# Milestones: a rapid assessment method for the Clinical Competency Committee

Christopher Nabors[1], Leanne Forman[1], Stephen J. Peterson[2], Melissa Gennarelli[1], Wilbert S. Aronow[1], Lawrence DeLorenzo[3], Dipak Chandy[3], Chul Ahn[4], Sachin Sule[1], Gary W. Stallings[1], Sahil Khera[1], Chandrasekar Palaniswamy[1], William H. Frishman[1]

[1]Department of Medicine, Division of Cardiology, New York Medical College at Westchester Medical Center, Valhalla, New York, USA
[2]Department of Medicine, New York Methodist Hospital, Brooklyn, NY, USA
[3]Department of Medicine, Division of Pulmonary and Critical Care, New York Medical College at Westchester Medical Center, Valhalla, New York, USA
[4]Department of Clinical Sciences, UT Southwestern Medical Center, Dallas, Texas, USA

**Corresponding author:**
Christopher Nabors
Wilbert S. Aronow MD
Department of Medicine
Division of Cardiology
New York Medical College
at Westchester
Medical Center
Macy Pavilion
Room 138
10595 Valhalla
New York, USA
Phone: 011914 493 5311
Fax: 011914 235 6274
E-mail: wsaronow@aol.com

## Abstract

**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.

**Key words:** educational milestones, clinical competency committee.

## Introduction

All US residency training programs are required to assess the development of their trainees using educational milestones [1]. Twice annually each program's Clinical Competency Committee (CCC) must review and report each resident's progress to the Accreditation Council for Graduate

C. Nabors, L. Forman, S.J. Peterson, M. Gennarelli, W.S. Aronow, L. DeLorenzo, D. Chandy, C. Ahn, S. Sule, G.W. Stallings, S. Khera, C. Palaniswamy, W.H. Frishman

Medical Education (ACGME). Certain features of the process are mandated by the ACGME, while others are left to the discretion of individual programs. The ideal way to make these determinations is not currently known.

Our internal medicine residency gained early experience with milestones as a participant in the Educational Innovations Project (EIP) [2, 3]. Our initial work focused on a set of detailed or "curricular milestones" that were released to the internal medicine community in 2009 [4, 5]. Prior to implementation of the Next Accreditation System (NAS), we pilot tested a second set of milestones that are now used for ACGME reporting. We (and others) found their use to be time intensive. This prompted us to devise and test a new approach which we hypothesized would permit rapid but accurate clinical competency assessments. This report describes the new approach.

## Material and methods

The study took place as an EIP initiative within a medium-sized residency program at an academic medical center [6]. Participants included academic files of 13 categorical residents (then interns), a clinical competency committee chair, three program leaders and six faculty members. We modified our usual clinical competency process to incorporate several new elements and tested the new approach during a special clinical competency session in 2013. Outcomes included pre- and post-intervention deliberation times, participant survey results and comparison of assessments made by individual committee members and the collective committee using class rank lists.

### The intervention

Our usual CCC included a chair, program director and eight faculty members. Meetings lasted 2–3 h, during which one class of 13 residents was evaluated. Faculty reviewed portfolios, *de novo*, in the meeting and presented evaluation data and other information to the group. Reference to other information not necessarily in the portfolio helped to inform decisions.

During this study, the committee composition was similar. However, six faculty members were assigned to serve as "presenters." As such, they were tasked with reviewing a standard data report for 2–3 assigned subjects prior to the meeting and rating their progress along the 22 internal medicine milestone subcompetencies. The program director did the same for each member of the class. At the meeting, the presenters announced the subcompetency ratings (and underlying reasoning) for each of their assigned subjects in sequence. The program director provided his rating for the same subcompetency, and through negotiated consensus the group reconciled any discrepancies between the committee member and program director's ratings to achieve final scores for each subcompetency. Other information supplemented the discussions as needed. Key distinctions between the traditional and new approaches are displayed in Table I.

