2009

Creating a path for high quality care: implementation of the mini-CEX in a student-run clinic

Paul Daniel Di Capua
Yale University

Creating a Path to High Quality Care: Implementation of the mini-CEX at a student run clinic

A Thesis Submitted to the

Yale University School of Medicine

in Partial Fulfillment of the Requirements for the

Degree of Doctor of Medicine

by

Paul Daniel Di Capua

MD, MBA

Class of 2009

Abstract

Creating a Path to High Quality Care: Implementation of the mini-CEX at a student run clinic.
Paul D. Di Capua, Michael L. Green, Kathleen P. White, Eric S. Holmboe. Department of Internal Medicine, Yale University, School of Medicine, New Haven, CT.

Purpose: This study examines the utility of implementing the mini-CEX at a student-run primary care clinic. The mini-CEX is a validated tool for clinical skills feedback and evaluation. The implementation of the mini-CEX sought to improve quality standards in providers' clinical skills, and therefore to set a model for continuous quality improvement in medical education.

Hypothesis: The primary hypothesis tested is that a minimum of three mini-CEX's per week will produce a positive "utility" for both students and faculty.

Methods: A needs-based analysis focus group solicited opinions on the mission and function of the clinic from the student-providers. The mini-CEX was then presented as a feedback tool that could create a more structured learning environment. After the providers agreed to a trial implementation of the mini-CEX, the faculty at the clinic were instructed on observation and feedback using the mini-CEX. During the trial study, the authors collected a copy of all mini-CEX forms which were used to document number of evaluations per week, characteristics of the visit, and the feedback given. Interviews conducted during the trial study with both students and faculty were documented on a weekly basis. The utility of the mini-CEX was determined according to a previously published model which examines the reliability, validity, educational impact, acceptability, cost and feasibility.

Results: The student-providers agreed to the trial implementation of the mini-CEX in their clinic. During the trial period, there was a mean of 3.82 (median = 3) evaluations. Interviews with faculty revealed a lack of experience in observing trainees with patients but also increased attention to the teaching they could provide the students. Interviews with students revealed frustration at the lack of consistency in the teaching, but appreciation of the feedback process.

Conclusions: The mini-CEX was successfully implemented in this student-run clinic. The feasibility is evidenced in the number of weeks in which the minimum of 3 evaluations were performed. The positive feedback during and after the trial period from both faculty and students evidences the acceptability of the mini-CEX. The implementation of a feedback and evaluation tool by students for students represents a movement in structuring clinical education in a self-regulatory manner to ensure higher quality clinical skills by future physicians.

Acknowledgments

This project would not be possible without the support, enthusiasm and encouragement of my esteemed colleagues and peers at the Wednesday Evening Clinic: Janet Abou, Jacob Applebaum, Mary Badon, Sanjay Basu, Jessica Berwick, Jonathan Chen, Joanna Chin, Justin Cohen, Neil Desai, Laura Dichtel, Benjamin Marks, David Merrick, Yasha Modi, Kuda Mutyambizi, Deepak Rao, Swapna Reddy, Jonathan Romanyshyn, Rachel Rutishauser, Raj Sawh, Robert Stavert, and Elizabeth Wahl, MD.

I would also like to thank Dr. Kathleen White, the Director the Clinic, whose unflagging energy over the past year helped fuel this project and without whose commitment the Clinic would not be possible. Dr. Dillard, thank you for your wisdom and dedication which have been essential to the clinic since its inception. I owe many thanks as well to Dr. Michael Green who helped shape this project and whose commitment over the year brought expertise and realism to ensure successful implementation of improvements in the Clinic. Your dedication to teaching stands out, perhaps only second to your capacity to lead a challenging discussion. Dr. Holmboe, thank you for your encouragement, support and expertise as I refined the goals of the project and the analysis of the work; your vision for bringing quality improvement measures to medical education are a blessing for those of us entering our training, and I aspire to rely on superb clinical skills in my care for patients.

Lisa Sander helped in infinite ways in launching this project, and I wish her well in her own medical career. Matthew Kim provided critical support during difficult months to ensure the implementation went smoothly.

The Yale University School of Medicine, Department of Internal Medicine, provided important support of the completion of my thesis. This project would not have been possible without the financial support of Yale University School of Medicine Medical Student Research Fellowship.

Over the past year, I have depended on my friends in both the Schools of Medicine and Management to make this possible. Finally, I would like to thank my parents, brother and sister for their love and patience throughout the year and perhaps particularly at the end, in putting this project together.

