION:
VSP score was correlated with robotic simulation skills, but not with FLS skills or ACGME milestones. This suggests that the ability of VSP score to predict competence differs between tasks. Therefore, further investigation of aptitude testing is needed, especially before its integration as an entry examination into a surgical subspecialty.
Timing of Milestone Competency Acquisition in Neurology Residency: What by When?


OBJECTIVE:
To determine the stage of training at which neurology residents should achieve individual elements of the Accreditation Council for Graduate Medical Education neurology Milestones and to examine the relationship between perceived importance of Milestones and the stage by which they should be achieved.

METHODS:
A modified Delphi technique was used to establish consensus postgraduate year (PGY) expectations for neurology Milestone competencies across 3 geographically and administratively distinct Mayo Clinic neurology residency programs. Timing expectations were examined for relationships to perceived importance of the individual Milestones and effects of participant characteristics.

RESULTS:
PGY expectations for neurology Milestone elements ranged from PGY 1.3 to PGY 4.1. Extent of rater educational seniority had no effect on PGY competency expectations. There was a moderate inverse relationship between perceived importance of the Milestone element and the PGY by which it should be achieved ($r_s = -0.74$, $p < 0.0001$).

CONCLUSIONS AND RELEVANCE:
Expectations for neurology Milestone competency acquisition can be measured and may help inform individual program design, educational expectations, and future Milestone design.
Predicting Performance of First-Year Residents: Correlations between Structured Interview, Licensure Exam, and Competency Scores in a Multi-Institutional Study


PURPOSE:
To determine whether scores on structured interview (SI) questions designed to measure non-cognitive competencies in physicians (1) predict subsequent first-year resident performance on Accreditation Council for Graduate Medical Education (ACGME) milestones and (2) add incremental validity over United States Medical Licensing Examination (USMLE) Step 1 and Step 2 Clinical Knowledge (CK) scores in predicting performance.

METHOD:
The authors developed 18 behavioral description questions to measure key non-cognitive competencies (e.g., teamwork). In 2013-2015, 14 programs (13 residency, 1 fellowship) from 6 institutions used subsets of these questions in their selection processes. The authors conducted analyses to determine the validity of SIs and USMLE scores in predicting first-year resident milestone performance in the ACGME’s core competency domains and overall.

RESULTS:
SI scores predicted mid-year and year-end overall performance (r = .18 and .19, respectively, P < .05) and year-end performance on patient care, interpersonal and communication skills, and professionalism competencies (r = .23, r = .22, and r = .20, respectively, p < .05). SI scores contributed incremental validity over USMLE scores in predicting year-end performance on patient care (ΔR = .05), interpersonal and communication skills (ΔR = .09), and professionalism (ΔR = .09; all P < .05). USMLE scores contributed incremental validity over SI scores in predicting year-end performance overall and on patient care and medical knowledge.

CONCLUSIONS:
SI scores predict first-year resident year-end performance in the interpersonal and communication skills, patient care, and professionalism competency domains. Future research should investigate whether SIs predict a range of clinically relevant outcomes.
Surgical Simulation: Markers of Proficiency


OBJECTIVE:
Surgical simulation has become an integral component of surgical training. Simulation proficiency determination has been traditionally based upon time to completion of various simulated tasks. We aimed to determine objective markers of proficiency in surgical simulation by comparing novel assessments with conventional evaluations of technical skill.

DESIGN:
Categorical general surgery residents completed 10 laparoscopic cholecystectomy modules using a high-fidelity simulator. We recorded and analyzed simulation task times, as well as number of hand movements, instrument path length, instrument acceleration, and participant affective engagement during each simulation. Comparisons were made to Objective Structured Assessment of Technical Skill (OSATS) and Accreditation Council for Graduate Medical Education Milestones, as well as previous laparoscopic experience, duration of laparoscopic cholecystectomies performed by participants, and postgraduate year. Comparisons were also made to Fundamentals of Laparoscopic Surgery task times. Spearman's rho was utilized for comparisons, significance set at >0.50.

SETTING:
University of Missouri, Columbia, Missouri, an academic tertiary care facility.

PARTICIPANTS:
Fourteen categorical general surgery residents (postgraduate year 1-5) were prospectively enrolled.

RESULTS:
One hundred forty simulations were included. The number of hand movements and instrument path lengths strongly correlated with simulation task times (ρ 0.62-0.87, p < 0.0001), FLS task completion times (ρ 0.50-0.53, p < 0.0001), and prior real-world laparoscopic cholecystectomy experience (ρ -0.51 to -0.53, p < 0.0001). No significant correlations were identified between any of the studied markers with Accreditation Council for Graduate Medical Education Milestones, Objective Structured Assessment of Technical Skill evaluations, total previous laparoscopic experience, or postgraduate year level. Neither instrument acceleration nor participant engagement showed significant correlation with any of the conventional markers of real-world or simulation skill proficiency.

CONCLUSIONS:
Simulation proficiency, measured by instrument and hand motion, is more representative of simulation skill than simulation task time, instrument acceleration, or participant engagement.
Pediatric Program Director Minimum Milestone Expectations before Allowing Supervision of Others and Unsupervised Practice


BACKGROUND:
The Accreditation Council for Graduate Medical Education requires semiannual Milestone reporting on all residents. Milestone expectations of performance are unknown.

OBJECTIVE:
Determine pediatric program director (PD) minimum Milestone expectations for residents prior to being ready to supervise and prior to being ready to graduate.

METHODS:
Mixed methods survey of pediatric PDs on their programs' Milestone expectations before residents are ready to supervise and before they are ready to graduate, and in what ways PDs use Milestones to make supervision and graduation decisions. If programs had no established Milestone expectations, PDs indicated expectations they considered for use in their program. Mean minimum Milestone level expectations adjusted for program size, region, and clustering of Milestone expectations by program were calculated for prior to supervise and prior to graduate. Free-text questions were analyzed using thematic analysis.

RESULTS:
The response rate was 56.8% (113/199). Most programs had no required minimum Milestone level before residents are ready to supervise (80%; 76/95) or ready to graduate (84%; 80/95). For readiness to supervise, minimum Milestone expectations PDs considered establishing for their program were highest for humanism (2.46, 95% CI: 2.21-2.71) and professionalization (2.37, 2.15-2.60). Minimum Milestone expectations for graduates were highest for help-seeking (3.14, 2.83-3.46). Main themes included the use of Milestones in combination with other information to assess learner performance and Milestones are not equally weighted when making advancement decisions.

CONCLUSIONS:
Most PDs have not established program minimum Milestones, but would vary such expectations by competency.
National Internal Medicine Milestone Ratings: Validity Evidence from Longitudinal Three-Year Follow-Up


PURPOSE:
To evaluate validity evidence for internal medicine milestone ratings across programs for three resident cohorts by quantifying "not assessable" ratings; reporting mean longitudinal milestone ratings for individual residents; and correlating medical knowledge ratings across training years with certification examination scores to determine predictive validity of milestone ratings for certification outcomes.

METHOD:
This retrospective study examined milestone ratings for postgraduate year (PGY)-1-3 residents in United States internal medicine residency programs. Data sources included milestone ratings, program characteristics, and certification examination scores.

RESULTS:
Among 35,217 participants, there was a decreased percentage with "not assessable" ratings across years: 1,566 (22.5%) PGY-1s in 2013-2014 versus 1,219 (16.6%) in 2015-2016 (P = .01), and 342 (5.1%) of PGY-3s in 2013-2014 versus 177 (2.6%) in 2015-2016 (P = .04). For individual residents with three years of ratings, mean milestone ratings increased from around 3 (behaviors of an early learner or advancing resident) in PGY-1 (ranging from a mean of 2.73 to 3.19 across subcompetencies) to around 4 (ready for unsupervised practice) in PGY-3 (mean of 4.00 to 4.22 across subcompetencies, P < .001 for all subcompetencies). For each increase of 0.5 units in two medical knowledge (MK1, MK2) subcompetency ratings, the difference in examination scores for PGY-3s was 19.5 points for MK1 (P < .001) and 19.0 for MK2 (P < .001).

CONCLUSIONS:
These findings provide evidence of validity of the milestones by showing how training programs have applied them over time and how milestones predict other training outcomes.
Simulation for Assessment of Milestones in Emergency Medicine Residents


OBJECTIVES:
All residency programs in the United States are required to report their residents' progress on the milestones to the Accreditation Council for Graduate Medical Education (ACGME) biannually. Since the development and institution of this competency-based assessment framework, residency programs have been attempting to ascertain the best ways to assess resident performance on these metrics. Simulation was recommended by the ACGME as one method of assessment for many of the milestone subcompetencies. We developed three simulation scenarios with scenario-specific milestone-based assessment tools. We aimed to gather validity evidence for this tool.

METHODS:
We conducted a prospective observational study to investigate the validity evidence for three mannequin-based simulation scenarios for assessing individual residents on emergency medicine (EM) milestones. The subcompetencies (i.e., patient care [PC]1, PC2, PC3) included were identified via a modified Delphi technique using a group of experienced EM simulationists. The scenario-specific checklist (CL) items were designed based on the individual milestone items within each EM subcompetency chosen for assessment and reviewed by experienced EM simulationists. Two independent live raters who were EM faculty at the respective study sites scored each scenario following brief rater training. The inter-rater reliability (IRR) of the assessment tool was determined by measuring intraclass correlation coefficient (ICC) for the sum of the CL items as well as the global rating scales (GRSs) for each scenario. Comparing GRS and CL scores between various postgraduate year (PGY) levels was performed with analysis of variance.

RESULTS:
Eight subcompetencies were chosen to assess with three simulation cases, using 118 subjects. Evidence of test content, internal structure, response process, and relations with other variables were found. The ICCs for the sum of the CL items and the GRSs were >0.8 for all cases, with one exception (clinical management GRS = 0.74 in sepsis case). The sum of CL items and GRSs (p < 0.05) discriminated between PGY levels on all cases. However, when the specific CL items were mapped back to milestones in various proficiency levels, the milestones in the higher proficiency levels (level 3 [L3] and 4 [L4]) did not often discriminate between various PGY levels. L3 milestone items discriminated between PGY levels on five of 12 occasions they were assessed, and L4 items discriminated only two of 12 times they were assessed.

CONCLUSION:
Three simulation cases with scenario-specific assessment tools allowed evaluation of EM residents on proficiency L1 to L4 within eight of the EM milestone subcompetencies. Evidence of test content, internal structure, response process, and relations with other variables were found. Good to excellent IRR and the ability to discriminate between various PGY levels was found for both the sum of CL items and the GRSs. However, there was a lack of a positive relationship between advancing PGY level and the completion of higher-level milestone items (L3 and L4).
An Analysis of Residents' Self Evaluation and Faculty Evaluation in Internal Medicine Standardized Residency Training Program using Milestones Evaluation System


OBJECTIVE:
To assess the value of internal medicine residency training program at Peking Union Medical College Hospital (PUMCH), and the feasibility of applying revised Milestones evaluation system.

METHODS:
Postgraduate-year-one to four (PGY-1 to PGY-4) residents in PUMCH finished the revised Milestones evaluation scales in September 2017. Residents' self-evaluation and faculty-evaluation scores were calculated. Statistical analysis was conducted on the data.

RESULTS:
A total of 207 residents were enrolled in this cross-sectional study. Both self and faculty scores showed an increasing trend in senior residents. PGY-1 residents were assessed during their first month of residency with scores of 4 points or higher, suggesting that residents have a high starting level. More strikingly, the mean score in PGY-4 was 7 points or higher, proving the career development of residency training program. There was no statistically significant difference between total self- and faculty-evaluation scores. Evaluation scores of learning ability and communication ability were lower in faculty group (t=2.627, -4.279, all P<0.05). The scores in graduate students were lower than those in standardized training residents.

CONCLUSIONS:
The goal of national standardized residency training is to improve the quality of healthcare and residents' career development. The evaluation results would guide curriculum design and emphasize the importance and necessity of multi-level teaching. Self-evaluation contributes to the understanding of training objectives and personal cognition.
Case Reporting, Competence, and Confidence: A Discrepancy in the Numbers


PURPOSE:
The Accreditation Council for Graduate Medical Education (ACGME) continues to play an integral role in accreditation of surgical programs. The institution of case logs to demonstrate competency of graduating residents is a key component of evaluation. This study compared the number of vascular cases a surgical resident has completed according to the ACGME operative log to their operative proficiency, quality of anastomosis, operative experience, and confidence in both a simulation and operative setting.

