

Competencies, Milestones and Observable Activities: Practical Implications for an Undergraduate Medical Education Program

DeWaay DJ, Duckett AA, Friesinger MK, Ledford CH and Walsh KJ*

Department of Clinical Internal Medicine, Ohio State University, Ohio, USA

*Corresponding author: Katherine Walsh, M.D., Assistant Professor, Department of Clinical Internal Medicine, Ohio State University, A350A Starling Loving Hall, 320 West 10th Ave, Columbus, Ohio, 43210, USA, Tel: 3522135219, Fax: 3522135219; E-mail: katherine.walsh@osumc.edu

Received date: May 18, 2017; Accepted date: May 31, 2017; Published date: June 09, 2017

Copyright: © 2017 DeWaay DJ, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Problem: Although many medical schools describe themselves as being competency based, the use of milestones is still a new concept in undergraduate medical education (UME).

Intervention: In the 2012-2013 academic year the Medical University of South Carolina (MUSC) and The Ohio State University College of Medicine (OSUCOM) independently implemented milestone-based innovations in the internal medicine core clerkships intended to improve teaching and measure student progress across the clerkship. The programs were interested in the feasibility and educational impact of the use of milestones in the clerkship curriculum.

Context: This curricular change was implemented in two third year Internal Medicine clerkship rotations at two large academic tertiary care medical schools.

Outcome: Milestones were successfully integrated into both clerkships as measured by improved or steadily excellent student evaluation of the courses. Student performance on exams or clinical assessments was unchanged.

Lessons learned: A milestone-based system can be successfully integrated into 3rd year Internal Medicine Clerkships. This type of system may be a way to improve clarity of objectives. However, the transition to a milestone-based curriculum requires increased faculty time and a need for direct observation of students.

Keywords: Medical education; Undergraduate Medical Education; Milestones; Direct observation

Introduction

Over the past 15 years there has been a movement toward competency-based medical education (CBME) with its increased focus on outcomes and clear milestones by which to monitor progress. In competency-based education, the concept of milestones has been defined as a series of markers "on the way to the explicit outcome goals of training" [1]. As part of the Next Accreditation System (NAS), implemented in July of 2013, the Accreditation Council of Graduate Medical Education (ACGME) created educational milestones that are developmental and based on observation of specialty-specific behaviors, with intention to use these milestones to measure the progress of residents through training [2]. These ACGME milestones, mapped to core competencies, are now the basis of a reporting structure for resident assessment in the accreditation process. More recently the Association of American Medical Colleges released a list of Core Entrustable Professional Activities for Entering Residency (CEPAER). This consensus statement defined activities that a resident should be able to accomplish unsupervised, the first day of residency [3].

One concern related to the competency based educational paradigm is the need it creates for increased direct observation and assessment; however, since the 1960's the time devoted to bedside teaching and

direct observation has drastically declined [4,5]. We need to reverse this trend to meet the needs of the competency based paradigm. From a graduate medical education (GME) standpoint, the use of milestones has led to innovations and efforts to develop observational tools and to improve quality of feedback. Although many medical schools describe themselves as being competency based, the use of milestones is still a new concept in undergraduate medical education (UME). Santen et al. reported use of milestones at the sub-internship level [6].

In the 2012-2013 academic year the Medical University of South Carolina (MUSC) and The Ohio State University College of Medicine (OSUCOM) independently implemented milestone-based innovations in the internal medicine core clerkships intended to improve teaching and measure student progress across the clerkship. The programs were interested in the feasibility and educational impact of the use of milestones in the clerkship curriculum. Additional goals and contextual considerations specific to each institution are discussed in depth below.

Of note, our curricular changes preceded the ACGME's efforts to define milestones. Subsequently, GME milestones have been defined as competency milestones, i.e. markers on the way to achieving a goal in a single competency. The milestones that we implemented at the UME level at our institutions are milestones toward more synthetic outcomes and observable professional activities.