Table of Contents

Introduction

The 1998 report Medical School Objectives Project (MSOP) released by the Association of American Medical Colleges (AAMC) states that "the goal of medical education is to produce physicians who are prepared to serve the fundamental purposes of medicine. To this end, physicians must possess the attributes that are necessary to meet their individual and collective responsibilities to society."[1] National consensus unambiguously agrees that the American health care system is in drastic need of reform. As the American public and professional organizations increasingly call for greater emphasis on patient safety and higher quality of care, it is incumbent on future physicians to rise to these higher standards in order to meet our responsibilities to the society we are training to serve. Superior clinical skills are an integral component of these higher standards for quality of care. This thesis discusses the role of medical education in preparing future physicians to rise to these levels of service by emphasizing the acquisition of effective clinical skills during medical training.

As a graduating medical student, I have been blessed with a superb education at the Yale School of Medicine and am thankful to the School and my dedicated mentors for guiding me to this point. Throughout my medical education, the faculty and residents at Yale have shown extraordinary attention to teaching, often taking time away from other important priorities to ensure we, the medical students, were engaged in our learning, had all of our questions answered and had the knowledge and confidence to take good care of our patients. I believe I will graduate from medical school ready to begin my training as a physician.

However, as a management student of Dr. W.E. Deming, the founder of Total Quality Management, I ascribe to his Fourteen Points for Transforming Management Effectiveness (see

Appendix A), one of which is, “Improve constantly and forever the system of production and service.”

With the humility of knowing that the art of medicine is infinitely deep and the science of medicine equally complex, this thesis project is an example of a student taking responsibility for improving the system in which we learn our craft and through which we provide care to our patients. This project has put me in the self-critical position of seeking to improve the same system in which I have trained, but I do so only because I have come to value that system and its lessons so profoundly so as to want to leave them even better. The voice from inside the short coat, ever optimistic and eager, may prove refreshing for the many thankless hours our educators have dedicated to our growth.

In particular, one part of the system in which I renewed and honed my clinical skills was in the student-run Wednesday Evening Clinic (WEC). Seeking to improve the clinical and educational objectives of the clinic, this thesis presents a model in which the interests of the WEC stakeholders are aligned to further the mission of the clinic by the implementation of a validated evaluation tool, the mini-CEX. The success of the implementation will be evaluated by its utility, according to a modified version of the Utility Index [2, 3]. The utility index, initially developed by van der Vleuten, examines the reliability, validity, educational impact, acceptability, cost and feasibility of assessment methods in medical education. The primary hypothesis tested is a minimum of three mini-CEX’s per week will produce a positive “utility” for both students and faculty. The goals of the project, however, are loftier, seeking to improve our own medical education, to raise the standards we set for ourselves as future physicians, and to build a culture of quality control into the system through which we provide care to our patients.

Background

Quality improvement is the fundamental driver of this endeavor to test the utility of the implementation of the mini-CEX at the WEC. Quality in healthcare has taken on its own definition, focusing on technical results such as hemoglobin A1C levels, door-to-cath times, and so on. However, "quality" in industry has a broader definition, and its history offers many lessons for improving patient care, and more specifically, medical education. In order to frame why we need more structured teaching and assessment of clinical skills in our medical education, this background section provides a brief introduction to quality in industry and then focuses on quality in medical education. This is followed by a more detailed introduction of the mini-CEX. Finally, to provide a better context for the project, I also briefly discuss the particularities of medical education at the Yale School of Medicine and of the WEC in particular.

Quality: A brief history with instrumental lessons

The term quality was defined for industrial purposes by Joseph Juran as "fitness for use" and by Philip Crosby as "conformance to requirements." In industry, these definitions lay the foundation for the theory of quality, which seeks to minimize errors and to create systems of production or service which conform to set standards. The notion of high quality, for which firms such as Toyota or the Ritz Carlton have become famous for, set narrow standards with minimal variation to consistently achieve remarkable results.

Every process exhibits variation. The quality gurus of the post-WWII era such as W.E. Deming and Joseph Juran sought to minimize variation to achieve higher quality. Variation was classified as either: (1) *systemic variation*, or the inherent variability of the process, which could never be eliminated but kept in conformance to requirement and (2) *assignable cause*

variation, the result of a specific cause that changes the process. If there are assignable causes to variation, such as an untrained employee or defective raw materials, then these assignable causes can be fixed to return production or service back to standards. However, if variation is due to systemic causes, then the process of production or service must be changed in order to comply with standards.

In the process of helping Japanese auto makers establish their superiority in producing high quality cars, Dr. W.E. Deming summarized his expectations of his clients in the aforementioned 14 point program (see Appendix A). He insisted that managers should focus their efforts on tomorrow, rather than being preoccupied with today by orienting them towards continuous improvement of products and services to meet the needs of their customers.