MATERIALS AND METHODS:
General surgery residents ranging from PGY 1 to 5 participated in a simulation laboratory in which they completed an end-to-side vascular anastomosis. Each participant was given a weighted score based on technical proficiency and anastomosis quality using a previously validated Global Rating Scale (Duran et al, 2014). These scores were correlated to the General Surgery Milestones. Participants completed preoperative and postoperative surveys assessing resident operative experience using the 4-level Zwisch scale (DaRosa et al., 2013), confidence with vascular procedures and confidence performing simulated anastomoses. Confidence was assessed on a scale from 1 to 9 (not confident to extremely confident). Case logs were recorded for each participant. An IRB approved questionnaire was distributed to assess preoperative and postoperative roles of both the resident physician and faculty, with a defined goal. Univariate and multivariate analysis was performed.

RESULTS:
Twenty-one general surgery residents were evaluated in the simulation laboratory and 8 residents were assessed intraoperatively. The residents were evenly distributed throughout clinical years. Groups of residents were divided into quartiles based upon the number of vascular cases recorded in the ACGME database. No correlation was found between number of cases, Milestones score and the weighted score (p = 0.94). No statistical significance was found between confidence and quality of anastomosis (p = 0.1). Resident operative experience per the Zwisch scale was categorized most commonly as "Smart Help" by both the trainee and attending surgeon, despite mean resident confidence ratings of 6.67 (± 1.61) with vascular procedures.

CONCLUSIONS:
ACGME case logs, which are utilized to assess readiness for completion of general surgery residency, may not be indicative of a resident's operative competency and technical proficiency. Confidence is not correlated with technical ability. Faculty and resident insight as to their role in a procedure differ, as faculty feel that they are providing less help than the resident perceives. Careful examination of resident operative technique is the best measure of competency.
Competency Assessment in Family Medicine Residency: Observations, Knowledge-Based Examinations, and Advancement


BACKGROUND:
The Family Medicine (FM) Milestones are competency-based assessments of residents in key dimensions relevant to practice in the specialty. Residency programs use the milestones in semiannual reviews of resident performance from the time of entry into the program to graduation.

OBJECTIVE:
Using a national sample, we investigated the relationship of FM competency-based assessments to resident progress and the complementarity of milestones with knowledge-based assessments in FM residencies.

METHODS:
We used midyear and end-of-year milestone ratings for all FM residents in Accreditation Council for Graduate Medical Education-accredited programs during academic years 2014-2015 and 2015-2016. The milestones contain 22 items across 6 competencies. We created a summative index across the milestones. The American Board of Family Medicine database provided resident demographics and in-training examination (ITE) scores. We linked information to the milestone data.

RESULTS:
The sample encompassed 6630 FM residents. The summative milestone index increased, on average, for each cohort (postgraduate year 1 [PGY-1] to PGY-2 and PGY-2 to PGY-3) at each assessment. The correlation between the milestone index that excluded the medical knowledge milestone and ITE scores was \( r = .195 \) (\( P < .001 \)) for PGY-1 to PGY-2 cohort and \( r = .254 \) (\( P < .001 \)) for PGY-2 to PGY-3 cohort. For both cohorts, ITE scores and composite milestone assessments were higher for residents who advanced than for those who did not.

CONCLUSIONS:
Competency-based assessment using the milestones for FM residents seems to be a viable multidimensional tool to assess the successful progression of residents.
OBJECTIVE:
The Accreditation Council for Graduate Medical Education requires accredited residency programs to implement competency-based assessments of medical trainees based upon nationally established Milestones. Clinical competency committees (CCC) are required to prepare biannual reports using the Milestones and ensure reporting to the Accreditation Council for Graduate Medical Education. Previous research demonstrated a strong correlation between CCC and resident scores on the Milestones at 1 institution. We sought to evaluate a national sampling of general surgery residency programs and hypothesized that CCC and resident assessments are similar.

DESIGN:
Details regarding the makeup and process of each CCC were obtained. Major disparities were defined as an absolute mean difference of ≥0.5 on the 4-point scale. A negative assessment disparity indicated that the residents evaluated themselves at a lower level than did the CCC. Statistical analysis included Wilcoxon rank sum and Sign tests.

SETTING:
CCCs and categorical general surgery residents from 15 residency programs completed the Milestones document independently during the spring of 2016.

RESULTS:
Overall, 334 residents were included; 44 (13%) and 43 (13%) residents scored themselves ≥0.5 points higher and lower than the CCC, respectively. Female residents scored themselves a mean of 0.08 points lower, and male residents scored themselves a mean of 0.03 points higher than the CCC. Median assessment differences for postgraduate year (PGY)-1-5 were 0.03 (range: -0.94 to 1.28), -0.11 (range: -1.22 to 1.22), -0.08 (range: -1.28 to 0.81), 0.02 (range: -to 1.00), and -0.19 (range: -1.16 to 0.50), respectively. Residents in university vs. independent programs had higher rates of negative assessment differences in medical knowledge (15% vs. 6%; P = 0.015), patient care (17% vs. 5%; P = 0.002), professionalism (23% vs. 14%; P = 0.013), and system-based practice (18% vs. 9%; P = 0.031) competencies. Major assessment disparities by sex or PGY were similar among individual competencies.

CONCLUSIONS:
Surgery residents in this national cohort demonstrated self-awareness when compared to assessments by their respective CCCs. This was independent of program type, sex, or level of training. PGY 5 residents, female residents, and those from university programs consistently rated themselves lower than the CCC, but these were not major disparities and the significance of this is unclear.
A Survey on Recent Medical School Graduate Comfort with the Level 1 Milestones


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.
Validity Evidence from Ratings of Pediatric Interns and Subinterns on a Subset of Pediatric Milestones


PURPOSE:
To investigate evidence for validity of faculty members’ pediatric milestone (PM) ratings of interns (first-year residents) and subinterns (fourth-year medical students) on nine subcompetencies related to readiness to serve as a pediatric intern in the inpatient setting.

METHOD:
The Association of Pediatric Program Directors Longitudinal Educational Assessment Research Network (APPD LEARN) and the National Board of Medical Examiners collaborated to investigate the utility of assessments of the PMs for trainees’ performance. Data from 32 subinterns and 179 interns at 17 programs were collected from July 2012 through April 2013. Observers used several tools to assess learners. At each site, a faculty member used these data to make judgments about the learner’s current developmental milestone in each subcompetency. Linear mixed models were fitted to milestone judgments to examine their relationship with learner's rank and subcompetency.

RESULTS:
On a 5-point developmental scale, mean milestone levels for interns ranged from 3.20 (for the subcompetency Work effectively as a member of a team) to 3.72 (Humanism) and for subinterns from 2.89 (Organize and prioritize care) to 3.61 (Professionalization). Mean milestone ratings were significantly higher for the Professionalism competency (3.59-3.72) for all trainees compared with Patient Care (2.89-3.24) and Personal and Professional Development (3.33-3.51). Mean intern ratings were significantly higher than mean subintern ratings for all nine subcompetencies except Professionalization, Humanism, and Trustworthiness.

CONCLUSIONS:
The PMs had a coherent internal structure and could distinguish between differing levels of trainees, which supports their validation for documenting developmental progression of pediatric trainees.
Examining the Functioning and Reliability of the Family Medicine Milestones


BACKGROUND:
The Family Medicine (FM) Milestones are a framework designed to assess development of residents in key dimensions of physician competency. Residency programs use the milestones in semiannual reviews of resident performance from entry toward graduation.

OBJECTIVE:
To examine the functioning and reliability of the FM Milestones and to determine whether they measure the amount of a latent trait (eg, knowledge or ability) possessed by a resident or simply indicate where a resident falls along the training sequence.

METHODS:
This study utilized the Rasch Partial Credit model to examine academic year 2014-2015 ratings for 10563 residents from 476 residency programs (postgraduate year [PGY] 1 = 3639; PGY-2 = 3562; PGY-3 = 3351; PGY-4 = 11).

RESULTS:
Reliability was exceptionally high at 0.99. Mean scores were 3.2 (SD = 1.3) for PGY-1; 5.0 (SD = 1.3) for PGY-2; 6.7 (SD = 1.2) for PGY-3; and 7.4 (SD = 1.0) for PGY-4. Keyform analysis showed a rating on 1 item was likely to be similar for all other items.

CONCLUSIONS:
Our findings suggest that FM Milestones seem to largely function as intended. Lack of spread in item difficulty and lack of variation in category probabilities show that FM Milestones do not measure the amount of a latent trait possessed by a resident, but rather describe where a resident falls along the training sequence. High reliability indicates residents are being rated in a stable manner as they progress through residency, and individual residents deviating from this rating structure warrant consideration by program leaders.
Correlations between Ratings on the Resident Annual Evaluation Summary and the Internal Medicine Milestones and Association with ABIM Certification Examination Scores among US Internal Medicine Residents, 2013-2014


IMPORTANCE: US internal medicine residency programs are now required to rate residents using milestones. Evidence of validity of milestone ratings is needed.

OBJECTIVE: To compare ratings of internal medicine residents using the pre-2015 resident annual evaluation summary (RAES), a non-developmental rating scale, with developmental milestone ratings.

DESIGN, SETTING, AND PARTICIPANTS: Cross-sectional study of US internal medicine residency programs in the 2013-2014 academic year, including 21,284 internal medicine residents (7048 postgraduate-year 1 [PGY-1], 7233 PGY-2, and 7003 PGY-3).

EXPOSURES: Program director ratings on the RAES and milestone ratings.

MAIN OUTCOMES AND MEASURES: Correlations of RAES and milestone ratings by training year; correlations of medical knowledge ratings with American Board of Internal Medicine (ABIM) certification examination scores; rating of unprofessional behavior using the 2 systems.

RESULTS: Corresponding RAES ratings and milestone ratings showed progressively higher correlations across training years, ranging among competencies from 0.31 (95% CI, 0.29 to 0.33) to 0.35 (95% CI, 0.33 to 0.37) for PGY-1 residents to 0.43 (95% CI, 0.41 to 0.45) to 0.52 (95% CI, 0.50 to 0.54) for PGY-3 residents (all P values <.05). Linear regression showed ratings differed more between PGY-1 and PGY-3 years using milestone ratings than the RAES (all P values <.001). Of the 6260 residents who attempted the certification examination, the 618 who failed had lower ratings using both systems for medical knowledge than did those who passed (RAES difference, -0.9; 95% CI, -1.0 to -0.8; P < .001; milestone medical knowledge 1 difference, -0.3; 95% CI, -0.3 to -0.3; P < .001; and medical knowledge 2 difference, -0.2; 95% CI, -0.3 to -0.2; P < .001). Of the 26 PGY-3 residents with milestone ratings indicating deficiencies on either of the 2 medical knowledge subcompetencies, 12 failed the certification examination. Correlation of RAES ratings for professionalism with residents' lowest professionalism milestone ratings was 0.44 (95% CI, 0.43 to 0.45; P < .001).

CONCLUSIONS AND RELEVANCE: Among US internal medicine residents in the 2013-2014 academic year, milestone-based ratings correlated with RAES ratings but with a greater difference across training years. Both rating systems for medical knowledge correlated with ABIM certification examination scores. Milestone ratings may better detect problems with professionalism. These preliminary findings may inform establishment of the validity of milestone-based assessment.
Using the ACGME Milestones for Resident Self-Evaluation and Faculty Engagement


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 a 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 a: 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.
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 First Look at the Accreditation Council for Graduate Medical Education Anesthesiology Milestones: Implementation of Self-Evaluation in a Large Residency Program


STUDY OBJECTIVE:
The objective was to determine if there is a correlation between resident postgraduate year (PGY) of training and self-evaluation of performance using the Accreditation Council for Graduate Medical Education milestones.

DESIGN:
Survey.

SETTING:
Residency program at a large academic center.

PATIENTS:
Residents and Faculty Clinical Competency Committee (CCC).

INTERVENTIONS:
None.

MEASUREMENTS:
Resident and CCC milestone scores.

MAIN RESULTS:
Correlation coefficients for average score for each milestone vs PGY level ranged from 0.80 for receiving and giving feedback to 0.95 for anesthetic choice and conduct. All milestones showed a relatively linear relationship with PGY of training, and none were found to be consistently reached very late or very early in training. When examining variation across the scores for the individual residents, the distributions for PGY-2 and -3 appeared to be wider than those for PGY-1 and -4. The intraclass correlation coefficients ranged from 0.718 to 0.928.