### Evaluation data

Standard data reports used by the CCC were generated in Microsoft Access (Microsoft Corp, Mountain View, CA) using data exported from New Innovations (New Innovations, Uniontown, Ohio). One report was generated for each study subject. Data were organized by core competency, milestone subcompetency, curricular milestone, rotation type and time frame and evaluator type. Figure 1. Characteristics of the resident evaluation data set are displayed in Table II. Figure 2 displays the CCC data consideration process.

Table I. Clinical competency deliberations

| Feature | Usual deliberation process | Study deliberation process |
| --- | --- | --- |
| Data sources | Custom Data Report in tandem with portfolio | Custom data report as primary source with portfolio as backup |
| Data review timing | In meeting | Pre-meeting emphasis with less on meeting review |
| Data presenters | Random Faculty Committee members | Assigned Faculty Committee members and Program Director |
| Data presentation | Entire portfolio:<br>• Curricular milestones<br>• Non-milestones<br>• General competency | Milestone subcompetency scores and basis therefor; less emphasis on remainder of portfolio |
| Analytic framework | General competencies | Milestones subcompetencies |
| Analytic process | General discussion | Reconciliation of reviewer ratings supplemented by other discussion |

**Subject 1**

program start date:-/-/20--

| Core Competency - Milestone Competency | N | Faculty ONLY | Std dev | N | nurse, peer -other | Std dev | Self |
|---|---|---|---|---|---|---|---|
| **20** General Competency - Interpersonal and Communication Skills | | 4.76 | —Narrative avgs— | 5.50 | | | |

**Communicates effectively with patients and caregivers. (ICS1)**

| Applies evidence to decision-making for individual patients | | 4.76 | —competency avgs— | 5.50 | | | |
|---|---|---|---|---|---|---|---|
| September 2012 | Consults | 1 | 7.00 | | 0 | | |

level 3 - PBLI-E3  -  Communicate risks and benefits of alternatives to patients *

| Communicate effectively | | 4.76 | —competency avgs— | 5.50 | | | |
|---|---|---|---|---|---|---|---|
| October 2012 | Clinic | 6 | 3.00 | 1.55 | 0 | | |
| November 2012 | Floors | 6 | 4.33 | 0.82 | 4 | 5.50 | 2.89 |
| December 2012 | Floors | 3 | 5.00 | 0.00 | 0 | | |
| January 2013 | Floors | 3 | 7.33 | 0.58 | 0 | | |
| March 2013 | Clinic | 2 | 6.00 | 0.00 | 0 | | |

level 1 - ICS-A1  -  Provide timely and comprehensive verbal and written communication to patients/advocates **
level 1 - ICS-A2  -  Effectively use verbal and nonverbal skills to create rapport with patients/families **
level 1 - ICS-A3  -  Use communication skills to build a therapeutic relationship **
level 2 - ICS-A4  -  Engage patients/advocates in shared decision making for uncomplicated diagnostic and therapeutic scenarios  ***
level 2 - ICS-A5  -  Use patient-centered education strategies  ***
level 3 - ICS-A7  -  Appropriately counsel patients about the risks and benefits of tests and procedures, highlighting cost awareness and  * resource allocation

program start date:-/-/20--

| Core Competency - Milestone Competency | N | Faculty ONLY | Std dev | N | nurse, peer -other | Std dev | Self |
|---|---|---|---|---|---|---|---|
| **21** General Competency - Interpersonal and Communication Skills | | 5.07 | —Narrative avgs— | 5.25 | | | |

**Communicates effectively in interprofessional teams (e.g. peers, consultants, nursing, ancillary professionals and other support personnel). (ICS2)**

| Communitcate to interprofessional team | | 5.07 | —competency avgs— | 5.25 | | | |
|---|---|---|---|---|---|---|---|
| September 2012 | Consults | 2 | 7.00 | 0.00 | 0 | | |
| October 2012 | Clinic | 2 | 3.00 | 1.41 | 0 | | |
| November 2012 | Floors | 4 | 4.00 | 1.15 | 4 | 5.25 | 3.20 |
| December 2012 | Floors | 2 | 4.00 | 0.00 | 0 | | |
| December 2012 | Floors - Nights | 2 | 6.00 | 0.00 | 0 | | |
| January 2013 | Floors | 2 | 6.50 | 0.71 | 0 | | |
| January 2013 | Floors Teaching | 1 | 7.00 | | 0 | | |

level 1 - ICS-D1  -  Deliver appropriate, succinct, hypothesis-driven oral presentations  **
level 1 - ICS-D2  -  Effectively communicate plan of care to all members of the health care team  **
level 3 - ICS-D3  -  Engage in collaborative communication with all members of the health care team *

```
* Aspirational Milestone(s)
** Curricular Milestone language
   corresponds to Narrative Language
*** Mapped to Narrative by Program
**** Self Assessments excluded
```