Process and Outcomes–The Medical University of South Carolina

Process-MUSC

At MUSC, we defined milestones to aid students in their goal of being able to admit a patient to the inpatient general internal medicine service as part of a six week clinical assignment in a traditional inpatient internal medicine clerkship. Our goal was to create a list of observable behaviors that would be a guide to both students and faculty (Table 1). The milestones were designed to be behavioral learning outcomes. Our intention was to provide students with tangible and concrete operations that when performed would help them to meet the objectives of the course. At our institution, faculty physicians were already doing observations and using a competency based vocabulary in graduate medical education, and so, we sought to align the assessment methods at both the undergraduate and graduate levels in order to streamline faculty development efforts.

Faculty physicians observed student progress through milestones informed via direct observation of history taking and examination skills with patients and of the students' presentations of patient cases on teaching rounds. In preparation for the use of milestones in the UME program, attending physicians were instructed on how to assess the milestones using paper cards that contained a checklist and an area for written feedback. Attending physicians observed and assessed student progress through a series of milestones, specifically milestones related to history-taking by the end of week two, physical exam by the end of week 3-4, and, then, integration into an entire admission by the end of the rotation, week 6. The marks and feedback on assessment cards served as formative feedback to the student. End of rotation observations of clinical performance were collected via a centralized evaluation form that is used for all clinical courses across the College of Medicine.

Weeks	OSU		MUSC			
	Milestones		Assessment	Milestones		Assessment
1-2	Reporter/Data Collector	Obtain a medical history Obtain a physical exam Report in notes (reporter level)	All students observed, at least once. Notes reviewed from all	History	Acquire relevant history	Direct Observation
					Seek and obtain data from secondary sources	Team Rounds
					Document and report clinical information truthfully	Student Note
3-4		Obtain a medical history Obtain a physical exam Report in notes (competent to proficient reporter)	All students observed, at least twice-both history and exam	Physical exam	Perform a physical examination and identify pertinent abnormalities using common maneuvers	Direct Observation
					Track important changes in the physical examination over time	Team Rounds; Review of Note
5	Interpreter	Interpret clinical data within patient contexts (advancing into interpreter level)	Data-interpretation and clinical reasoning in small groups, additional bedside observations and teaching	Admit a patient to the General Medicine Service	Synthesize all available data to define each patient's central clinical problem	Direct Observation
					Formulate differential diagnoses, and an evidence based plan	Team Rounds; Review of Note
					Understand and interpret common diagnostic testing	Direct Observation; Team Rounds
					Demonstrate appropriate use of information technology	Team Rounds; Review of Note
					Present the patient on rounds in standardized format.	Direct Observation; Team Rounds
					Write a progress note in standardized format.	Review of Note
Learn to recognize an unstable patient	Simulation Activity					

					Recognize when to seek additional guidance	Team Rounds; Simulation Activity
					Customize care in the context of the patient's preferences	Team Rounds; Faith Assessment Activity
					Demonstrate an ability to build a healing relationship	Direct Observation; Night Shift Activity
					Communicate effectively with non-physician members of the health care team	Team Rounds; Inter-professional Activity
					Communicate effectively with non-physician members of the health care team	Team Rounds; Inter-professional Activity
6	Interpreter	Interpret clinical data within patient contexts (nearly competent to proficient interpreter)	Additional observations and small group teaching (flexible, based on student needs)			
7-8						

Table 1: Milestones for OSU and MUSC internal medicine clerkships.

Outcomes-MUSC

Our pre-intervention group consisted of 155 students from academic year 2011-2012 (Class of 2013) and our post-intervention group consisted of 166 students from academic year 2012-2013 (Class of 2014). Data from the AAMC Graduation Questionnaire (GQ), internal clerkship evaluations, and NBME Shelf scores were used to compare the two groups. Internal clerkship evaluation data was collected in E*Value, a healthcare education enterprise software. All analyses were conducted using SAS version 9.3 (SAS Institute, Cary, NC). We used independent t-tests to tabulate and compare mean differences between the two student groups. In addition, we used student t-tests to determine whether the pre-intervention and post-intervention group's mean score differences varied significantly in the dependent variables of the evaluation categories and their NBME shelf

score. AAMC GQ data from the Class of 2013 and Class of 2014 were compared using Fisher Exact Test, two-tailed analysis to determine significance. This study was deemed "not human subjects research" by the Health Sciences South Carolina (HSSC) electronic Institutional Review Board (eIRB).