CONCLUSIONS:
There was a remarkable degree of consistency in the relationship between level of training and resident self-assessment score for every milestone, as well as strong agreement between the resident and CCC faculty scores. Examination of the variance in the scores, when interpreted in light of our particular training program’s characteristics, suggests that the milestones accurately reflect the progression in skill across the residency. In addition, given the concordance between the self-evaluation scores and the CCC faculty scores, self-evaluation may be a reasonable starting point as programs begin the daunting task of determining scores for each of the 25 milestones as part of the biannual evaluation process.
Pathology Milestones: Assessing Clinical Competency by Committee


ABSTRACT
All Accreditation Council for Graduate Medical Education accredited pathology residency training programs are now required to evaluate residents using the new Pathology Milestones assessment tool. Similar to implementation of the 6 Accreditation Council for Graduate Medical Education competencies a decade ago, there have been challenges in implementation of the new milestones for many residency programs. The pathology department at the University of Iowa has implemented a process that divides the labor of the task in rating residents while also maintaining consistency in the process. The process is described in detail, and some initial trends in milestone evaluation are described and discussed. Our experience indicates that thoughtful implementation of the Pathology Milestones can provide programs with valuable information that can inform curricular changes.
Practicing Emergency Physicians Report Performing Well on Most Emergency Medicine Milestones


BACKGROUND:
The Accreditation Council for Graduate Medical Education's Next Accreditation System endorsed specialty-specific milestones as the foundation of an outcomes-based resident evaluation process. These milestones represent five competency levels (entry level to expert), and graduating residents will be expected to meet Level 4 on all 23 milestones. Limited validation data on these milestones exist. It is unclear if higher levels represent true competencies of practicing emergency medicine (EM) attendings.

OBJECTIVE:
Our aim was to examine how practicing EM attendings in academic and community settings self-evaluate on the new EM milestones.

METHODS:
An electronic self-evaluation survey outlining 9 of the 23 EM milestones was sent to a sample of practicing EM attendings in academic and community settings. Attendings were asked to identify which level was appropriate for them.

RESULTS:
Seventy-nine attendings were surveyed, with an 89% response rate. Sixty-one percent were academic. Twenty-three percent (95% confidence interval [CI] 20%-27%) of all responses were Levels 1, 2, or 3; 38% (95% CI 34%-42%) were Level 4; and 39% (95% CI 35%-43%) were Level 5. Seventy-seven percent of attendings found themselves to be Level 4 or 5 in eight of nine milestones. Only 47% found themselves to be Level 4 or 5 in ultrasound skills (p = 0.0001).

CONCLUSIONS:
Although a majority of EM attendings reported meeting Level 4 milestones, many felt they did not meet Level 4 criteria. Attendings report less perceived competence in ultrasound skills than other milestones. It is unclear if self-assessments reflect the true competency of practicing attendings. The study design can be useful to define the accuracy, precision, and validity of milestones for any medical field.
The Pathology Milestones and the Next Accreditation System


BACKGROUND:
In the late 1990s, the Accreditation Council for Graduate Medical Education developed the Outcomes Project and the 6 general competencies with the intent to improve the outcome of graduate medical education in the United States. The competencies were used as the basis for developing learning goals and objectives and tools to evaluate residents' performance. By the mid-2000s the stakeholders in resident education and the general public felt that the Outcomes Project had fallen short of expectations.

OBJECTIVE:
To develop a new evaluation method to track trainee progress throughout residency using benchmarks called milestones. A change in leadership at the Accreditation Council for Graduate Medical Education brought a new vision for the accreditation of training programs and a radically different approach to the evaluation of residents.

DATA SOURCES:
The Pathology Milestones Working Group reviewed examples of developing milestones in other specialties, the literature, and the Accreditation Council for Graduate Medical Education program requirements for pathology to develop pathology milestones. The pathology milestones are a set of objective descriptors for measuring progress in the development of competency in patient care, procedural skill sets, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice.

CONCLUSIONS:
The milestones provide a national standard for evaluation that will be used for the assessment of all residents in Accreditation Council for Graduate Medical Education-accredited pathology training programs.
Content - Do the Milestones Represent my Specialty/Program?
Aligning Geriatric Medicine Fellowships with the Program of All-Inclusive Care for the Elderly (PACE)


ABSTRACT
Geriatric medicine fellowship programs provide comprehensive training to one-year clinical fellows and must demonstrate successful progression of competence among fellows by reporting on 23 milestones to the Accreditation Council for Graduate Medical Education (ACGME). The Program of All-inclusive Care for the Elderly (PACE) is a model of care located throughout the United States and can serve as a training venue for fellows. We surveyed 113 fellowship program directors with a response rate of 42% (n = 48). The purpose of the survey was to assess: (1) familiarity and access to PACE and (2) perceived value of PACE to the fellowship program with regard to training and ability to achieve success in the 23 reporting milestones. Milestones involving communication and team management skills were most consistently identified as very valuable with a PACE clinical rotation. We then convened a focus group of four PACE medical directors who developed a fellowship curriculum for use in training fellows at PACE. We discuss the limitations of our design as well as the opportunities to build on the strengths of that model as a training site for fellows.
Radiology Education in Medical School and Residency: The Views and Needs of Program Directors


RATIONALE AND OBJECTIVES:
The authors of this study used the perspectives of residency program directors (PDs) nationally to explore whether trainees are adequately prepared to utilize and interpret medical imaging as interns, to identify the types of imaging skills most important for residency, and to begin to address current shortcomings in radiology education.

MATERIALS AND METHODS:
The authors created a survey using a modified version of Accreditation Council for Graduate Medical Education radiology milestones and sent it to 100 randomly selected PDs each in pediatrics, internal medicine, obstetrics and gynecology, and general surgery. The survey asked PDs to assess the actual and desired imaging skills of their incoming interns, the incoming interns’ variability of skill level upon matriculation, and which imaging skills were most important from the PDs' perspective.

RESULTS:
PDs from all specialties identified a significant shortcoming relative to their expectations for both image interpretation and utilization skills. Additionally, PDs identified a significant variability in imaging skills, and described that variability as a hindrance to their programs. All of the potential imaging skills were rated as highly important with little clinically relevant difference between them.

DISCUSSION:
This multidisciplinary national survey found a deficiency in imaging education among interns across specialties and substantiates calls for formalized and improved radiology education in undergraduate medical education. Additionally, PDs had difficulty distinguishing which skills were most important, suggesting an unclear understanding of imaging ability needs for interns in respective specialties. More specific needs assessments are warranted on a national level.
Gender Bias in Simulation-Based Assessments of Emergency Medicine Residents


BACKGROUND:
Gender-related disparities persist in medicine and medical education. Prior work has found differences in medical education assessments based on gender.

OBJECTIVE:
We hypothesized that gender bias would be mitigated in a simulation-based assessment.

METHODS:
We conducted a retrospective cohort study of emergency medicine residents at a single, urban residency program. Beginning in spring 2013, residents participated in mandatory individual simulation assessments. Twelve simulated cases were included in this study. Rating forms mapped milestone language to specific observable behaviors. A Bayesian regression was used to evaluate the effect of resident and rater gender on assessment scores. Both 95% credible intervals (CrIs) and a Region of Practical Equivalence approach were used to evaluate the results.

RESULTS:
Participants included 48 faculty raters (25 men [52%]) and 102 residents (47 men [46%]). The difference in scores between male and female residents (M = -0.58, 95% CrI -3.31-2.11), and male and female raters (M = 2.87, 95% CrI -0.43-6.30) was small and 95% CrIs overlapped with 0. The 95% CrI for the interaction between resident and rater gender also overlapped with 0 (M = 0.41, 95% CrI -3.71-4.23).

CONCLUSIONS:
In a scripted and controlled system of assessments, there were no differences in scores due to resident or rater gender.
The American Society of Hematology and ASCO Curricular Milestones for Assessment of Fellows in Hematology/Oncology: Development, Reflection, and Next Steps


ABSTRACT
The American Society of Hematology (ASH)/ASCO Curricular Milestones is a tool for assessment and teaching for fellows in hematology/oncology. The expectations of the Next Accreditation System of the Accreditation Council of Graduate Medical Education (ACGME) was developed over years from the creation of the six core competencies in 1999 to the current data-driven outcomes-based system. The current internal medicine subspecialty milestones (ACGME reporting milestones) follow the general rubric of the general internal medicine milestones. The ASH/ASCO curricular milestones were developed from the foundational elements of the specialty, and they are interwoven with the ACGME reporting milestones. The 2017 ACGME Milestones Report shows that the milestones display progression in performance through clear anchors. Educational outcomes are available in many specialties. The internal medicine subspecialties have been given the opportunity to update the ACGME reporting milestones. The ACGME has acknowledged that these milestones may be different for each of the specialties. The program committees of ASH and ASCO agree that revision of the ACGME reporting milestones would decrease the overlap of domains, lack of clarity, and negative language that is present in version 1.0. ASH and ASCO are working with the ACGME and American Board of Internal Medicine (ABIM) to develop Curricular Milestones, version 2.0.
Teaching and assessing trainees’ professionalism now represents an explicit expectation for Accreditation Council Graduate Medical Education-accredited radiology programs. Challenges to meeting this expectation include variability in defining the construct of professionalism; limits of traditional teaching and assessment methods, used for competencies historically more prominent in medical education, for professionalism; and emerging expectations for credible and feasible professionalism teaching and assessment practices in the current context of health-care training and practice. This article identifies promising teaching resources and methods that can be used strategically to augment traditional teaching of the cognitive basis for professionalism, including role modeling, case-based scenarios, debriefing, simulations, narrative medicine (storytelling), guided discussions, peer-assisted learning, and reflective practice. This article also summarizes assessment practices intended to promote learning, as well as to inform how and when to assess trainees as their professional identities develop over time, settings, and autonomous practice, particularly in terms of measurable behaviors. This includes assessment tools (including mini observations, critical incident reports, and appreciative inquiry) for authentic assessment in the workplace; engaging multiple sources (self, peer, other health professionals, and patients) in assessment; and intentional practices for trainees to take responsibility for seeking our actionable feedback and reflection. This article examines the emerging evidence of the feasibility and value added of assessment of medical competency milestones, including professionalism, coordinated by the Accreditation Council Graduate Medical Education in radiology and other medical specialties. Radiology has a strategic opportunity to contribute to scholarship and inform policies in professionalism teaching and assessment practices.
OBJECTIVES:
Emergency medicine (EM) physicians commonly care for patients with serious life-limiting illness. Hospice and palliative medicine (HPM) is a subspecialty pathway of EM. Although a subspecialty level of practice requires additional training, primary-level skills of HPM such as effective communication and symptom management are part of routine clinical care and expected of EM residents. However, unlike EM residency curricula in disciplines like trauma and ultrasound, there is no nationally defined HPM curriculum for EM resident training. An expert consensus group was convened with the aim of defining content areas and competencies for HPM primary-level practice in the ED setting. Our overall objective was to develop HPM milestones within a competency framework that is relevant to the practice of EM.

METHODS:
The American College of Emergency Physicians Palliative Medicine Section assembled a committee that included academic EM faculty, community EM physicians, EM residents, and nurses, all with interest and expertise in curricular design and palliative medicine.

RESULTS:
The committee peer reviewed and assessed HPM content for validity and importance to EM residency training. A topic list was developed with three domains: provider skill set, clinical recognition of HPM needs, and logistic understanding related to HPM in the ED. The group also developed milestones in HPM-EM to identify relevant knowledge, skills, and behaviors using the framework modeled after the Accreditation Council for Graduate Medical Education (ACGME) EM milestones. This framework was chosen to make the product as user-friendly and familiar as possible to facilitate use by EM educators.

CONCLUSIONS:
Educators in EM residency programs now have access to HPM content areas and milestones relevant to EM practice that can be used for curriculum development in EM residency programs. The HPM-EM skills/competencies presented herein are structured in a familiar milestone framework that is modeled after the widely accepted ACGME EM milestones.
Mapping the Balint Groups to the Accreditation Council for Graduate Medical Education Family Medicine Competencies


ABSTRACT:
Introduction Balint group discussions provide learning opportunities for many of the competencies and milestones put forward by the Accreditation Council for Graduate Medical Education. The current literature is mixed concerning the effect of Balint groups on communication skills and professionalism. Aim To map the content of the Balint discussion to the competencies and milestones put forward by the Accreditation Council for Graduate Medical Education.