*Rotation and General comments on last pages*

**Figure 1.** Sample Reporting Milestones Report Subcompetencies 20 and 21

### Deliberation times

Baseline CCC deliberation times for each subject were established during an ACGME/ABIM reporting milestones feasibility pilot. During this study, deliberation times were recorded by the program coordinator. Time spent by committee members and the program director during pre-meeting assessments was recorded by members and compiled by leadership.

### Class rank lists

We compared class rank lists made during this study to estimate reliability of competence determinations. A gold standard or "master" rank list

C. Nabors, L. Forman, S.J. Peterson, M. Gennarelli, W.S. Aronow, L. DeLorenzo, D. Chandy, C. Ahn, S. Sule, G.W. Stallings, S. Khera, C. Palaniswamy, W.H. Frishman

**Table II.** Characteristics of evaluation system

| Time frame – July 1, 2013 to March 1, 2013 | All evaluations | Milestone-based | Non-milestone-based |
|---|---|---|---|
| Total evaluation tools employed | 31 | 29 | 2 |
| Total evaluations for class | 710 | 513 | 197 |
| Total faculty evaluations for class | 228 | 180 | 48 |
| Percent of requested faculty evaluations completed | 228/288 (79%) | | |
| Directly observed faculty evaluations for class | 80 | 80 | 0 |
| Supervising peer evaluations (PGY-2 or PGY-3) | 31 | 31 | 0 |
| Same training level peer evaluations (PGY 1) | 135 | 0 | 135 |
| Nursing evaluations for class | 48 | 48 | 0 |
| Self-assessments for class | 87 | 87 | 0 |
| Patient satisfaction | 20 | 20 | 0 |
| Commendations or concerns | 6 | 0 | 6 |
| Clinical Competency Committee Evaluations | 13 | 13 | 0 |
| Synthetic Evals – Program Director Formative; Clinic Director ready for distance supervision | 26 | 26 | 0 |
| Documentation reviews (Progress Notes/ H&P, DC Summary) | 66 | 66 | 0 |
| Avg. No. of milestones used to rate each resident in PGY-1 year | 64 | | |
| Individual milestone-based ratings for class (excludes self-assessments and chart reviews) | 7608 | | |
| Avg. no. of milestone-based ratings per resident | 585 | | |

was generated by the CCC just after deliberations. Using open discussion to achieve consensus, the group rated each member of the subject class from most (rank 1) to least competent (rank 13). Competence was defined as aggregate effectiveness in the application of knowledge, skills, and attitudes within the scope of medical training across all possible settings. To the master list, we compared rank lists produced by individual presenters prior to any data review (pre-meeting) and after full deliberations (post-meeting), but prior to discussions which led to the master list. Additional rank lists were compiled from overall rotation evaluation scores (Faculty Rotation Evaluations) and from overall scores assigned to subcompetencies during the CCC (CCC Milestone Subcompetencies) (see Table III). Rank lists by two program leaders were excluded because of their substantial familiarity with the subjects' evaluations prior to the study.

### Participant surveys

At meeting close, non-leadership faculty (6/6) completed a voluntary and anonymous survey. To help establish content validity, items were patterned after a questionnaire used in a prior re-

port [3] and were piloted for clarity by a former program director. Focus sections included: 1) demographics; 2) comparison of milestones versus non-milestone based clinical competency deliberations; 3) effectiveness of data organization and analytic approaches.