Quality had to be built in to the system. One of the fourteen points of his program states "Cease dependence on mass inspection." Inspection implies planning on defects, at which point it is too late and ineffective. Instead, Deming insisted, processes must be improved so that the product or service can be guaranteed to be of high quality and render inspection superfluous.

It would take decades for his ideas to catch on in the United States though he had found eager learners in Japan. By 1986 Deming predicted that it would take 30 years for Americans to match the progress of the Japanese and that the United States was still falling behind. Today, in 2008, Deming might be smirking if he could see the United States auto manufacturers facing the possibility of bankruptcy and the American public largely unsympathetic to what is largely perceived as badly managed companies producing sub-standard cars. The implications for the American health care system, and for medical education in particular, suggest better quality control in our system of education to identify and correct both systemic and assignable causes

in order to reliably produce high quality physicians, lest we fall 30 years behind and face an even more uncertain future.

Quality in medical education

As the national discussion on healthcare increasingly calls for greater emphasis on higher quality of care, attention has turned to medical education in determining whether future physicians are adequately prepared to provide high quality healthcare. High quality health care, taken in the broader context described above, implies setting rigorous standard requirements for graduating medical students such that future physicians are reliably fit to provide the high quality care demanded from the American public.

The 1998 MSOP report reached a consensus on four essential attributes that physicians need to fulfill their responsibilities to society, of which one was "Physicians must be skillful." In delineating the characteristics of a "skillful" physician, the MSOP includes history and physical exam skills, diagnostic skills, synthesis of a differential diagnosis, management of information from imaging or laboratory studies, and skills in understanding patient needs and skills in communicating with patients and their families about their concerns regarding the patients' health [1]. The synthesis of an ambitious medical curriculum, however, is not sufficient: Moercke and Eika have shown that 75% of newly graduated physicians in Denmark had learned the expected curriculum, demonstrating a sizeable gap between expected and learned skills [4]. Similarly, graduating US medical students [5] and residents [6] have been shown to miss important patient information in interactions due to inadequate or incomplete interviews or physical exams.

The Institute of Medicine's report on medical errors, *To Err is Human*, raised concerns regarding physician training and education, serving as a high-profile emphasis on the need for increased supervision for physicians in training. Graduating medical students entering residency have shown an alarmingly wide variability in clinical skills[7], and residents are commonly assumed to possess skills they do not have [1, 4]. In fact, the Association of American Medical Colleges (AAMC) visited 97 medical schools and found that faculty rarely observed medical student patient encounters; instead evaluators' impressions of the students' presentation skills and knowledge formed the basis of their evaluations [8]. And despite the widespread call for higher quality health care, lower costs and greater patient safety, the medical literature evidences a general paucity of studies examining the effectiveness of clinical teaching or how to measure its long-term outcomes [9-11].

Moreover, residents often point out that as medical students they were told that they would learn certain skills or processes as residents, but as residents they are expected to have learned them in medical school [7]. Without a well-organized infrastructure to ensure that all trainees are well-equipped with the necessary clinical skills to distinguish them as high quality physicians, ad hoc clinical teaching instead results in haphazard results revealed by wide variability, and assessments based on impressions and indirect evaluations. Deming and Duran would cite the widely variable results of the medical education process as evidence of a poor-quality system, with "assignable causes" that could be addressed to improve results.

In the clinical years of medical school as with residency, it seems that a large part of learning the critical skills which distinguish an average physician from a high quality physician are taught ad hoc, with scant processes in place to assess and improve trainees' clinical skills. However, the 1998 MSOP report emphasized institutional accountability to ensure that graduates

possessed the necessary clinical skills upon graduation; an important, if symbolic shift towards systems-based approaches to quality control in medical education [1]. This change is indicative of the new direction to improve quality standards and measures in a system otherwise seemingly plagued with wide variability and few quality control measures of the trainees' clinical skills. The new direction towards greater emphasis on clinical skills evidenced in changes in Internal Medicine clerkship curricula around the country [12].