METHODOLOGY:
A group who were both experts in Balint and members of the clinical competency committee of residency programs rated narratives that summarized Balint group discussions. Credentialed Leaders of the American Balint Society were invited via email to submit narratives (250 words) about Balint groups that they have led, or were leading, with residents.

RESULTS:
Only four narratives were submitted. Additional cases were recruited through literature review of published Balint discussion cases. A total of 25 cases were rated by the committee. There was agreement between three out of four raters on at least one core milestone in every case. The most frequent milestones were C1 (develops meaningful therapeutic relationships with patients and families), C2 (communicated effectively with patients, families, and public), Prof1 (completes a process of professionalization), and Prof3 (demonstrates humanism and cultural proficiency). Balint groups provided a learning opportunity for a subset of milestones in at least 36% of the cases.

CONCLUSION:
This pilot research suggests that Balint groups and the discussions of complex and challenging cases provide learning opportunities for multiple family medicine milestones, mainly communication skills and professionalism. Further research is needed to refine the methodology and the rating system.
Trapped as a Group, Escape as a Team: Applying Gamification to Incorporate Teambuilding Skills Through an 'Escape Room' Experience


ABSTRACT
Teamwork, a skill critical for quality patient care, is recognized as a core competency by the Accreditation Council for Graduate Medical Education (ACGME). To date, there is no consensus on how to effectively teach these skills in a forum that engages learners, immerses members in life-like activities, and builds both trust and rapport. Recreational 'Escape Rooms' have gained popularity in creating a life-like environment that rewards players for working together, solving puzzles, and completing successions of mindbending tasks in order to effectively 'escape the room' in the time allotted. In this regard, escape rooms share many parallels with the multitasking and teamwork that is essential for a successful emergency department (ED) shift. A pilot group of nine emergency medicine (EM) residents and one senior EM faculty member underwent a commercial escape room as part of a teambuilding exercise in January 2018. The escape room required participants to practice teamwork, communication, task delegation, and critical thinking to tackle waves of increasingly complex puzzles, ranging from hidden objects, physical object assembly (i.e., jigsaw puzzles), and symbol matching.

Activities required members to recognize and utilize the collective experiences, skills, knowledge base, and physical abilities of the group. After the game, players underwent a structured 'game-master' debriefing facilitated by an employee of the commercial escape room; this was followed by a post-event survey facilitated by a faculty member, which focused on participants' feelings, experiences, and problem-solving techniques. Escape rooms afford learners the opportunity to engage in an activity that rewards teamwork and effective leadership through experiences that directly link to specific ACGME milestones and educational learning theories. EM participants were engaged in the activity and felt that the escape room reproduced an environment analogous to the ED. The debriefing that followed the activity provided a satisfactory conclusion to the experience; but learners preferred a more organized debriefing format that provided them with constructive and specific feedback on their performance.
Consensus Statement by the Congenital Cardiac Anesthesia Society: Milestones for the Pediatric Cardiac Anesthesia Fellowship


ABSTRACT

Pediatric cardiac anesthesiology has evolved as a subspecialty of both pediatric and cardiac anesthesiology and is devoted to caring for individuals with congenital heart disease ranging in age from neonates to adults. Training in pediatric cardiac anesthesia is a second-year fellowship with variability in both training duration and content and is not accredited by the Accreditation Council on Graduate Medical Education. Consequently, in this article and based on the Accreditation Council on Graduate Medical Education Milestones Model, an expert panel of the Congenital Cardiac Anesthesia Society, a section of the Society of Pediatric Anesthesiology, defines 18 milestones as competency-based developmental outcomes for training in the pediatric cardiac anesthesia fellowship.
**Competency Crosswalk: A Multispecialty Review of the Accreditation Council for Graduate Medical Education Milestones across Four Competency Domains**


**PURPOSE:**
To identify common and overlapping themes among the interpersonal and communication skills (ICS), practice-based learning and improvement (PBLI), professionalism (PROF), and systems-based practice (SBP) milestones of the transitional year and 26 specialties.

**METHOD:**
In May 2017, milestones were accessed from the Accreditation Council for Graduate Medical Education specialties website. A thematic analysis of the ICS, PBLI, PROF, and SBP milestones was performed to determine unique and common themes across these competencies and across specialties. Keywords from the common program requirements were initially applied as codes to the milestones. Codes were then grouped into common themes.

**RESULTS:**
Twenty-two themes were identified: 15 (68%) were unique to a given competency (3 related to ICS, 4 related to PBLI, 5 related to PROF, and 3 related to SBP), and 7 (32%) appeared in the milestones of more than one core competency. Eleven themes (50%) were used by 20 or more specialties, and 6 themes (27%) by 10 or fewer specialties. No theme was present across all specialties.

**CONCLUSIONS:**
The ICS, PBLI, PROF, and SBP milestones contain multiple themes with areas of overlap among these four competencies and substantial variability across specialties. This variability may create differential expectations of residents across specialties, complicate faculty development, and make sharing assessment tools difficult. The thematic analysis provides important insights into how individual specialties interpret and operationalize the ICS, PBLI, PROF, and SBP competency domains and can inform future revisions of milestones to enable harmonization and shared understanding of these competencies across specialties where appropriate.
What Every Graduating Resident Needs to Know About Quality Improvement and Patient Safety: A Content Analysis of 26 Sets of ACGME Milestones


PURPOSE:
Quality improvement (QI) and patient safety (PS) are broadly relevant to the practice of medicine, but specialty-specific milestones demonstrate variable expectations for trainee competency in QI/PS. The purpose of this study was to develop a unifying portrait of QI/PS expectations for graduating residents irrespective of specialty.

METHOD:
Milestones from 26 residency programs representing the 24 member boards of the American Board of Medical specialties were downloaded from the Accreditation Council for Graduate Medical Education (ACGME) website in 2015. A codebook was generated by in-depth reading of all milestone sets by two authors. Using a content analytic approach, milestones were then coded by a single author, with a 10% sample double-coded by another author. Descriptive statistics were used to characterize frequency counts.

RESULTS:
Of 612 total milestones, 249 (40.7%) made mention of QI/PS. A median 10 milestones per specialty (interquartile range, 5.25-11.75) mentioned QI/PS. There were 446 individual references to QI, 423 references to PS, and another 1,065 references to QI/PS-related concepts, including patient-centered care, cost-effective practice, documentation, equity, handoffs and care transitions, and teamwork. QI/PS references reflected expectations about both individual-level practice (531/869, 61.1%) and practice within a healthcare system (338/869, 38.9%). QI and PS references were linked to all six ACGME core competencies.

CONCLUSIONS:
Although there is variability in the emphasis placed on QI/PS across specialties, overall, QI/PS is reflected in more than 40% of residency milestones. Graduating residents in all specialties are expected to demonstrate competence in QI, PS, and multiple related concepts.
Aligning In-Service Training Examinations in Plastic Surgery and Orthopaedic Surgery with Competency-Based Education


BACKGROUND:
In-service training examinations (ITEs) are used to assess residents across specialties. However, it is not clear how they are integrated with the Accreditation Council for Graduate Medical Education Milestones and competencies.

OBJECTIVE:
This study explored the distribution of specialty-specific milestones and competencies in ITEs for plastic surgery and orthopaedic surgery.

METHODS:
In-service training examinations were publicly available for plastic surgery (PSITE) and orthopaedics (OITE). Questions on the PSITE for 2014-2016 and the OITE for 2013-2015 were mapped to the specialty-specific milestones and the 6 competencies.

RESULTS:
There was an uneven distribution of milestones and competencies in ITE questions. Nine of the 36 Plastic Surgery Milestones represented 52% (341 of 650) of questions, and 3 were not included in the ITE. Of 41 Orthopaedic Surgery Milestones, 7 represented 51% (201 of 394) of questions, and 5 had no representation on the ITE. Among the competencies, patient care was the most common (PSITE = 62% [403 of 650]; OITE = 59% [233 of 394]), followed by medical knowledge (PSITE = 34% [222 of 650]; OITE = 31% [124 of 394]). Distribution of the remaining competencies differed between the 2 specialties (PSITE = 4% [25 of 650]; OITE = 9% [37 of 394]).

CONCLUSIONS:
The ITEs tested slightly more than half of the milestones for the 2 specialties, and focused predominantly on patient care and medical knowledge competencies.
Gender Differences in Attending Physicians’ Feedback to Residents: A Qualitative Analysis


BACKGROUND:
Prior research has shown a gender gap in the evaluations of emergency medicine (EM) residents' competency on the Accreditation Council for Graduate Medical Education (ACGME) milestones, yet the practical implications of this are not fully understood.

OBJECTIVE:
To better understand the gender gap in evaluations, we examined qualitative differences in the feedback that male and female residents received from attending physicians.

METHODS:
This study used a longitudinal qualitative content analysis of narrative comments by attending physicians during real-time direct observation milestone evaluations of residents. Comments were collected over 2 years from 1 ACGME-accredited EM training program.

RESULTS:
In total, 1317 direct observation evaluations with comments from 67 faculty members were collected for 47 postgraduate year 3 EM residents. Analysis of the comments revealed that the ideal EM resident possesses many stereotypically masculine traits. Additionally, examination of a subset of the residents (those with 15 or more comments, n = 35) showed that when male residents struggled, they received consistent feedback from different attending physicians regarding aspects of their performance that needed work. In contrast, when female residents struggled, they received discordant feedback from different attending physicians, particularly regarding issues of autonomy and assertiveness.

CONCLUSIONS:
Our study revealed qualitative differences in the kind of feedback that male and female EM residents received from attending physicians. The findings suggest that attending physicians should endeavor to provide male and female residents with consistent feedback and guard against gender bias in their perceptions of residents' capabilities.
Ways to Write a Milestone: Approaches to Operationalizing the Development of Competence in Graduate Medical Education


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

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

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

CONCLUSIONS:
Initial ACGME milestones operationalize the development of competence in many ways. These findings offer insights into how physicians understand and assess the developmental progression of competence and an opportunity to consider how different approaches may affect the validity of milestone-based assessments. The results of this analysis can inform the work of educators developing or revising milestones, interpreting milestone data, or creating assessment tools to inform milestone-based performance measures.
Competency-Based Medical Education and the Ghost of Kuhn: Reflections on the Messy and Meaningful Work of Transformation


ABSTRACT
The transition, if not transformation, to outcomes-based medical education likely represents a paradigm shift struggling to be realized. Paradigm shifts are messy and difficult but ultimately meaningful if done successfully. This struggle has engendered tension and disagreements, with many of these disagreements cast as either-or polarities. There is little disagreement, however, that the health care system is not effectively achieving the triple aim for all patients. Much of the tension and polarity revolve around how more effectively to prepare students and residents to work in and help change a complex health care system. Competencies were an initial attempt to facilitate this shift by creating frameworks of essential abilities needed by physicians. However, implementation of competencies has proven to be difficult. Entrustable professional activities (EPAs) in undergraduate and graduate medical education and Milestones in graduate medical education are recent concepts being tried and studied as approaches to guide the shift to outcomes. Their primary purpose is to help facilitate implementation of an outcomes-based approach by creating shared mental models of the competencies, which in turn can help to improve curricula and assessment. Understanding whether and how EPAs and Milestones effectively facilitate the shift to outcomes has been and will continue to be an iterative and ongoing reflective process across the entire medical education community using lessons from implementation and complexity science. In this Invited Commentary, the author reflects on what got the community to this point and some sources of tension involved in the struggle to move to outcomes-based education.
Critical Deficiency Ratings in Milestone Assessment: A Review and Case Study


PURPOSE:
The Accreditation Council for Graduate Medical Education (ACGME) requires programs to report learner progress using specialty-specific milestones. It is unclear how milestones can best identify critical deficiencies (CDs) in trainee performance. Specialties developed milestones independently of one another; not every specialty included CDs within milestones ratings. This study examined the proportion of ACGME milestone sets that include CD ratings, and describes one residency program's experiences using CD ratings in assessment.

METHOD:
The authors reviewed ACGME milestones for all 99 specialties in November 2015, determining which rating scales contained CDs. The authors also reviewed three years of data (July 2012-June 2015) from the University of Cincinnati Medical Center (UCMC) internal medicine residency assessment system based on observable practice activities mapped to ACGME milestones. Data were analyzed by postgraduate year, assessor type, rotation, academic year, and core competency. The Mantel-Haenszel chi-square test was used to test for changes over time.