### Statistical analysis

Survey results and deliberation times were analyzed using the Mann-Whitney $U$ test with significance accepted for $p < 0.05$. Survey results were not adjusted for multiple comparisons. Reliability estimates for competence assessments were made by comparing the CCC master rank list ordering with rank lists based on faculty pre-/post-meeting assessments, faculty ward ratings and aggregated CCC milestone subcompetency ratings using Bland-Altman analysis with significance accepted for $p < 0.05$. Associations between the master list and other rank lists were investigated using Spearman's rank correlation coefficient. Statistical analysis was conducted at the University of Texas Southwestern Medical Center using SAS software version 9.3 (SAS Institute, Cary, North Carolina). The project was approved by the Institutional Review Board of New York Med-
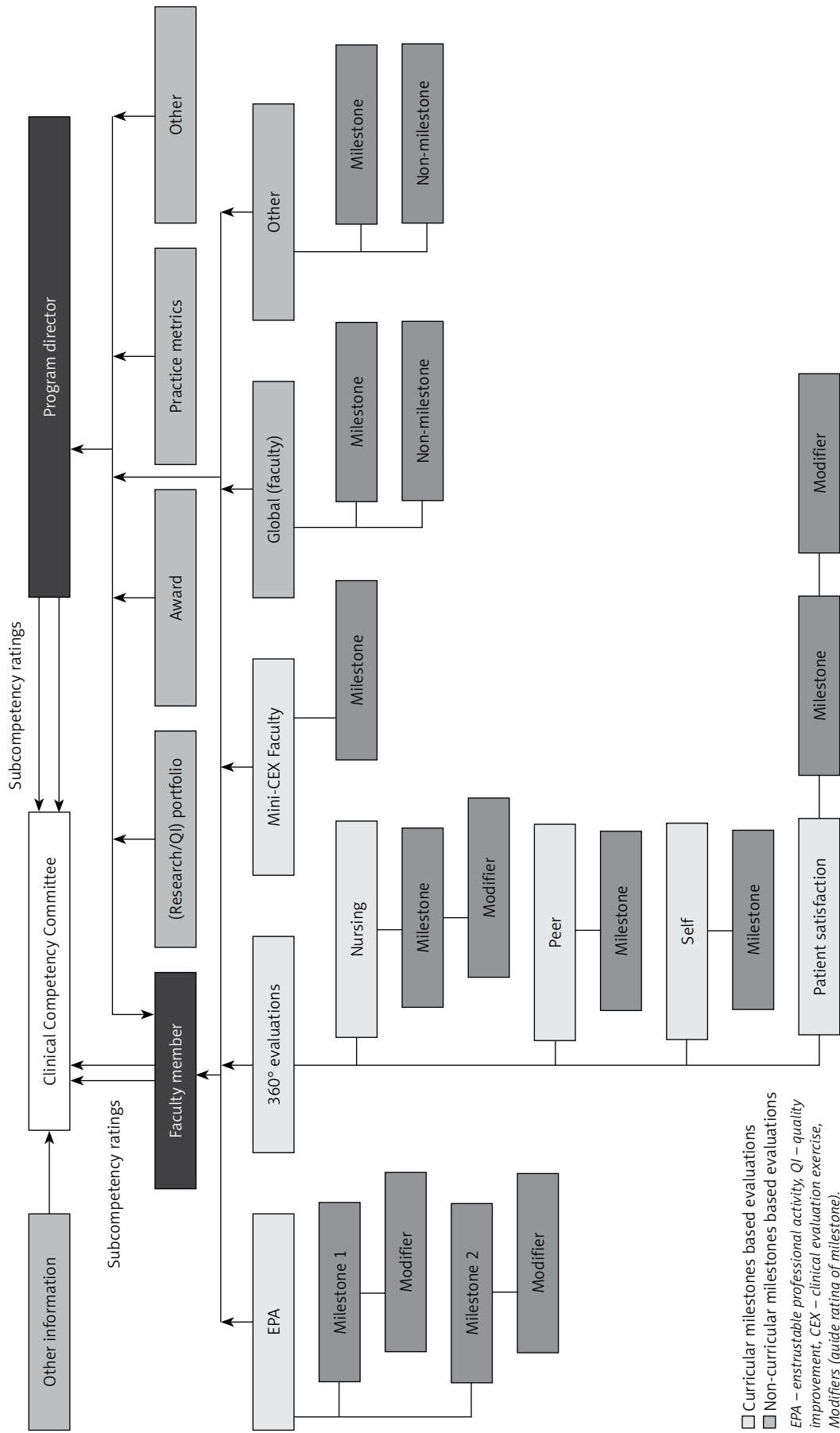
**Figure 2.** Study clinical competency data collection process