The mini-CEX

The American Board of Internal Medicine developed the mini-CEX as a means to efficiently evaluate competence in clinical skills of residents over a number of cases of variable complexity in various settings (ambulatory, emergency department, and inpatient)[13]. The execution of a mini-CEX evaluation requires one faculty member to evaluate a trainee in a 15- to 20-minute patient encounter. Over the span of the training period, the trainee has multiple opportunities for observation and feedback, and should therefore accumulate several evaluations from different faculty with a variety of patients with a broader spectrum of clinical problems and required skills. The scores produced by the mini-CEX have been noted to be more reliable than those based on the traditional CEX as an evaluation tool [13]. In a large observational study using the mini-CEX in 21 different residency programs, Norcini et al quantified the reliability of the mini-CEX by calculating the 95% confidence interval for scores on the mini-CEX as the number of encounters increased, summarized in Table 1 [13]. For example, for a 95% confidence interval of less than one point on the nine point scale used in the mini-CEX questions, trainees need a minimum of four evaluations for "pass-fail" decisions, but only from the perspective of reliability. Furthermore, the mini-CEX has been shown to discriminate

between pre-existing levels of seniority between residents, and has had high satisfaction rates in both evaluators and evaluated[14].

One study qualitatively examining the value of the mini-CEX in first-year residents has quotations demonstrating both the demand for more supervised learning as well as the value of the mini-CEX:

> …Medicine is an apprenticeship. You can't learn a physical exam from a book, you have to learn it from another expert you can speak with, so I do think this is a useful exercise… it's a good opportunity to get one on one with a staff person and go through whatever case scenarios it is, go through a physical exam. As a resident we often time we don't get to do that… [15]

Previous studies have shown the mini-CEX to have construct validity [13, 16, 17]. Knowing that the mini-CEX is an effective clinical skills feedback tool, this project sought to assess the utility of implementing the mini-CEX at the WEC. The mini-CEX has been previously used in the instruction of medical students[18, 19]. However, its use on medical student clerkships limits the ability of students to take advantage of serial feedback to continually hone their skills over months or a year. To our knowledge this is the first implementation of the method in the setting of a student-run longitudinal care clinic.

Table 1: 95% Confidence interval on mini-CEX questions

Table 3. The 95% CI for the Total Score Based on 1 to 14 Encounters and for a Total Score of 5.0*

Encounters, *n*	95% CI	95% CI for a Total Score of 5.0
1	±1.47	3.53–6.47
2	±1.04	3.96–6.04
4	±0.73	4.27–5.73
6	±0.60	4.40–5.60
8	±0.52	4.48–5.52
10	±0.46	4.54–5.46
12	±0.42	4.58–5.42
14	±0.39	4.61–5.39

* This table provides an example of how the variability for a resident with a total score of 5.0 changes with the number of mini–clinical evaluation exercise (mini-CEX) encounters completed. The components that influence the variability include the resident, program, patient, and examiner. Therefore, the CIs would probably be different for another sample of programs, patients, and examiners.

Table 1: 95% Confidence interval on mini-CEX questions based on number of encounters. This table allows implementation of mini-CEX to estimate the accuracy of the scores received by trainees and furthermore demonstrates the value a tool that can be readily used repeatedly in routine patient care. (Table taken from Norcini, Ann Intern Med, 1995)[13]

Utility Index

The 'utility index' described by Cees van der Vleuten in 1996 outlined the key components involved in developing assessment methods [3]. The utility (U) of an assessment is defined as the multiplicative product of five variables, each with an associated weight (w) depending on context. The five variables van der Vleuten included in his model are as follows: Reliability (R), Validity (V), Educational impact (E), Acceptability (A), and Cost (C).

The Postgraduate Medical Education and Training Board (PMETB) of the U.K. expanded this model to include feasibility (F) in recognition of the burden each additional assessment adds to

the system. Although the variables in the model can only rarely be quantified, the model can be summarized as follows:

$$U(assessment) = w_R R \times w_V V \times w_E E \times w_A A \times w_C C \times w_F F$$

Reliability refers to the reproducibility of the results of the assessment. Validity refers to an assessment's capacity to test what it is supposed to test. Validity requires multiple sources of evidence, and is conventionally subdivided into various facets: face validity, content validity, construct validity, predictive validity, consequential validity (see Table 2 for description of each of these levels). Educational impact refers to the assessment's capacity to maximize learning in the intended area of training. The reliability, validity and educational impact of the mini-CEX have been previously established [13-17, 20, 21]. Although some studies have examined the acceptability and feasibility indirectly [14, 18, 19], they were not assessed in the context of the utility model, nor do they examine the cost of the assessment.

Table 2: Traditional Facets of Validity

Face validity	Compatibility with the curriculum's educational philosophy	What is the test's face value? Does it match up with the educational intentions?
Content validity **(may also be referred to as direct validity)**	The content of the curriculum	Does the test include a representative sample of the subject matter
Construct validity **(may also be referred to as indirect validity)**	Does the evidence in relation to assessment support a sensible underpinning construct or constructs?	What is the construct, does the evidence support it? E.g. Differentiation between novice and expert on a test of overall clinical assessment – do two assessments designed to test different things have a low correlation?
Predictive validity	The ability to predict an outcome in the future, e.g. professional success after graduation	Does the test predict future performance and level of competency?
Consequential validity	The educational consequence or impact of the test	Does the test produce the desired educational outcome?