RESULTS:
Specialties demonstrated heterogeneity in accounting for CDs in ACGME milestones, with 22% (22/99) of specialties having no language describing CDs in milestones assessment. Thirty-three percent (63/189) of UCMC internal medicine residents received at least one CD rating, with CDs accounting for 0.18% (668/364,728) of all assessment ratings. The authors identified CDs across multiple core competencies and rotations.

CONCLUSIONS:
Despite some specialties not accounting for CDs in milestone assessment, UCMC's experience demonstrates that a significant proportion of residents may be rated as having a CD during training. Identification of CDs may allow programs to develop remediation and improvement plans.
“The Actualized Neurosurgeon”: A Proposed Model of Surgical Resident Development


BACKGROUND:
Modern neurosurgical training is both physically and emotionally demanding, posing significant challenges, new and old, to residents as well as programs attempting to train safe, competent surgeons. Models to describe resident development, such as the Accreditation Council for Graduate Medical Education competencies and milestones, address the acquisition of specific skills but largely ignore the stresses and pressures unique to each stage of resident training.

METHODS:
We propose an alternative model of resident development adapted from the developmental psychology literature.

RESULTS:
Our model identifies the challenges that must be met at each stage of junior, intermediate, and senior and chief residency, leading ultimately to an "actualized" neurosurgeon (i.e., one who has maximized his or her potential). Failure to overcome any 1 of these challenges can lead to specific long-lasting consequences, including regret, identity crisis, incompetence, and bitterness. In contrast, the actualized surgeon is one who has successfully acquired the virtues of hope, will, purpose, fidelity, productivity, leadership, competence, and wisdom. The actualized surgeon not only functions safely, confidently, and professionally, but also successfully navigates the challenges of residency and emerges from them having fulfilled his or her maximal potential.

CONCLUSIONS:
This developmental perspective provides an individualized description of healthy surgical development. Our model allows programs to identify the basis for residents who fail to progress, counsel residents during their training, and perhaps help identify resident candidates who are better prepared to meet the developmental challenges of residency training.
Implementation of a Needs-Based, Online Feedback Tool for Anesthesia Residents with Subsequent Mapping of the Feedback to the ACGME Milestones


BACKGROUND:
Optimizing feedback that residents receive from faculty is important for learning. The goals of this study were to (1) conduct focus groups of anesthesia residents to define what constitutes optimal feedback; (2) develop, test, and implement a web-based feedback tool; and (3) then map the contents of the written comments collected on the feedback tool to the Accreditation Council for Graduate Medical Education (ACGME) anesthesiology milestones.

METHODS:
All 72 anesthesia residents in the program were invited to participate in 1 of 5 focus groups scheduled over a 2-month period. Thirty-seven (51%) participated in the focus groups and completed a written survey on previous feedback experiences. On the basis of the focus group input, an initial online feedback tool was pilot-tested with 20 residents and 62 feedback sessions, and then a final feedback tool was deployed to the entire residency to facilitate the feedback process. The completed feedback written entries were mapped onto the 25 ACGME anesthesiology milestones.

RESULTS:
Focus groups revealed 3 major barriers to good feedback: (1) too late such as, for example, at the end of month-long clinical rotations, which was not useful because the feedback was delayed; (2) too general and not specific enough to immediately remedy behavior; and (3) too many in that the large number of evaluations that existed that were unhelpful such as those with unclear behavioral anchors compromised the overall feedback culture. Thirty residents (42% of 72 residents in the program) used the final online feedback tool with 121 feedback sessions with 61 attendings on 15 rotations at 3 hospital sites. The number of feedback tool uses per resident averaged 4.03 (standard deviation 5.08, median 2, range 1-21, 25th-75th % quartile 1-4). Feedback tool uses per faculty averaged 1.98 (standard deviation 3.2, median 1, range 1-25, 25th-75th % quartile 1-2). For the feedback question item "specific learning objective demonstrated well by the resident," this yielded 296 milestone-specific responses. The majority (71.3%) were related to the patient care competency, most commonly the anesthetic plan and conduct (35.8%) and airway management (11.1%) milestones; 10.5% were related to the interpersonal and communication skills competency, most commonly the milestones communication with other professionals (4.4%) or with patients and families (4.4%); and 8.4% were related to the practice-based learning and improvement competency, most commonly self-directed learning (6.1%). For the feedback tool item "specific learning objective that resident may improve," 67.0% were related to patient care, most commonly anesthetic plan and conduct (33.5%) followed by use/interpretation of monitoring and equipment (8.5%) and airway management (8.5%); 10.2% were related to practice-based learning and improvement, most commonly self-directed learning (6.8%); and 9.7% were related to the systems-based practice competency.

CONCLUSIONS:
Resident focus groups recommended that feedback be timely and specific and be structured around a tool. A customized online feedback tool was developed and implemented. Mapping of the free-text feedback comments may assist in assessing milestones. Use of the feedback tool was lower than expected, which may indicate that it is just 1 of many implementation steps required for behavioral and culture change to support a learning environment with frequent and useful feedback.
New Roadmap for the Journey From Internist to Rheumatologist


OBJECTIVE:
Measurement is necessary to gauge improvement. US training programs have not previously used shared standards to assess trainees’ mastery of the knowledge, skills, and attitudes necessary to practice rheumatology competently. In 2014, the Accreditation Council for Graduate Medical Education (ACGME) Next Accreditation System began requiring semiannual evaluation of all medicine subspecialty fellows on 23 internal medicine subspecialty reporting milestones. Since these reporting milestones are not subspecialty specific, rheumatology curricular milestones were needed to guide rheumatology fellowship training programs and fellows on the training journey from internist to rheumatologist.

METHODS:
Rheumatology curricular milestones were collaboratively composed by expanding the internal medicine reporting milestones to delineate the specific targets of rheumatology fellowship training within 6 ACGME core competencies. The 2006 American College of Rheumatology core curriculum for rheumatology training programs was updated.

RESULTS:
A total of 80 rheumatology curricular milestones were created, defining progressive learning through training; most focus on patient care and medical knowledge. The core curriculum update incorporates the new curricular milestones and rheumatology entrustable professional activities.

CONCLUSION:
Rheumatology curricular milestones are now available for implementation by rheumatology fellowship training programs, providing a clear roadmap for specific training goals and a guide to track each fellow’s achievement over a 2-year training period. The comprehensive core curriculum delineates the essential breadth of knowledge, skills, and attitudes that define rheumatology, and provides a guide for educational activities during fellowship training. These guiding documents are now used to train and assess fellows as they prepare for independent rheumatology practice as the next generation of rheumatologists.
Emergency Medicine Resident Assessment of the Emergency Ultrasound Milestones and Current Training Recommendations


OBJECTIVES
Emergency ultrasound (EUS) has been recognized as integral to the training and practice of emergency medicine (EM). The Council of Emergency Medicine Residency-Academy of Emergency Ultrasound (CORD-AEUS) consensus document provides guidelines for resident assessment and progression. The Accredited Council for Graduate Medical Education (ACGME) has adopted the EM Milestones for assessment of residents' progress during their residency training, which includes demonstration of procedural competency in bedside ultrasound. The objective of this study was to assess EM residents' use of ultrasound and perceptions of the proposed ultrasound milestones and guidelines for assessment.

METHODS
This study is a prospective stratified cluster sample survey of all U.S. EM residency programs. Programs were stratified based on their geographic location (Northeast, South, Midwest, West), presence/absence of ultrasound fellowship program, and size of residency with programs sampled randomly from each stratum. The survey was reviewed by experts in the field and pilot tested on EM residents. Summary statistics and 95% confidence intervals account for the survey design, with sampling weights equal to the inverse of the probability of selection, and represent national estimates of all EM residents.

RESULTS
There were 539 participants from 18 residency programs with an overall survey response rate of 85.1%. EM residents considered several applications to be core applications that were not considered core applications by CORD-AEUS (quantitative bladder volume, diagnosis of joint effusion, interstitial lung fluid, peritonsillar abscess, fetal presentation, and gestational age estimation). Of several core and advanced applications, the Focused Assessment with Sonography in Trauma examination, vascular access, diagnosis of pericardial effusion, and cardiac standstill were considered the most likely to be used in future clinical practice. Residents responded that procedural guidance would be more crucial to their future clinical practice than resuscitative or diagnostic ultrasound. They felt that an average of 325 (301-350) ultrasound examinations would be required to be proficient, but felt that number of examinations poorly represented their competency. They reported high levels of concern about medicolegal liability while using EUS. Eighty-nine percent of residents agreed that EUS is necessary for the practice of EM.

CONCLUSIONS
EM resident physicians' opinion of what basic and advanced skills they are likely to utilize in their future clinical practice differs from what has been set forth by various groups of experts. Their opinion of how many ultrasound examinations should be required for competency is higher than what is currently expected during training.
Goal-directed Focused Ultrasound Milestones Revised: A Multiorganizational Consensus


ABSTRACT
In 2012 the Accreditation Council for Graduate Medical Education and the American Board of Emergency Medicine released the emergency medicine milestones. The Patient Care 12 (PC12) subcompetency delineates staged and progressive accomplishment in emergency ultrasound. While valuable as an initial framework for ultrasound resident education, there are limitations to PC12. This consensus paper provides a revised description of criteria to define the subcompetency. A multiorganizational task force was formed between the American College of Emergency Physicians Ultrasound Section, the Council of Emergency Medicine Residency Directors, and the Academy of Emergency Ultrasound of the Society for Academic Emergency Medicine. Representatives from each organization created this consensus document and revision.
ASDS Cosmetic Dermatologic Surgery Fellowship Milestones


BACKGROUND:
The American Council of Graduate Medical Education, which oversees much of postgraduate medical education in the United States, has championed the concept of "milestones," standard levels of achievement keyed to particular time points, to assess trainee performance during residency.

OBJECTIVE:
To develop a milestones document for the American Society for Dermatologic Surgery (ASDS) Cosmetic Dermatologic Surgery (CDS) fellowship program.

METHODS:
An ad hoc milestone drafting committee was convened that included members of the ASDS Accreditation Work Group and program directors of ASDS-approved Cosmetic Dermatologic Surgery (CDS) fellowship training programs. Draft milestones were circulated through email in multiple rounds until consensus was achieved.

RESULTS:
Thirteen milestones were developed in the 6 Accreditation Council for Graduate Medical Education (ACGME) competency areas, with 8 of these being patient-care milestones. Additional instructions for milestone administration more specific to the CDS fellowship than general ACGME instructions were also approved. Implementation of semiannual milestones was scheduled for the fellowship class entering in July 2018.

CONCLUSION:
Milestones are now available for CDS fellowship directors to implement in combination with other tools for fellow evaluation.
A Survey of Ultrasound Milestone Incorporation into Emergency Medicine Training Programs


OBJECTIVES:
With the introduction of the Emergency Medicine Milestone Project in 2013, residencies now assess emergency ultrasound (US) skills at regular intervals. However, it is unclear how programs are implementing the emergency US milestones and assessing competency. With the use of the milestone tool, a survey was distributed to emergency US educators to determine when programs are providing emergency US education, when residents are expected to attain competency, and whether the milestones reflect their expectations of trainees.

METHODS:
We conducted a prospective cross-sectional survey study distributed electronically to designated emergency US experts at 169 programs. Participants were queried on education and competency evaluation within the context of the milestones by designating a postgraduate year when the 5 milestone levels were taught and competency was expected. Survey findings were reported as percentages of total respondents from descriptive statistics.

RESULTS:
Responses were received from 53% of programs, and 99% were familiar with the milestones. Most programs provide level 1 (88%) and 2 (85%) instruction during postgraduate year 1. Most programs expect level 1 competency before residency (61%) and expect mastery of level 2 by the end of postgraduate year 1 (60%). Sixty-two percent believe the milestones do not accurately reflect their expectations, citing insufficient minimum scan numbers, lack of specificity, and unattainable level 5 requirements.