EPA – enstrustable professional activity, QI – quality improvement, CEX – clinical evaluation exercise, Modifiers (guide rating of milestone).

C. Nabors, L. Forman, S.J. Peterson, M. Gennarelli, W.S. Aronow, L. DeLorenzo, D. Chandy, C. Ahn, S. Sule, G.W. Stallings, S. Khera, C. Palaniswamy, W.H. Frishman

**Table III.** Comparison of rank lists

| Variable 1 | Variable 2 | Correlation | *P*-value |
|---|---|---|---|
| Master rank list | Faculty 1 post-meeting rank list | 0.74725 | 0.0033* |
| Master rank list | Faculty 1 pre-meeting rank list | 0.06593 | 0.8305 |
| Master rank list | Faculty 2 post-meeting rank list | 0.67582 | 0.0112* |
| Master rank list | Faculty 2 pre-meeting rank list | 0.45055 | 0.1223 |
| Master rank list | Faculty 3 post-meeting rank list | 0.58791 | 0.0346* |
| Master rank list | Faculty 3 pre-meeting rank list | 0.43956 | 0.1329 |
| Master rank list | Faculty 4 post-meeting rank list | 0.93407 | < 0.0001* |
| Master rank list | Faculty 4 pre-meeting rank list | 0.40110 | 0.1744 |
| Master rank list | Faculty rotation evaluations | 0.73626 | 0.0041* |
| Master rank list | CCC Milestone Subcompetencies | 0.76374 | 0.0024* |

*P-value < 0.05.*

ical College and the Office of Clinical Trials and Westchester Medical Center.

## Results

### Participant surveys

Participants rated overall satisfaction with clinical competency deliberations prior to and after milestones introduction at 6.0 and 7.7 ($p = 0.041$). Members found that use of milestones for CCC deliberations improved their ability to know and identify expected elements of competency development from 6.2 to 8.3 ($p = 0.015$), their ability to specifically assess a resident's competency development from 5.5 to 7.8 ($p = 0.026$) and their ability to identify particular strengths or weaknesses from 5.3 to 7.5 ($p = 0.026$) (Table IV).

Faculty rated their ability to rate competence at 6.2 using chart review alone, 7.3 using only data collected into a standard report and 8.3 using the study deliberation process which included pre-meeting review coupled with in-meeting committee reconciliation of disparate ratings (Table V). Committee members agreed strongly (2 of 6, 33%) or somewhat (4 of 6, 67%) that a standard data report facilitated deliberations; all agreed strongly (4 of 6, 67%) or somewhat (2 of 6, 33%) that such a report facilitated rapid CCC data analysis. All agreed somewhat (3 of 6, 50%) or strongly (3 of 6, 50%) that consideration of curricular milestone evaluations within the current milestone framework represented an effective evaluation strategy (Table VI).

### Deliberation time

Study CCC deliberation time for 13 subjects was 2 h and 5 min (9.6 min per subject). Using our prior approach, per subject deliberation time was 25 min. The difference in review time (15.4 min) was statistically significant ($p \leq 0.001$). During this study, presenters and the program director, respectively, required 32 min (mean) and 30 min pre-meeting time to rate subject milestone subcompetency achievement. Five man-hours per subject (25 min × 12 members) were required to generate milestone ratings using our usual approach, while 3 man-hours per subject (30 min faculty, 32 min program director, 10 min × 12 members) were required using the study approach.