Table 2: from Joshi, 2007. The various levels of validity reflect the different kinds of data necessary to measure validity.

The Yale System

All students and faculty at Yale know that Yale's approach to medical education is unique, that it has a history, and that we call it, generally without knowing what it means, the "Yale System." As historian Arthur Viseltear writes in his history of the Yale System of Medical Education,

> Most agree that the Yale System is indeed a system, in that it comprises a set of doctrines and principles which explain the working of a systemic whole, but there are others who argue that it is more a philosophy, a peculiar mind set, something intangible. [22]

The Yale System dates back to 1920, with the election of Dr. Milton C. Winternitz as Dean of the School. Dean Winternitz is credited for brining Yale School of Medicine to prominence, his legacy persists through his innovative approach to medical education, and he is perhaps best remembered by then Yale President James Angell's description of the Dean as a "steam-engine in pants."

More specifically, Dean Winternitz sought to keep "the eyes of medical educators" on Yale. He did so first, by designing, in Viseltear's words, an "elastic curriculum which liberated students' time and was adaptable to individual abilities and needs." [22] Secondly, to keep their attention focused on Yale, he sought to develop "pedagogic experiments which, if judiciously implemented, would have the twofold benefit of aiding medical education and of giving 'character' to Yale" [22].

Students and faculty today would certainly recognize features envisioned by the System's founder of elasticity and liberated student time adaptable to individuals. The process of evaluating educational procedures as exemplified by this thesis, almost 90 years later, embodies the spirit of "pedagogic experiments" through which the Yale System evolved over time.

In their informational website for prospective students, the Office of Education requires more than 650 words to define the System, summarized perhaps best in the following paragraph:

> The fundamental element of the system is the concept that medical students are mature individuals, strongly motivated to learn, requiring guidance and stimulation rather than compulsion or competition for relative standing in a group. The corollary of this concept is that students must assume more than usual responsibility for their education. Students should be considered adults in a graduate school and be permitted to enjoy as large a degree of freedom as is consistent with the fulfillment of requirements for the degree of Doctor of Medicine. Memorization of facts should be far less important than a well-rounded education in fundamental principles, training in methods of investigation, and the acquisition of the scientific habit of mind.[23]

Critical to the Yale System is the lack of grades and formal examinations, which reflects the presumed maturity and self-motivation to learn: "Freed from the usual anxieties provoked by examinations, students tend to learn for their future rather than for tests." [23] In this context, Yale medical students are extremely territorial over attempts to impose a testing regimen on us, but, by assuming "more than usual responsibility for [our] education," this thesis exemplifies that we moreover take an active role in improving the process in which we are learning.

This aversion has been institutionalized at Yale. In 1941, Dr. Samuel Harvey, chairman of Yale's Department of Surgery since 1924, wrote a paper analyzing the objectives of a medical education. With regards to evaluation, Dr. Harvey wrote that examinations were important, but, "if the purpose of [an exam was] disciplinary and direct control of the students' activities [then there was] no place for such [at Yale]." [22] Evaluations at Yale, as a result, are carefully vetted against the values of the highly valued Yale System by its protectors, the students, the faculty, the alumni, and the administration.

The Wednesday Evening Clinic

The Wednesday Evening Clinic (WEC) is a student run clinic in which senior medical students follow a cohort of patients in a primary care setting for a minimum of one year. Participating students can choose to work at the WEC to fulfill the primary care requirement. Students who choose to participate are typically taking an extra year(s) for a dual degree or for research. The clinic was started over 25 years ago and has been a perennial success from the point of view of patients, students, faculty and the Yale School of Medicine. The WEC's patients come largely from the underserved population of New Haven. The objectives of the clinic are the following: (1) to enforce and strengthen the clinical skills of the students who work in the clinic, (2) to serve as a longitudinal learning opportunity for Yale medical students not involved in clinical rotations, and (3) to provide high quality primary care to our patients.

Senior medical students administer the Wednesday Evening Clinic and follow primary patients longitudinally. We convene each week at 5 pm on Wednesday and each week one student prepares a didactic presentation on a topic of their choice. By 5:45 we begin seeing patients.

A total of twelve students staff the clinic, and each sees approximately three patients per night on average. The students are divided into three teams; "team leaders" are assigned by seniority and are typically M.D., Ph.D. students. Under their leadership, the team divides new patients assigned to the team, will cover for others in the case of absentees, or will discuss particularly interesting or difficult cases. Returning patients are scheduled to see the student who had seen them on their previous visit. Students call the patient into a room and perform a history and a physical exam. They then ask the patient to wait in the room while they review the case with an attending.