CONCLUSIONS:
There is substantial variability in the frequency and methods of competency evaluation using the emergency US milestones. However, most responders agree that residents should obtain level 2 competency by postgraduate year 1. Variation exists regarding what year and what skills define level 3 or greater competency.
OBJECTIVE:
The Accreditation Council of Graduate Medical Education (ACGME) Milestone Project is the next step in a series of changes revamping the system of graduate medical education. In 2013 the ACGME completed the general psychiatry milestones. The ACGME then pursued creation of milestones for accredited psychiatric subspecialty fellowships. This article documents the work of the geriatric psychiatry subspecialty milestones work group. It reports the history and rationale supporting the milestones, the milestone development process, and the implications for geriatric psychiatry fellowship training.

METHODS:
In consultation with the American Association for Geriatric Psychiatry, the American Board of Psychiatry and Neurology, and the ACGME Psychiatry Residency Review Committee, the ACGME appointed a working group to create the geriatric psychiatry milestones using the general psychiatry milestones as a guide.

CONCLUSION:
The geriatric psychiatry milestones are the result of an iterative process resulting in the definition of the characteristics vital to a fellowship-trained geriatric psychiatrist. It is premature to assess their effect on psychiatric training. The true impact of the milestones will be determined as each training director uses the milestones to re-evaluate their program curriculum and the educational and clinical learning environment. The ACGME is currently collecting the information about the milestone performance of residents and fellows to further refine and determine how the milestones can best be used to assist programs in improving training.
Milestones for the Final Mile: Interspecialty Distinctions in Primary Palliative Care Skills Training


CONTEXT:
Primary palliative care (PPC) skills are useful in a wide variety of medical and surgical specialties, and the expectations of PPC skill training are unknown across graduate medical education.

OBJECTIVES:
We characterized the variation and quality of PPC skills in residency outcomes-based Accreditation Council for Graduate Medical Education (ACGME) milestones.

METHODS:
We performed a content analysis with structured implicit review of 2015 ACGME milestone documents from 14 medical and surgical specialties chosen for their exposure to clinical situations requiring PPC. For each specialty milestone document, we characterized the variation and quality of PPC skills in residency outcomes-based ACGME milestones.

RESULTS:
We identified 959 occurrences of 29 palliative search terms within 14 specialty milestone documents. Within these milestone documents, implicit review characterized 104 milestones with direct saliency to PPC skills and 196 milestones with indirect saliency. Initial interrater agreement of the saliency rating among the primary reviewers was 89%. Specialty milestone documents varied widely in their incorporation of PPC skills within milestone documents. PPC milestones were most commonly found in milestone documents for Anesthesiology, Pediatrics, Urology, and Physical Medicine and Rehabilitation. PPC-relevant milestones were most commonly found in the Interpersonal and Communication Skills core competency with 108 (36%) relevant milestones classified under this core competency.

CONCLUSIONS:
Future revisions of specialty-specific ACGME milestone documents should focus on currently underrepresented, but important PPC skills.
What Is a Rheumatologist and How Do We Make One?


OBJECTIVE:
Graduate medical education is a critical time in the training of a rheumatologist, and purposeful evaluation of abilities during this time is essential for long-term success as an independent practitioner. The internal medicine subspecialties collectively developed a uniform set of reporting milestones by which trainees can be assessed and receive formative feedback, providing clarity of accomplishment as well as areas for improvement in training. Furthermore, the reporting milestones provide a schema for assessment and evaluation of fellows by supervisors. The internal medicine subspecialties were also tasked with considering entrustable professional activities (EPAs), which define the abilities of a subspecialty physician who has attained sufficient mastery of the field to be accountable to stakeholders and participate in independent practice. Although EPAs have been established for a few specialties, they had not yet been described for rheumatology. EPAs have value as descriptors of the comprehensive abilities, knowledge, and skills of a practicing rheumatologist. The rheumatology EPAs have a role in defining a specialist in rheumatology upon completion of training, and also represent the ways our specialty defines our abilities that are enduring throughout practice.

METHODS:
We describe the collaborative process of the development of both the subspecialty reporting milestones and the rheumatology EPAs. The reporting milestones evolved through discussions and collaborations among representatives from the Association of Specialty Professors, the Alliance for Academic Internal Medicine, the American Board of Internal Medicine, and the Accreditation Council for Graduate Medical Education. The EPAs were a product of deliberations by the Next Accreditation System (NAS) working group of the American College of Rheumatology (ACR) Committee on Rheumatology Training and Workforce Issues.

RESULTS:
Twenty-three subspecialty reporting milestones and 14 rheumatology EPAs were advanced and refined over the course of 3 subspecialty reporting milestone development summits and 3 ACR NAS working group meetings, respectively.

CONCLUSION:
The subspecialty reporting milestones and rheumatology EPAs presented here stipulate reasonable and measurable expectations for rheumatologists-in-training. Together, these tools aim to promote enrichment and greater accountability in the training of fellows. Additionally, the EPAs define, for all stakeholders, the expertise of a rheumatologist in practice.
What's New in 10 Years? A Revised Cardiothoracic Curriculum for Diagnostic Radiology Residency with Goals and Objectives Related to General Competencies


ABSTRACT
This is a cardiothoracic curriculum document for radiology residents meant to serve not only as a study guide for radiology residents but also as a teaching and curriculum reference for radiology educators and radiology residency program directors. This document represents a revision of a cardiothoracic radiology resident curriculum that was published 10 years ago in Academic Radiology. The sections that have been significantly revised, expanded, or added are (1) lung cancer screening, (2) lung cancer genomic profiling, (3) lung adenocarcinoma revised nomenclature, (4) lung biopsy technique, (5) nonvascular thoracic magnetic resonance, (6) updates to the idiopathic interstitial pneumonias, (7) cardiac computed tomography updates, (8) cardiac magnetic resonance updates, and (9) new and emerging techniques in cardiothoracic imaging. This curriculum was written and endorsed by the Education Committee of the Society of Thoracic Radiology. This curriculum operates in conjunction with the Accreditation Council for Graduate Medical Education (ACGME) milestones project that serves as a framework for semiannual evaluation of resident physicians as they progress through their training in an ACGME-accredited residency or fellowship programs. This cardiothoracic curriculum document is meant to serve not only as a more detailed guide for radiology trainees, educators, and program directors but also complementary to and guided by the ACGME milestones.
Mapping Direct Observations From Objective Structured Clinical Examinations to the Milestones Across Specialties


BACKGROUND:
Little is known about residents' performance on the milestones at the institutional level. Our institution formed a work group to explore this using an institutional-level curriculum and residents' evaluation of the milestones.

OBJECTIVE:
We assessed whether beginner-level milestones for interpersonal and communication skills (ICS) related to observable behaviors in ICS-focused objective structured clinical examinations (OSCEs) for postgraduate year (PGY) 1 residents across specialties.

METHODS:
The work group compared ICS subcompetencies across 12 programs to identify common beginner-level physician-patient communication milestones. The selected ICS milestone sets were compared for common language with the ICS-OSCE assessment tool—the Kalamazoo Essential Elements of Communication Checklist-Adapted (KEECC-A). To assess whether OSCE scores related to ICS milestone scores, all PGY-1 residents from programs that were part of Next Accreditation System Phase 1 were identified; their OSCE scores from July 2013 to June 2014 and ICS subcompetency scores from December 2014 were compared.

RESULTS:
The milestones for 10 specialties and the transitional year had at least 1 ICS subcompetency that related to physician-patient communication. The language of the ICS beginner-level milestones appears similar to behaviors outlined in the KEECC-A. All 60 residents with complete data received at least a beginner-level ICS subcompetency score and at least a satisfactory score on all 3 OSCEs.

CONCLUSIONS:
The ICS-OSCE scores for PGY-1 residents appear to relate to beginner-level milestones for physician-patient communication across multiple specialties.
Pathology Informatics Essentials for Residents: A Flexible Informatics Curriculum Linked to Accreditation Council for Graduate Medical Education Milestones (a secondary publication)


BACKGROUND:
Recognition of the importance of informatics to the practice of pathology has surged. Training residents in pathology informatics has been a daunting task for most residency programs in the United States because faculty often lacks experience and training resources. Nevertheless, developing resident competence in informatics is essential for the future of pathology as a specialty.

OBJECTIVE:
To develop and deliver a pathology informatics curriculum and instructional framework that guides pathology residency programs in training residents in critical pathology informatics knowledge and skills, and meets Accreditation Council for Graduate Medical Education Informatics Milestones.

DESIGN:
The College of American Pathologists, Association of Pathology Chairs, and Association for Pathology Informatics formed a partnership and expert work group to identify critical pathology informatics training outcomes and to create a highly adaptable curriculum and instructional approach, supported by a multiyear change management strategy.

RESULTS:
Pathology Informatics Essentials for Residents (PIER) is a rigorous approach for educating all pathology residents in important pathology informatics knowledge and skills. PIER includes an instructional resource guide and toolkit for incorporating informatics training into residency programs that vary in needs, size, settings, and resources. PIER is available at http://www.apcprods.org/PIER (accessed April 6, 2016).

CONCLUSIONS:
PIER is an important contribution to informatics training in pathology residency programs. PIER introduces pathology trainees to broadly useful informatics concepts and tools that are relevant to practice. PIER provides residency program directors with a means to implement a standardized informatics training curriculum, to adapt the approach to local program needs, and to evaluate resident performance and progress over time.
Direct Observation Assessment of Milestones: Problems with Reliability


INTRODUCTION:
Emergency medicine (EM) milestones are used to assess residents' progress. While some milestone validity evidence exists, there is a lack of standardized tools available to reliably assess residents. Inherent to this is a concern that we may not be truly measuring what we intend to assess. The purpose of this study was to design a direct observation milestone assessment instrument supported by validity and reliability evidence. In addition, such a tool would further lend validity evidence to the EM milestones by demonstrating their accurate measurement.

METHODS:
This was a multi-center, prospective, observational validity study conducted at eight institutions. The Critical Care Direct Observation Tool (CDOT) was created to assess EM residents during resuscitations. This tool was designed using a modified Delphi method focused on content, response process, and internal structure validity. Paying special attention to content validity, the CDOT was developed by an expert panel, maintaining the use of the EM milestone wording. We built response process and internal consistency by piloting and revising the instrument. Raters were faculty who routinely assess residents on the milestones. A brief training video on utilization of the instrument was completed by all. Raters used the CDOT to assess simulated videos of three residents at different stages of training in a critical care scenario. We measured reliability using Fleiss' kappa and interclass correlations.

RESULTS:
Two versions of the CDOT were used: one used the milestone levels as global rating scales with anchors, and the second reflected a current trend of a checklist response system. Although the raters who used the CDOT routinely rate residents in their practice, they did not score the residents' performances in the videos comparably, which led to poor reliability. The Fleiss' kappa of each of the items measured on both versions of the CDOT was near zero.

CONCLUSION:
The validity and reliability of the current EM milestone assessment tools have yet to be determined. This study is a rigorous attempt to collect validity evidence in the development of a direct observation assessment instrument. However, despite strict attention to validity evidence, inter-rater reliability was low. The potential sources of reducible variance include rater- and instrument-based error. Based on this study, there may be concerns for the reliability of other EM milestone assessment tools that are currently in use.
PURPOSE:
Clinical competency committees (CCCs) are now required in graduate medical education. This study examined how residency programs understand and operationalize this mandate for resident performance review.

METHOD:
In 2013, the investigators conducted semi-structured interviews with 34 residency program directors at five public institutions in California, asking about each institution's CCCs and resident performance review processes. They used conventional content analysis to identify major themes from the verbatim interview transcripts.

RESULTS:
The purpose of resident performance review at all institutions was oriented toward one of two paradigms: a problem identification model, which predominated; or a developmental model. The problem identification model, which focused on identifying and addressing performance concerns, used performance data such as red-flag alerts and informal information shared with program directors to identify struggling residents. In the developmental model, the timely acquisition and synthesis of data to inform each resident's developmental trajectory was challenging. Participants highly valued CCC members' expertise as educators to corroborate the identification of struggling residents and to enhance credibility of the committee's outcomes. Training in applying the milestones to the CCC's work was minimal. Participants were highly committed to performance review and perceived the current process as adequate for struggling residents but potentially not for others.