### Clinical competency rank lists

Four faculty pre-/post-intervention class rank lists were compared with the committee's master list. While none of the four pre-meeting rank lists showed statistically significant agreement with the master rank list, each of the four post-meeting rank lists agreed with the master list (Table III). Significant correlation was also noted between the master list and class rank lists based on current academic year faculty rotation-evaluations and each subject's summed CCC's reporting milestone subcompetency scores.

## Discussion

In this study, CCC deliberations that were informed by a custom data report and guided by pre-meeting milestone assessments with in-meeting reconciliation produced rapid and reliable milestone ratings. Committee deliberation time was reduced from 25 min to less than 10 min per subject using the new method. Time savings resulted primarily from the committee's ability to focus on resolving discrepancies between program director and faculty competence determinations rather than on conducting a de novo data review.

**Table IV.** Clinical Competency Committee Member perceptions of milestones use (*N* = 6)

| Rate your level of satisfaction with clinical competency deliberation prior to and after introduction of Milestones (10 Maximally Satisfied, 1 Minimally Satisfied) | Mean | Range[a] | *P*-value |
|---|---|---|---|
| The Committee's format (milestones versus no milestones) for evaluating resident performance: | | | |
| Pre-milestones | 6 | 3–8 | 0.041 |
| Using milestones | 7.7 | 7–8 | |
| How data were presented to you for advancement decisions: | | | |
| Pre-milestones | 5.8 | 3–8 | |
| Using milestones | 7.5 | 6–9 | 0.093 |
| Your ability to know and identify expected elements of competency development (knowledge, skills and attitudes) for use in competency committee deliberations: | | | |
| Pre-milestones | 6.2 | 4–7 | 0.004 |
| Using milestones | 8.3 | 7–9 | |
| Your ability to specifically assess a resident's level of competency development (i.e., attainment of required knowledge, skills and attitudes): | | | |
| Pre-reporting milestones | 5.5 | 4–7 | 0.015 |
| Using milestones | 7.8 | 6–9 | |
| Your ability to identify residents ready for an accelerated training curriculum (complete training in 2 years rather than 3 years: | | | |
| Pre-milestones | 5 | 3–7 | 0.026 |
| Using milestones | 7.3 | 6–9 | |
| Your ability to describe and quantify differences in level of performance between house officers at the same level of training: | | | |
| Pre-milestones | 6 | 4–9 | 0.132 |
| Using milestones | 7.7 | 7–8 | |
| Your ability to identify particular strengths or weaknesses in the trainees' developmental progress: | | | |
| Pre-milestones | 5.3 | 3–7 | 0.026 |
| Using milestones | 7.5 | 6–9 | |

[a]*Ten-point scale where 10 = maximally satisfied, 1 = minimally satisfied.*

**Table V.** Data consideration method

| Rate how well you were able to assess trainee competence | Mean | Range[a] | *N* |
|---|---|---|---|
| Based on manual chart review | 6.2 | 5–8 | 6 |
| Using data from a standard milestones report | 7.3 | 5–10 | 6 |
| Based on pre-meeting review with in-meeting reconciliation | 8.3 | 7–10 | 6 |

[a]*Ten-point scale where 10 = very well, 1 = very poorly.*

Despite the reduced deliberation time, committee members believed the new process improved their ability to assess a resident's level of competency development and that consideration of evaluation data through a standard data report was useful. Support for these perceptions came from comparison of class rank lists produced during the study. Lists generated by faculty prior to data review lacked significant association with the master class rank list generated by the collective CCC. On the other hand, each of the faculty generated rank lists which followed the CCC deliberation process bore a close statistical association with the master list. This suggests that analysis within

the milestone framework permits development of a shared mental model [7, 8] of competence attainment during CCC deliberations.