The attending and student then return to see the patient. At this point we review the history with the patient and the attending often ask additional questions and performs a focused physical examination. We then put together a plan with the patient for how to proceed. Finally, the student completes any paper work – such as ordering tests, referrals, prescriptions – and reviews the plan with the patient again. After having seen all the patients for that night, the student then writes the note into the electronic medical record system. Over the span of the week the student may sometimes need to follow-up on lab results, referrals or any other pending items.

The WEC presents various opportunities for participating in longitudinal care. First, it allows for developing a relationship with patients over many visits. This creates a greater sense of responsibility to patients, but that is accompanied by greater reward. Moreover, because it is student-run, it also empowers us as students to envision better ways of providing care to our patients. The founding students, for example, found that the primary clinic closed just as people were getting out of work, and so filled an important niche by offering primary care to both the underserved as well as those who could not afford to leave work for a medical visit[24].

Patients who want after-hours primary care remain a significant segment of the patient population at the WEC.

The model for implementation of the mini-CEX at the WEC

The model described below demonstrates that the implementation of the mini-CEX at the WEC aligns the interests of all stakeholders to further the mission of the clinic. Dr. Deming's philosophy to "constantly and forever improve the systems of production and service" served as the initial impetus for the creation of this model. The WEC has a dual mission to provide high quality care to patients while serving as an educational forum to teach high quality medicine to students. Any attempted improvements in the system of service should seek to improve both missions, or at the very least not compromise one for the betterment of the other.

Opening conversations with students on how to improve the system of service at the WEC revealed a broad opportunity for better clinical teaching. A 2001 article in Yale Medicine revealed that this issue had a history [24], and conversation with faculty confirmed that despite previous discussions and previous attempts towards instituting better feedback mechanisms, little had been realized. Indeed, prior to the implementation of this model, the WEC had little structured assessment of our performance and generally lacked formal infrastructure for helping students improve their clinical skills. Both of these gaps must be filled for the WEC to achieve its primary objectives as a teaching and learning experience and as a clinic for high quality patient care. The notion of systemically assessing and improving students' clinical skills addressed the dual mission of the clinic.

The mini-CEX could provide a structure for both: assessment and feedback. Through direct observation of multiple patient visits, students would gain insight into the maturation,

strengths, and weakness of their clinical skills and identify areas for remediation. Moreover, patients would doubly benefit: firstly, from additional the direct supervision of an attending physician during their visit, and secondly, from the improved clinical skills of their provider. However, the Yale System itself provides another philosophical tension that is important to consider before attempting any changes in medical education at Yale. The stated objective of the Yale School of Medicine is "to develop physicians who are highly competent and compassionate practitioners of the medical arts, schooled in the current state of knowledge of both medical biology and patient care" [23], which aligns well with the educational goals of the clinic. However, as a proud product of the Yale System, I was also well aware that any systemic changes must be sensitive to our well-ingrained skepticism against testing, requirements, and evaluations. Especially in a student-run clinic, changes in our medical education not well-aligned with our cherished approach to learning would face considerable resistance and I knew would be doomed to failure. Medical education at Yale had to be constructive, open-ended and non-competitive. The tension between the Yale system's ideal vision of medical education and the goal of becoming a highly competent physician backdrops against a strong body of evidence showing that medical students are consistently poor at evaluating their own competence [25, 26]. In a system that insists on self-directed learning, students will at a minimum need feedback to identify the gaps they need to address to maximize their potential. If presented as a summative assessment tool, Yale medical students would have a knee-jerk aversion against the mini-CEX; I know because I am a Yale medical student and I wouldn't stand for it at a clinic staffed and run by my peers, who volunteer their time. Indeed, the mini-CEX was not developed as a summative evaluation. But the elegance of the mini-CEX lies in its simplicity and its adaptability. Designed for myriad clinical settings and to adjust to the realities

of practicing medicine, it could similarly be adapted as a non-obtrusive feedback tool where students could use it to help them learn without feeling they were being graded or forced to compete with one another. On the contrary, it engages both trainee and trainer into an active dialogue about how to provide the highest quality patient care. The mini-CEX can be construed by both faculty and students as an evaluation tool, which may not align with a student-run clinic, it was not designed as such. If presented correctly, it could be repositioned to work well within a Yale clinic.