CONCLUSIONS:
Institutions orient resident performance review toward problem identification; a developmental approach is uncommon. Clarifying the purpose of resident performance review and employing efficient information systems that synthesize performance data and engage residents and faculty in purposeful feedback discussions could enable the meaningful implementation of milestones-based assessment.
Implementation of Nephrology Subspecialty Curricular Milestones


**ABSTRACT**

Beginning in the 2014-2015 training year, the US Accreditation Council for Graduate Medical Education (ACGME) required that nephrology Clinical Competency Committees assess fellows' progress toward 23 subcompetency "context nonspecific" internal medicine subspecialty milestones. Fellows' advancement toward the "ready for unsupervised practice" target milestone now is tracked in each of the 6 competencies: Patient Care, Medical Knowledge, Professionalism, Interpersonal Communication Skills, Practice-Based Learning and Improvement, and Systems-Based Practice. Nephrology program directors and subspecialty societies must define nephrology-specific "curricular milestones," mapped to the nonspecific ACGME milestones. Although the ACGME goal is to produce data that can discriminate between successful and underperforming training programs, the approach is at risk to produce biased, inaccurate, and unhelpful information. We map the ACGME internal medicine subspecialty milestones to our previously published nephrology-specific milestone schema and describe entrustable professional activities and other objective assessment tools that inform milestone decisions. Mapping our schema onto the ACGME subspecialty milestone reporting form allows comparison with the ACGME subspecialty milestones and the curricular milestones developed by the American Society of Nephrology Program Directors. Clinical Competency Committees may easily adapt and directly translate milestone decisions reached using our schema onto the ACGME internal medicine subspecialty competency milestone-reporting format.
The New Milestones: Do We Need to Take a Step Back to Go a Mile Forward?


ABSTRACT
The Milestones Project, like all previous systems and changes in graduate psychiatric education, for example, moving from 3 to 4 years of training or adopting six competency domains, has been devised without any supporting data and does not assess meaningful outcomes, such as improved patient outcomes. No evidence is presented that Milestones-based training will produce better psychiatrists. There is a path forward. First, replace unproven expert consensus with scientific and evidence-based approaches. Second, exchange endpoints that are easy to assess but uncorrelated with real world functioning (e.g., multiple-choice examinations) for outcomes that are meaningful and external to the training program (e.g., patient outcomes). Finally, to prevent possible waste, excess burden, or harm, no changes should be mandated until proven in prospective studies.
The Milestones for Psychosomatic Medicine Subspecialty Training


BACKGROUND:
The Accreditation Council of Graduate Medical Education Milestones project is a key element in the Next Accreditation System for graduate medical education. On completing the general psychiatry milestones in 2013, the Accreditation Council of Graduate Medical Education began the process of creating milestones for the accredited psychiatric subspecialties.

METHODS:
With consultation from the Academy of Psychosomatic Medicine, the Accreditation Council of Graduate Medical Education appointed a working group to create the psychosomatic medicine milestones, using the general psychiatry milestones as a starting point.

RESULTS:
This article represents a record of the work of this committee. It describes the history and rationale behind the milestones, the development process used by the working group, and the implications of these milestones on psychosomatic medicine fellowship training.

CONCLUSIONS:
The milestones, as presented in this article, will have an important influence on psychosomatic medicine training programs. The implications of these include changes in how fellowship programs will be reviewed and accredited by the Accreditation Council of Graduate Medical Education and changes in the process of assessment and feedback for fellows.
ABSTRACT
The article offers information on the implementation of the Milestones program outlined by the Accreditation Council on Graduate Medical Education (ACGME) which provides a detailed framework for determining residency knowledge within specific core competencies in Orthopaedic medical education. Topics discussed include a residency program initiated by surgeon, William S. Halsted, development of formal orthopaedic postgraduate education, and expansion of the education after World War II.
Purpose
Milestones for the assessment of residents in graduate medical education mark a change in our evaluation paradigms. The Accreditation Council for Graduate Medical Education has created milestones and defined them as significant points in development of a resident based on the 6 competencies. We propose that a similar approach be taken for resident assessment of teaching faculty. We believe this will establish parity and objectivity for faculty evaluation, provide improved data about attending surgeons' teaching, and standardize faculty evaluations by residents.

Methods
A small group of advanced surgery educators determined appropriate educational characteristics, resulting in creation of 11 milestones (Fig. 2) that were reviewed by faculty and residents. The residents have historically answered 16 questions, developed by our surgical education committee (Fig. 3), on a 5-point Likert score (never to very often). Three weeks after completing this Likert-type evaluation, the residents were asked to again evaluate attending faculty using the Faculty Milestones evaluation. The residents then completed a survey of 7 questions (scale of 1-9—disagree to strongly agree, neutral = 5), assessing the new milestones and compared with the previous Likert evaluation system.

Results
Of 32 surgery residents, 13 completed the Likert evaluations (3760 data points) and 13 completed the milestones evaluations (1800 data points). The number completing both or neither is not known, as the responses are anonymous when used for faculty feedback. The Faculty Milestones attending physicians' scores have far fewer top of range scores (21% vs 42%) and have a wider spread of data giving better indication of areas for improvement in teaching skills. The residents completed 17 surveys (116 responses) to evaluate the new milestones system. Surveys indicated that milestones were easier to use (average rating 6.13 ± 0.42 Standard Error (SE)), effective (6.82 ± 0.39) and efficient (6.11 ± 0.53), and more objective (6.69 ± 0.39/6.75 ± 0.38) than the Likert evaluations are. Average response was 6.47 ± 0.46 for overall satisfaction with the Faculty Milestones evaluation. More surveys were completed than evaluations, as all residents had an opportunity to review both evaluation systems.

Conclusions
Faculty Milestones are more objective in evaluating surgical faculty and mirror the new paradigm in resident evaluations. Residents found this was an easier, more effective, efficient, and objective evaluation of our faculty. Although our Faculty Milestones are designed for surgical educators, they are likely to be applicable with appropriate modifications to other medical educators as well.
Cytopathology Fellowship Milestones


ABSTRACT
The American Society of Cytopathology has provided guidelines for goals and objectives for cytopathology fellows. There are 90 Accreditation Council for Graduate Medical Education- accredited cytopathology fellowship training programs in the United States, each with its own unique curriculum designed to achieve these goals and objectives. The Accreditation Council for Graduate Medical Education cytopathology fellowship milestones were developed to ensure some uniformity in the outcomes of the various skill sets and competencies expected of a graduating cytopathology fellow. The rationale, development, and details of the cytopathology fellowship milestones are described herein.
Survey of Developmental Milestones in Internal Medicine among Residents and Faculty


OBJECTIVES:
The published Accreditation Council for Graduate Medical Education (ACGME) milestones represent a novel method of evaluation of trainees in graduate medical education. We surveyed a group of teaching faculty and residents, regarding the new ACGME milestones project. We obtained their input on the expected timeline for the developmental milestones and compared their responses to the ACGME recommendations.

METHODS:
A 42-item survey questionnaire, derived from the original 142 item publication, was completed by 26 internal medicine teaching faculty and 34 internal medicine residents.

RESULTS:
We found statistically significant differences in the responses given by residents and faculty compared to those in the standard recommendations. The differences were more pronounced with the residents than with the faculty.

CONCLUSIONS:
The results of our survey showed significantly different responses as compared to the standard recommended timelines. Since this is a novel evaluation process, substantial faculty development and resident education regarding the process can help improve its implementation. Future studies should focus on how learners might better understand and refine the milestone evaluation process.
Progress toward Improved Leadership and Management Training in Pathology


CONTEXT:
Competency gaps in leadership and laboratory management skills continue to exist between what training programs deliver and what recent graduates and future employers expect. A number of recent surveys substantiate this. Interest in delivering content in these areas is challenged by time constraints, the presence of knowledgeable faculty role models, and the necessary importance placed on diagnostic skills development, which overshadows any priority trainees have toward developing these skills.

OBJECTIVE:
To describe the problem, the near-future horizon, the current solutions, and the recommendations for improving resident training in laboratory management.

DATA SOURCES:
The demands of new health care delivery models and the value being placed on these skills by the Pathology Milestones and Next Accreditation System initiative of the Accreditation Council for Graduate Medical Education for training programs emphasizes their importance. This initiative includes 6 milestone competencies in laboratory management. Organizations like the American Society for Clinical Pathology, the American Pathology Foundation, the College of American Pathologists, and the Association of Pathology Chairs Program Directors Section recognize these competencies and are working to create new tools for training programs to deploy.

CONCLUSIONS:
It is our recommendation that (1) every training program develop a formal educational strategy for management training, (2) greater opportunity and visibility be afforded for peer-reviewed publications on management topics in mainstream pathology literature, and (3) pathology milestones-oriented tools be developed to assist program directors and their trainees in developing this necessary knowledge and skills.
The Pathology Milestones and the Next Accreditation System


BACKGROUND:
In the late 1990s, the Accreditation Council for Graduate Medical Education developed the Outcomes Project and the 6 general competencies with the intent to improve the outcome of graduate medical education in the United States. The competencies were used as the basis for developing learning goals and objectives and tools to evaluate residents' performance. By the mid-2000s the stakeholders in resident education and the general public felt that the Outcomes Project had fallen short of expectations.

OBJECTIVE:
To develop a new evaluation method to track trainee progress throughout residency using benchmarks called milestones. A change in leadership at the Accreditation Council for Graduate Medical Education brought a new vision for the accreditation of training programs and a radically different approach to the evaluation of residents.

DATA SOURCES:
The Pathology Milestones Working Group reviewed examples of developing milestones in other specialties, the literature, and the Accreditation Council for Graduate Medical Education program requirements for pathology to develop pathology milestones. The pathology milestones are a set of objective descriptors for measuring progress in the development of competency in patient care, procedural skill sets, medical knowledge, practice-based learning and improvement, interpersonal and communication skills, professionalism, and systems-based practice.

CONCLUSIONS:
The milestones provide a national standard for evaluation that will be used for the assessment of all residents in Accreditation Council for Graduate Medical Education-accredited pathology training programs.
The "Zing Factor" – How Do Faculty Describe the Best Pediatrics Residents?


BACKGROUND:
Faculty in graduate medical education programs may not have uniform approaches to differentiating the quality of residents, and reviews of evaluations suggest that faculty use different standards when assessing residents. Standards for assessing residents also do not consistently map to items on evaluation forms. One way to improve assessment is to reach consensus on the traits and behaviors that are (or should be) present in the best residents.

METHODS:
A trained interviewer conducted semistructured interviews with faculty affiliated with 2 pediatrics residency programs until content saturation was achieved. Interviewees were asked to describe specific traits present in residents they identify as the best. Interviews were recorded and transcribed. We used an iterative, inductive approach to generate a coding scheme and identify common themes.

RESULTS:
From 23 interviews, we identified 7 thematic categories of traits and behaviors: personality, energy, professionalism, team behaviors, self-improvement behaviors, patient-interaction behaviors, and medical knowledge and clinical skills (including a subcategory, knowledge integration). Most faculty interviewees focused on traits like passion, enthusiasm, maturity, and reliability. Examination score or intelligence was mentioned less frequently than traits and behaviors categorized under personality and professionalism.

CONCLUSIONS:
Faculty identified many traits and behaviors in the residents they define as the best. The thematic categories had incomplete overlap with Accreditation Council for Graduate Medical Education (ACGME) and CanMEDS competencies. This research highlights the ongoing need to review our assessment strategies, and may have implications for the ACGME Milestone Project.
Early Feedback on the Use of the Internal Medicine Reporting Milestones in Assessment of Resident Performance


BACKGROUND:
The educational milestones were designed as a criterion-based framework for assessing resident progression on the 6 Accreditation Council for Graduate Medical Education competencies.

OBJECTIVE:
We obtained feedback on, and assessed the construct validity and perceived feasibility and utility of, draft Internal Medicine Milestones for Patient Care and Systems-Based Practice.

METHODS:
All participants in our mixed-methods study were members of competency committees in internal medicine residency programs. An initial survey assessed participant and program demographics; focus groups obtained feedback on the draft milestones and explored their perceived utility in resident assessment, and an exit survey elicited input on the value of the draft milestones in resident assessment. Surveys were tabulated using descriptive statistics. Conventional content analysis method was used to assess the focus group data.

RESULTS:
Thirty-four participants from 17 programs completed surveys and participated in 1 of 6 focus groups. Overall, the milestones were perceived as useful in formative and summative assessment of residents. Participants raised concerns about the length and complexity of some draft milestones and suggested specific changes. The focus groups also identified a need for faculty development. In the exit survey, most participants agreed that the Patient Care and Systems-Based Practice Milestones would help competency committees assess trainee progress toward independent practice.