We compared the satisfaction data given by students at the end of the clerkship for the pre-intervention students to those of the post-intervention students and found that the post-intervention students were significantly more likely to agree that the clerkship was well organized (73 vs. 90; $p < 0.0001$), the learning objectives were clear (86 vs. 95; $p < 0.0001$) and assessed (73 vs. 85; $p = 0.02$), there was direct observation of their patient care (90 vs. 95; $p = 0.01$) and there was an opportunity to develop clinical skills (67 vs. 84; $p = 0.02$) (Table 2).

Clerkship Evaluation Criteria	MUSC		OSUCOM	
	Pre-Intervention	Post-Intervention	Pre-Intervention	Post-Intervention
	2011-2012	2012-2013	2011-2012	2012-2013
N	155	166	212	234
Clear learning objectives	86	95 ^b	98	97
Provided the opportunity to accomplish the learning objectives	83	94 ^b		
Performance assessed against the learning objectives	73	85 ^b	99	90
Graded activities assessed mastery of clinical skills	67	84 ^b		
Provided opportunity to develop clinical skills	90	96 ^b		
Clerkship activities were well organized	73	90 ^b		
Supervisor watched me perform a clinically pertinent history or physical exam	90	95 ^b		

A faculty member personally observed me taking a patient history			94	97
A faculty member personally observed me performing physical exam			96	96
Compared to other clerkships, I would rate the educational experience of this clerkship as: % Better or Much Better	46	58		
Rate the quality of your educational experience in this clinical clerkship: % Very Good or Excellent			86	93 ^a

^aStatistically significant difference for OSU data using two-tailed Fisher Exact Test, $p < 0.05$. Student t Test using mean, SD, N-No statistically significant differences for OSUCOM responses

^bStatistically significant differences for MUSC data using student t Test. I received clear learning objectives ($t = -4.16$, $p \leq 0.0001$), Provided the opportunity to accomplish the learning objectives ($t = -3.60$, $p = 0.0004$), My performance was assessed against these learning objectives ($t = -2.39$, $p = 0.02$), Graded activities accurately assessed mastery of clinical skills ($t = -2.58$, $p = 0.01$), Provided opportunity to develop clinical skills ($t = -2.34$, $p = 0.02$), Clerkship activities were well organized ($t = -4.95$, $p \leq 0.0001$). A supervisor (attending/resident) watched me perform a clinically pertinent history or physical exam ($t = -2.67$, $p = 0.01$).

Table 2: Student evaluation of clerkship (end of clerkship) % agree and strongly agree.

AAMC GQ data showed improved ratings of overall clerkship quality. For the post-intervention group, 178 (94%) students reported that they were observed performing both a history and a physical exam. The percentage of students rating the clerkship as good or excellent increased from 75% to 88% ($p = 0.006$) [Table 3]. There were no differences in clinical performance scores (ratings by supervising physicians on a scale from 1-4). Students in the pre-intervention group had a CPE average of 3.62, while the post-intervention students had an average of 3.63. In addition, there were no differences seen in the NBME subject exam scores between AY 11-12 and AY 12-13. Students in the pre-intervention group had a Shelf average of 76.06+6.99, while the post-intervention students had an average of 75.21+9.60.

The attempt to streamline faculty development efforts was not as successful for several reasons. First, the milestones we established for students are milestones toward performing an observable professional activity whereas the milestones for GME are competency milestones, thus, different frameworks. Second, the end of clerkship clinical performance assessment form used by the College of Medicine did not change and also was not aligned with the resident form. Although, faculty development was not streamlined due to differences in forms and frameworks, we were able to emphasize a similar use of direct observation techniques to assess both students and residents.