The final consideration in the implementation of the mini-CEX is the stakeholder who drives the entire process: the faculty ensures that patients get the high quality care they deserve, that students learn the high quality clinical skills they will need to become excellent physicians, and that the environment stays true to the values of the philosophy within which it functions. At the WEC, the faculty members volunteer their time, which implies a self-selected process of faculty dedicated to clinical education, but it also implies a higher opportunity cost. Implementation of the mini-CEX would have to respect the time and effort they dedicate to the clinic at the least, and ideally would augment the reward derived from teaching. However, the literature reassuringly evidences examiners routinely very satisfied with their experience with the mini-CEX. [14, 16] or their experiences during training as evaluators [27]. These findings were anecdotally confirmed by the subjective but unambiguous enthusiasm I received in informal conversations with attending physicians in the clinic conference room during the weeks I was learning more about medical education prior to the actual implementation trial. This gave me the confidence of knowing that at least a core of the faculty would support its implementation.

Satisfied that the various stakeholders in the clinic would all have aligned interests in the implementation of the mini-CEX at the WEC, we moved forward to test our hypothesis that its implementation would sustainably have a high level of utility. As always, theory is simpler than reality, and the implementation trial was begun with the knowledge that obstacles to its sustainability would require a realignment of interests to make the trial successful. Nonetheless, this would be the first reported student-initiated implementation of an assessment and feedback tool in medical education; it would be the first of its kind in a student-run clinic; and it would be the first within the Yale System.