Because CCC deliberations in the NAS rest on the application of the criterion-based milestone narratives, our finding of agreement between the CCC's milestone-based rank list and the committee's (Gestalt-based) master list was significant. This concordance between traditional and NAS-based deliberations provides preliminary evidence that application of the internal medicine milestone narratives permits evaluators to effectively discern differing developmental trajectories among trainees. Further work will be needed to verify this finding

C. Nabors, L. Forman, S.J. Peterson, M. Gennarelli, W.S. Aronow, L. DeLorenzo, D. Chandy, C. Ahn, S. Sule, G.W. Stallings, S. Khera, C. Palaniswamy, W.H. Frishman

**Table VI.** Effectiveness of standard data report

| Variable | Disagree strongly | Disagree somewhat | Uncertain or neutral | Agree somewhat | Agree strongly | Total | Average rating |
|---|---|---|---|---|---|---|---|
| | (–2) | (–1) | 0 | (+1) | (+2) | (*N*) | |
| The milestones report structure: milestone, core competency, curricular milestones, time frame, clinical rotation, effectively facilitates deliberations | 0% (0) | 0% (0) | 0% (0) | 67% (4) | 33% (2) | 6 | 1.33 |
| A standard report permits more rapid data analysis and presentation during clinical competency deliberations than is possible with manual chart review | 0% (0) | 0% (0) | 0% (0) | 33% (2) | 67% (4) | 6 | 1.67 |
| The milestones report permits more effective data analysis and presentation during clinical competency deliberations than is possible with manual chart review | 0% (0) | 0% (0) | 33% (2) | 33% (2) | 33% (2) | 6 | 1.00 |
| Most data central to consideration of progress along milestones is contained within the milestones data report | 0% (0) | 0% (0) | 0% (0) | 67% (4) | 33% (2) | 6 | 1.33 |
| The aggregation of curricular milestones based data and consideration within the framework of reporting milestones represents an effective evaluation strategy | 0% (0) | 0% (0) | 0% (0) | 50% (3) | 50% (3) | 6 | 1.50 |

and to determine whether application of individual subcompetencies can permit discrimination of more granular features of competence attainment.

This study had several limitations. First, it was completed at a single program which was atypical in some respects. At the time of the study, our program had already gained experience with milestone-based evaluations and had developed a robust mechanism by which to aggregate data for use in CCC deliberations. As such, the creation of a streamlined CCC process may have required less groundwork than would be necessary to generate a comparable process elsewhere. On the other hand, our findings appear to have general application in that current residency software systems such as New Innovations now offer data aggregating capability similar to that achieved by the custom data report used herein.

A second limitation was that this study relied on a small sample size and lacked a control group. Third, lack of benchmark milestone data limited our ability to make comparisons of our findings with those of others. Finally, the survey results derived from a non-validated instrument and were not fully substantiated by other objective measures. Future studies at our facility will focus on gathering data to further assess survey perceptions and to permit correlation of milestone ratings with objective patient care outcome measures.

In conclusion, the use of a modified clinical competency process which included pre-meeting assessment of milestone-based developmental progress and employed a custom data report significantly reduced committee deliberation time and permitted reliable milestone-based assessment.

## Acknowledgments

## Conflict of interest

The authors declare no conflict of interest.

References

1. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system: rationale and benefits. N Engl J Med 2012; 366: 1051-6.
2. Meade LB, Caverzagie KJ, Swing SR, et al. Playing with curricular milestones in the educational sandbox: Q-sort results from an internal medicine educational collaborative. Acad Med 2013; 88: 1142-8.
3. Nabors C, Peterson SJ, Forman L, et al. Operationalizing the internal medicine milestones – an early status report. J Grad Med Educ 2013; 5: 130-7.
4. Green ML, Aagaard EM, Caverzagie KJ, et al. Charting the road to competence: developmental milestones for internal medicine residency training. J Grad Med Educ 2009; 1: 5-20.
5. Caverzagie KJ, Iobst WF, Aagaard EM, et al. The internal medicine reporting milestones and the next accreditation system. Ann Intern Med 2013; 158: 557-9.
6. Mladenovic J, Bush R, Frohna J. Internal medicine's educational innovations project: improving health care and learning. Am J Med 2009; 122: 398-404.
7. Holmboe ES, Ward DS, Reznick RK, et al. Faculty development in assessment: the missing link in competency-based medical education. Acad Med 2011; 86: 460-7.
8. Aagaard E, Kane GC, Conforti L, et al. Early feedback on the use of the internal medicine reporting milestones in assessment of resident performance. J Grad Med Educ 2013; 5: 433-8.