CONCLUSIONS:
Draft reporting milestones for 2 competencies demonstrated significant construct validity in both the content and response process and the perceived utility for the assessment of resident performance. To ensure success, additional feedback from the internal medicine community and faculty development will be necessary.
The Emergency Medicine Milestones: A Validation Study


OBJECTIVES:
The Accreditation Council for Graduate Medical Education (ACGME) and the American Board of Medical Specialties sought to define milestones for skill and knowledge acquisition during residency training. Milestones are significant objective observable events. The milestones are listed within a structure that is derived from the ACGME general competencies. Major groups of milestones are called "subcompetencies." The original 24 subcompetencies containing 255 milestones for emergency medicine (EM) were developed through a multiorganizational group representing most EM stakeholder groups. To assure that the milestones reflected EM resident progress throughout training, the EM Milestones Working Group (EM MWG) sought to validate the individual milestones.

METHODS:
A computer-based survey was sent to all EM residency programs. The survey period began on April 30, 2012, and concluded on May 15, 2012. Respondents were asked to assign each milestone to a specific level of skill or knowledge acquisition. These levels ranged from a beginning resident to an accomplished clinician. There were two different forms that divided the milestones into two groups of 12 subcompetencies each. Surveys were randomly assigned to programs.

RESULTS:
There were five respondents (the program director and four key faculty) requested from each of the 159 residences. There were responses from 96 programs (60.4%). Of the 795 survey recipients, 28 were excluded due to prior exposure to the EM milestones. Of the remaining 767 potential respondents, 281 completed the survey (36.6%) within a 16-day period. Based on the survey results, the working group adjusted the milestones in the following ways: one entire subcompetency (teaching) was eliminated, six new milestones were created, 34 milestones were eliminated, 26 milestones were reassigned to a lower level score, and 20 were reassigned to a higher level. Nineteen milestones were edited to provide greater clarity. The final result was 227 discrete milestones among 23 subcompetencies.

CONCLUSIONS:
The EM milestones were validated through a milestone assignment process using a computer-based survey completed by program directors and key faculty. Milestones were revised in accordance with the results to better align assignment within each performance level.
Educational Milestone Development in the First 7 Specialties to Enter the Next Accreditation System


BACKGROUND:
The Accreditation Council for Graduate Medical Education (ACGME) Outcome Project introduced 6 general competencies relevant to medical practice but fell short of its goal to create a robust assessment system that would allow program accreditation based on outcomes. In response, the ACGME, the specialty boards, and other stakeholders collaborated to develop educational milestones, observable steps in residents' professional development that describe progress from entry to graduation and beyond.

OBJECTIVES:
We summarize the development of the milestones, focusing on 7 specialties, moving to the next accreditation system in July 2013, and offer evidence of their validity.

METHODS:
Specialty workgroups with broad representation used a 5-level developmental framework and incorporated information from literature reviews, specialty curricula, dialogue with constituents, and pilot testing.

RESULTS:
The workgroups produced richly diverse sets of milestones that reflect the community's consideration of attributes of competence relevant to practice in the given specialty. Both their development process and the milestones themselves establish a validity argument, when contemporary views of validity for complex performance assessment are used.

CONCLUSIONS:
Initial evidence for validity emerges from the development processes and the resulting milestones. Further advancing a validity argument will require research on the use of milestone data in resident assessment and program accreditation.
OBJECTIVE:
As the next step in competency-based medical education, the Pediatrics Milestone Project seeks to provide a learner-centered approach to training and assessment. To help accomplish this goal, this study sought to determine how pediatric residents understand, interpret, and respond to the Pediatrics Milestones.

METHODS:
Cognitive interviews with 48 pediatric residents from all training levels at 2 training programs were conducted. Each participant reviewed one Pediatrics Milestone document (PMD). Eight total Pediatrics Milestones, chosen for their range of complexity, length, competency domain, and primary author, were included in this study. Six residents, 2 from each year of residency training, reviewed each PMD. Interviews were transcribed and coded using inductive methods, and codes were grouped into themes that emerged.

RESULTS:
Four major themes emerged through coding and analysis: 1) the participants' degree of understanding of the PMDs is sufficient, often deep; 2) the etiology of participants' understanding is rooted in their experiences; 3) there are qualities of the PMD that may contribute to or detract from understanding; and 4) participants apply their understanding by noting the PMD describes a developmental progression that can provide a road map for learning. Additionally, we learned that residents are generally comfortable being placed in the middle of a series of developmental milestones. Two minor themes focusing on interest and practicality were also identified.

CONCLUSIONS:
This study provides initial evidence for the Pediatrics Milestones as learner-centered documents that can be used for orientation, education, formative feedback, and, ultimately, assessment.
Milestones for Apheresis Education


ABSTRACT
Milestones represent the essential knowledge, skills, and attitudes required for the practice of a medical discipline. Defining these milestones for each medical specialty has become a focus for the American Council of Graduate Medical Education (ACGME). Practitioners of Apheresis Medicine come from a variety of medical specialties making it challenging to establish the essential educational milestones for all. The American Society for Apheresis (ASFA) has an interest in promoting standards of excellence for Apheresis Medicine. ASFA's Physician's Curriculum Content Committee is a group of physician educators in the field of Apheresis Medicine, both donor and therapeutic, from across the United States, who have met regularly for several years to discuss the appropriate educational milestones in Apheresis training. The committee members teach residents and fellows from Pathology, Transfusion Medicine, Hematology/Oncology, Nephrology and other specialties. In this document, we have outlined the basic set of Apheresis milestones required in the ACGME defined competency areas of Patient Care and Medical Knowledge. We have also recommended methods of evaluation and estimated the time necessary for the acquisition of these cognitive and behavioral elements.
From Theory to Actual Practice: Creation and Application of Milestones in an Internal Medicine Residency Program, 2004-2010


BACKGROUND:
In the USA, the Accreditation Council of Graduate Medical Education, Educational Innovations Project is a partner in reshaping residency training to meet increasingly complex systems of health care delivery.

AIM:
We describe the creation and implementation of milestones as a vehicle for translating educational theory into practice in preparing residents to provide safe, autonomous patient care.

METHOD:
Six program faculty leaders, all with advanced medical education training, met in an iterative process of developing, implementing, and modifying milestones until a final set were vetted.

RESULTS:
We first formed the profile of a Master Internist. We then translated it into milestone language and implemented its integration across the program. Thirty-seven milestones were applied in all settings and rotations to reach explicit educational outcomes. We created three types of milestones: Progressive, build one on top of the other to mastery; additive, adding multiple behaviors together to culminate in mastery; and descriptive, using a proscribe set of complex, predetermined steps toward mastery.

CONCLUSIONS:
Using milestones, our program has enhanced an educational model into explicit, end of training goals. Milestone implementation has yielded positive results toward competency-based training and others may adapt our strategies in a similar effort.
How Residents Can Use Milestones Data
A Cross-Specialty Examination of Resident Error Disclosure and Communication Skills Using Simulation


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


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


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


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


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


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.
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


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


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 $\alpha$ 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 $\alpha$: 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


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


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


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.
Appendix:

Original Description of the Milestones by Specialty (Journal of Graduate Medical Education (JGME) supplementary issue, March 2013)

A Milestone in the Milestones Movement: the JGME Milestones Supplement
Gail Sullivan, Deborah Simpson, Thomas Cooney, and Eugene Beresin
Journal of Graduate Medical Education 2013 5:1s1, 1-4

Emergency Medicine Milestones
Journal of Graduate Medical Education 2013 5:1s1, 5-13

Internal Medicine Milestones
William Iobst, Eve Aagaard, Hasan Bazari, Timothy Brigham, Roger W. Bush, Kelly Caverzagie, Davoren Chick, Michael Green, Kevin Hinchey, Eric Holmboe, Sarah Hood, Gregory Kane, Lynne Kirk, Lauren Meade, Cynthia Smith, and Susan Swing
Journal of Graduate Medical Education 2013 5:1s1, 14-23

Neurological Surgery Milestones
Journal of Graduate Medical Education 2013 5:1s1, 24-35

Orthopaedic Surgery Milestones
Journal of Graduate Medical Education 2013 5:1s1, 36-58

Pediatrics Milestones
Carol Carraccio, Bradley Benson, Ann Burke, Robert Englander, Susan Guralnick, Patricia Hicks, Stephen Ludwig, Daniel Schumacher, and Jerry Vasilias
Journal of Graduate Medical Education 2013 5:1s1, 59-73

Diagnostic Radiology Milestones
Journal of Graduate Medical Education 2013 5:1s1, 74-78

Urology Milestones
Michael Coburn, Christopher Amling, Robert R. Bahnson, Philipp Dahm, B. Price Kerfoot, Louise King, Brian Lane, Michael L. Ritchey, Charles D. Scales, Jr, Chandru P. Sundaram, and Susan R. Swing
Journal of Graduate Medical Education 2013 5:1s1, 79-98

©2019 Accreditation Council for Graduate Medical Education (ACGME)
Original Description of the Milestones by Specialty *(Journal of Graduate Medical Education (JGME) supplementary issue, March 2014)*

**Formulating the Allergy and Immunology Milestones**
Michael R. Nelson
Journal of Graduate Medical Education 2014 6:1s1, 1-4

**The Allergy and Immunology Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 5-11

**Development of the Anesthesiology Educational Milestones**
Scott A. Schartel, Catherine Kuhn, Deborah J. Culley, Margaret Wood, and Neal Cohen
Journal of Graduate Medical Education 2014 6:1s1, 12-14

**The Anesthesiology Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 15-28

**The Development of the Colon and Rectal Surgery Milestones**
Charles B. Whitlow
Journal of Graduate Medical Education 2014 6:1s1, 29-30

**The Colon and Rectal Surgery Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 31-43

**The Dermatology Milestone Project**
George W. Turiansky and Daniel Loo
Journal of Graduate Medical Education 2014 6:1s1, 44-46

**The Dermatology Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 47-70

**Development of the Family Medicine Milestones**
Suzanne Allen
Journal of Graduate Medical Education 2014 6:1s1, 71-73

**The Family Medicine Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 74-86

**The Medical Genetics Residency Milestones**
Bruce R. Korf
Journal of Graduate Medical Education 2014 6:1s1, 87-90

**The Medical Genetics Milestone Project**
Journal of Graduate Medical Education 2014 6:1s1, 91-101

**Introducing the Neurology Milestones**
Steven L. Lewis, Ralph F. Józefowicz, Shannon Kilgore, Amar Dhand, and Laura Edgar
Journal of Graduate Medical Education 2014 6:1s1, 102-104
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Journal</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Neurology Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 105-115</td>
</tr>
<tr>
<td>Nuclear Medicine Milestones</td>
<td>Lorraine M. Fig and Laura Edgar</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 116-118</td>
</tr>
<tr>
<td>The Nuclear Medicine Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 119-125</td>
</tr>
<tr>
<td>Obstetrics and Gynecology Milestones</td>
<td>Jessica L. Bienstock, Laura Edgar, and Rebecca McAlister</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 126-128</td>
</tr>
<tr>
<td>The Obstetrics and Gynecology Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 129-143</td>
</tr>
<tr>
<td>Developing the Educational Milestones for Ophthalmology</td>
<td>Anthony Arnold</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 144-145</td>
</tr>
<tr>
<td>The Ophthalmology Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 146-161</td>
</tr>
<tr>
<td>Developing the Otolaryngology Milestones</td>
<td>Terance T. Tsue</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 162-165</td>
</tr>
<tr>
<td>The Otolaryngology Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 166-179</td>
</tr>
<tr>
<td>Pathology Milestones</td>
<td>Wesley Y. Naritoku and C. Bruce Alexander</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 180-181</td>
</tr>
<tr>
<td>The Pathology Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 182-203</td>
</tr>
<tr>
<td>The Development of the Physical Medicine and Rehabilitation Milestones</td>
<td>William L. Bockenek, Teresa L. Massagli, Susan R. Swing, and Caroline Fischer</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 204-206</td>
</tr>
<tr>
<td>The Physical Medicine and Rehabilitation Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 207-221</td>
</tr>
<tr>
<td>The Plastic Surgery Milestone Project</td>
<td>Mary H. McGrath</td>
<td>Journal of Graduate Medical Education</td>
<td>2014 6:1s1, 222-224</td>
</tr>
<tr>
<td>The Plastic Surgery Milestone Project</td>
<td></td>
<td></td>
<td>2014 6:1s1, 225-242</td>
</tr>
</tbody>
</table>

©2019 Accreditation Council for Graduate Medical Education (ACGME)