Process and Outcomes-The Ohio State University College of Medicine

Process-OSUCOM

At OSUCOM, we instituted a “Medicine Mentor” program as a specific educational innovation designed to aid students in their development along milestones across what had become short clinical assignments. The clerkship was an 8 week inpatient course in which a student is assigned four blocks that are two weeks each. While the shorter assignments allowed every student to care for patients as part of general medicine and cardiology services, as well as two other specialty settings, we were concerned about the loss of more longitudinal assessment and feedback supporting core clinical skills and the major goals of the clerkship. The purpose of this intervention was threefold: 1) to maintain and increase direct observations of core clinical skills, an ongoing area of challenge and effort, 2) to maintain and increase student satisfaction with the clerkship, and 3) to improve learning. As a secondary goal, we were preparing for an upcoming plan

to form a longer, more complexly integrated clerkship structure. The Medicine Mentor program offered an opportunity, not only for increased direct observation of progress based on milestones, but also provided students with a stable faculty-student relationship that was exclusively formative (ungraded) and supportive of learning.

To define milestones in the medicine clerkship, Ohio State used the Reporter-Interpreter-Manager-Educator framework described by Pangaro (Table 1) [7]. The structure of the Medicine Mentor program consisted of a faculty mentor assigned to 6-7 students, with a total of 5-6 mentors per rotation. The Medicine Mentors spent 20 hours cumulatively working with students over the 8 week clerkship, providing continuity of teaching and coaching across four 2 week clinical assignments in a pure teaching role, without patient care responsibility as in the attending physician model. The milestones served to guide the mentor to developmentally assess the student’s core clinical skills irrespective of the clinical focus of the service on which the student was rotating. Direct observations were conducted for history taking and physical exam skill. Verbal and written feedback was given at the time of the direct observations; additional observations as well as review of clinical documentation (notes) were encouraged, but optional. All assessments and feedback were formative with the intent to create a learning environment where students could comfortably ask questions and seek feedback. In addition to direct observations, the faculty also facilitated small group discussions with their students on topics selected with student input. Prior to the Medicine Mentor program, all feedback and direct observations on clinical skills were the responsibility of ward attending(s) to complete during the busy clinical assignments. The opportunity to practice skills in a non-graded atmosphere did not exist. For our pre-intervention group all clinical performance assessments were completed by supervising physicians on clinical wards, and used to measure student performance.

Outcomes-OSUCOM

Student evaluation data from clerkship evaluations and the AAMC Graduation Questionnaire were reviewed to determine if there was a measurable impact on the students’ perception of the quality of the clerkship, the amount of feedback received, and whether direct observations were performed. Our pre-intervention group consisted of 212 students from academic year 2011-2012 during which the Medicine Mentor program did not exist. Our post-intervention group consisted of 234 students from academic year 2012-2013 the year of intervention (the Medicine Mentor program). GraphPad was used to

analyze categorical data, using two-tailed Fisher Exact Test. Continuous data was analyzed using SPSS (v.22), using student t-test. Approval was granted through Ohio State's Human Behavioral and Social Sciences IRB.

At end-of-clerkship evaluation, the percentage of students rating the clerkship Very Good or Excellent increased from 86% to 93% (p=0.045). A high number of students reported being observed performing a history (227 or 97%) and examination (224 or 96%), and was not statistically different than the prior year (Table 2). For the AAMC GQ 2014, 180 (100%) students reported that they were

observed performing a history and 180 (99%) reported being observed performing an exam, i.e. 1 of 181 responded "no".

The percentage of students rating the clerkship as good or excellent increased from 93% to 98% (p=0.045) (Table 3). There were no differences in clinical performance assessment scores (CPAs), ratings by supervising physicians. Students in the pre-intervention group had a CPA average of 76.4, while the post-intervention students had an average of 76.1. In addition, there was no difference seen in the NBME subject exam scores between AY 11-12 (78.2 ± 7.8) and AY 12-13 (80.4 ± 7.8).

Standardized Graduation Questionnaire Criteria	MUSC GQ 2013	OSU GQ 2013	All Schools GQ 2013	MUSC GQ 2014	OSU GQ 2014	All Schools GQ 2014
N	124	189	14071	126	181	14264
Received clear learning objectives	94	98	93			
Performance assessed against the learning objectives	85	88	85			
A faculty member personally observed me taking a patient history	84	94	78			
Were you observed taking the relevant portions of the patient history (% Yes)				94	100	92
A faculty member personally observed me performing physical exam	85	94	82			
Were you observed performing the relevant portions of the physical or mental status exam (% Yes)				95	99	93
Rate the Quality of your Educational Experiences on this Clinical Clerkship: % Good + Excellent	75	93	91	88 ^a	98 ^a	92

^aNotes statistically significant difference compared to previous year, MUSC p=0.006 and OSU p=0.044 (Fisher exact test, two-tailed); Other items were unable to analyze for statistical differences due to change in item and response scale

Table 3: Clerkship evaluation items (AAMC graduate questionnaire) % agree or strongly agree.

Lessons Learned

Educational milestones can guide learning by framing the development of curricula and illustrating to faculty the specific behaviors and skills medical students should be expected to demonstrate at a particular time. At both OSUCOM and MUSC, we found that it is possible to successfully implement a milestone-based curriculum in an inpatient internal medicine clerkship. Results from each program independently showed that the interventions increased student perception of clerkship quality. MUSC's results suggest that a simple intervention related to clear, detailed milestones, coupled with specific assessment cards, can also increase observation and student clarity related to objectives and assessments. While student perception of quality and important learning process measures improved, there were no measurable changes in learning outcomes. Both the short duration of the clerkships (6-8 weeks) and the imprecision of our clinical assessments may limit our ability to detect small incremental improvements in the learning of core clinical skills, however.

Direct comparison of the differences in outcomes of these two interventions is interesting. Presumably, some differences in data between OSUCOM and MUSC are due to the baseline clerkship data; OSU started with high rates of direct observation and a high clarity of objectives; these high baseline levels likely resulted in a ceiling effect. An increase in observation of history-taking (69.2% to 93.5%) and examination (72.5% to 91.6%) was accomplished 2 years prior related to a new requirement that every attending physician observe students; this did not result in an improvement in clarity of objectives, feedback or clerkship quality, according to AAMC GQ data. In fact, while no statistical differences were seen on quality ratings, students and faculty both voiced dissatisfaction with the required observations. Also although the general approach of defining milestones and observing clinical skills was shared, the two schools used very different frameworks and level of detail in defining the milestones (Table 1). Despite the differences, both resulted in positive program outcomes with high rates of direct observation. We postulate that perhaps the framework is not as important as the intentional clarity and purposeful definition of an expected progression of learning that is achieved by a milestone-based approach. Further study is needed to determine whether the interventions and results will be sustainable. Direct

comparison of the relative value of the different components in the two interventions is not possible, as these were two studies at separate institutions. Further information is needed to fully understand the contextual factors, potential barriers, and adaptability of the approaches.

Furthermore, in defining EPAs for entering residency to serve as a critical milestone between undergraduate and graduate medical education, the AAMC has called for efforts to develop innovative teaching strategies and assessment methods related to the concept of entrustable activities [8]. Our work suggests that perhaps the core EPAs for entering residency might also result in more observation, more clarity and higher student perception of quality of the program. However, some of these benefits may be more effectively achieved by use of more timely and incremental milestones along the path of achieving the EPAs. The need for a more incremental milestone in residency education is described by Eric Warm et al. In his description of observable professional activities (OPAs). Our works with OSUCOM's Medicine Mentors program and MUSC's use of an assessment card system are two institution-specific examples of how milestones might be useful in even short educational experiences, such as an internal medicine clerkship. Development of milestone-based curricula in other specialties is ongoing [9,10].

As milestone-based initiatives continue to emerge, other institutions may be faced with similar challenges to ours. In both of our institutions, the transition to a milestone-based curriculum required an increase in faculty time - to attend faculty development sessions, to perform the direct observations, to give feedback to learners, and to document formative assessments. Further study is needed to measure the cost and value of such interventions, as well as the utility, reliability and learning impact of the more frequent assessments. As medical student education moves to a milestone or EPA based system, further attention will be needed to explore issues of variability in teaching and assessments and to identify effective educational practices related to curricular resources, teaching clinical skills, assessing correct behaviors, and providing useful feedback [11,12]. Programs like OSUCOM's Medicine Mentor program which focus on fewer faculty members and use of more longitudinal assessments may help address some of these concerns.

There are still many questions that remain, particularly related to the different languages currently in use in medical education. While our study indicates that we can increase direct observations, further study is needed to understand how many observations, by how many observers and across how many contexts, are needed for "entrustment" on the first day of residency training. Santen et al. have begun to explore what entrustment means for 4th year Sub-interns. Their findings are in keeping with our study, where third year students were able to obtain and report data in a supervised setting and progressed to obtain and report data without direct supervision. Medical schools will need to develop plans to address how the many smaller and more formative assessments might be used to generate a summative assessment at the end of medical school.

Both institutions focused on observable activities; however, entrustment was never directly assessed. We believe our assessments will be a helpful piece in the longitudinal assessment and cross-contextual observations necessary to grant entrustment. Also, if a student is able to demonstrate a higher level of skill earlier in an

assured manner, perhaps a more formal effort can be made to accelerate her or his level of responsibility accordingly.

The implementation of milestones at our programs highlights the schism that continues between UME and GME. Although both entities are evolving to be increasingly competency based, they are doing so using two different methods with a similar vocabulary. UME is using EPAs and, we predict, the use of incremental observable professional activities as meaningful milestones in our curricular reform. In contrast, GME is using competencies with developmental competency milestones for assessment. Dr. Warm's use of cross-mapping observable professional activities is one way to perhaps bridge these frameworks. We would suggest that while the complexity of these differing approaches may have curricular planning benefits, the two entities, or perhaps more realistically, those of us charged with the pragmatic role of implementing these ideals, could simplify the efforts and reduce the burden at the supervising physician level.

Overall, our goal was to demonstrate that use of clear milestones in clerkships will improve not only the process of learning, but also the trajectory of learning, the latter being one of the purported goals of competency based education. While we were not able to demonstrate a benefit to learning trajectory, hopefully, this pair of observational studies demonstrates one small step toward more meaningful teaching and assessment through use of milestones.

References

1. Frank JR, Mungroo R, Ahmad Y, Wang M, De Rossi S, et al. (2010) Toward a definition of competency-based education in medicine: A systematic review of published definitions. *Med Teach* 32: 631-637.
2. Nasca TJ, Philibert I, Brigham T, Flynn TC (2012) The next GME accreditation system-rationale and benefits. *N Engl J Med* 366: 1051-1056.
3. AAMC (2015) Core entrustable professional activities for entering residency.
4. Lacombe MA (1997) On bedside teaching. *Ann Int Med* 126: 217-220.
5. Stickrath C, Noble M, Prochazka A, Anderson M, Griffiths M, et al. (2013) Attending rounds in the current era: What is and is not happening. *JAMA Intern Med* 173: 1084-1089.
6. Santen SA, Seidelman JL, Miller CS, Brownfield ED, Houchens N, et al. (2015) Milestones for internal medicine sub-interns. *Am J Med* 128: 790-798.
7. Pangaro L (1999) A new vocabulary and other innovations for improving descriptive in-training evaluations. *Acad Med* 74: 1203-1207.
8. Chen HC, van den Broek WE, Ten Cate O (2015) The case for use of entrustable professional activities in undergraduate medical education. *Acad Med* 90: 431-436.
9. Morgan H, Marzano D, Lanham M, Stein T, Curran D, et al. (2014) Preparing medical students for obstetrics and gynecology milestone level one: A description of a pilot curriculum. *Med Educ Online* 19: 25746.
10. Santen SA, Peterson WJ, Khandelwal S, House JB, Manthey DE, et al. (2014) Medical student milestones in emergency medicine. *Acad Emerg Med* 21: 905-911.
11. Kogan JR, Conforti LN, Iobst WF, Holmboe ES (2014) Reconceptualizing variable rater assessments as both an educational and clinical care problem. *Acad Med* 89: 721-727.
12. Holmboe ES, Ward DS, Reznick RK, Katsfrakis PJ, Leslie KM, et al. (2011) Faculty development in assessment: The missing link in competency-based medical education. *Acad Med* 86: 460-467.