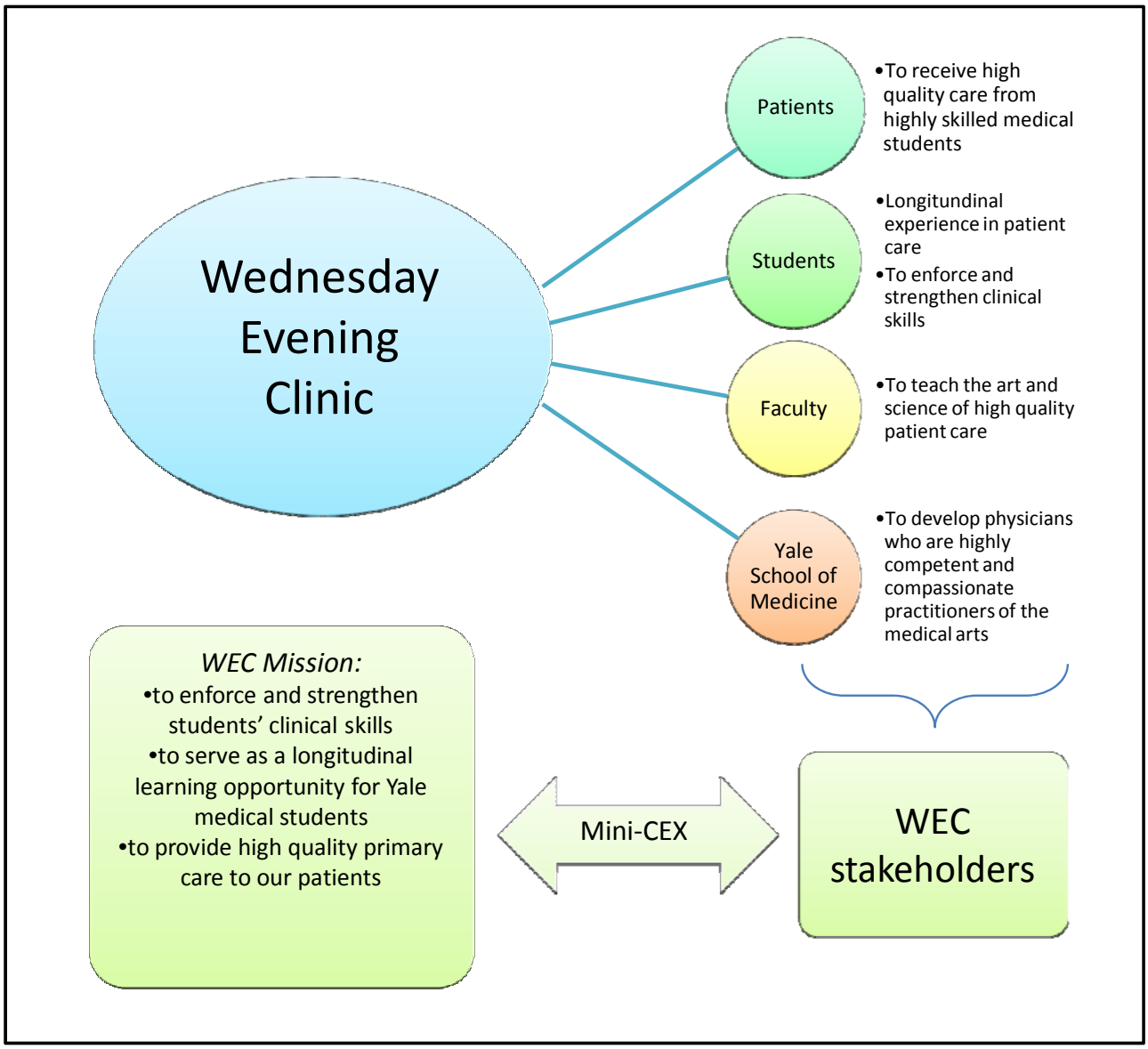

Figure 1: Model for implementation of the mini-CEX at the Wednesday Evening Clinic (WEC). The stakeholders of the Wednesday Evening Clinic are considered to be the patients who receive their care there, the students who organize and staff the clinic, the faculty who attend the clinic and the Yale School of Medicine who provides the infrastructure and support for the clinic. The primary objective for each stakeholder evidences each ones unique relationship with the WEC. At the same time, however, each one has an interest in a WEC staffed by providers with high quality clinical skills. The mini-CEX, a tool used for assessing and improving clinical skills, is in alignment with the interests of all involved stakeholders, and moreover furthers the WEC's mission to teaching and providing care.

Methods

Setting and participants

The implementation of the mini-CEX was done at the WEC, both are described in detail above. The students who staff the clinic volunteer to see patients from the underserved population of New Haven on a voluntary basis one night per week. The faculty are attending physicians either at Yale or in the community; with the exception of one faculty member, they had not previously been trained in use of the mini-CEX.

The patients at the clinic passively participated in the study. The Yale School of Medicine Human Investigation Committee reviewed the experimental design and found it to be exempt from IRB review under federal regulation. This study was conducted between May and September of 2008.

Focus groups

Implementation of the mini-CEX in a student-run clinic necessitates student willingness to add the tool to the administration of the clinic. A pre-intervention needs-based analysis focus group was arranged to gauge student response to this intervention. The structure of the focus group reflected those recommended and used in the literature [28, 29], with an opening question, an introductory question, key questions, and ending question and time to put all the parts together. The needs-based analysis sought to examine the perceived objectives of the clinic and whether students' priorities in the clinic were well aligned with implementation of the mini-CEX. Questions at the pre-intervention focus group assessed the perceived objectives of the clinic, assessment of achievement towards stated objectives, and possible interventions to more effectively achieve stated objectives. The mini-CEX was introduced after the needs-

based analysis as a teaching tool to facilitate teaching of clinical skills through direct observation, and the students were asked if they would want to incorporate the mini-CEX into the clinic.

Faculty training

Implementation of mini-CEX also necessitates faculty training. In accordance with previously published studies showing the importance of feedback training, faculty at the WEC were encouraged to voluntarily attend a four hour course offered through Yale's CME program on using the mini-CEX to evaluate students[27]. The faculty that could not attend the training session were trained by the faculty director of the clinic in the use of the mini-CEX on-site. This was recognized as sub-optimal but would serve as a preliminary training to be supplemented in the future if deemed necessary.

Intervention

The implementation trial of the mini-CEX was launched with the goal of a minimum of "one mini-CEX per team per week." In order to minimize interruption of patient flow, one student from each team would use the mini-CEX on their first visit. This was intended to reduce the possibility of unavailability of attending physicians for routine care. Initially the students were asked to find an attending physician to observe their first visit, and since there would be more than 3 attendings, this would have achieved the minimum goal. However, students were inconsistent in taking the initiative to seek out an observer, so the implementation model changed. Team leaders were put in charge of designating a student from their team for direct observation that week. Either the faculty clinic director or I would follow up with the team leaders to make sure someone from their team was observed that week.

Since the mini-CEX can be used even if only a fragment of the patient visit is observed, students were also encouraged to use of the mini-CEX with any visit by reaching out to faculty for direct observation if faculty were available. It was hoped that the group discussions framing the mini-CEX as a feedback tool to help us become better physicians would motivate students to seek out additional observation opportunities beyond the minimum standard.

Since the mini-CEX booklets used provide carbon copies of the evaluation, students were asked to submit one copy of their evaluation for data collection. The information was uploaded to a database on a weekly basis, with codes for student names unknown to faculty in order to emphasize the learning objectives of the mini-CEX and minimize student anxiety about being quantitatively evaluated.

Outcomes

As a utility study examining whether the mini-CEX could be sustainably implemented at the WEC, the primary outcome of the study was the number of mini-CEX's per week. A predefined threshold of success was set at "one mini-CEX per team per week." Another major outcome included whether the students would choose to keep the mini-CEX as part of their clinic after the trial period was over.

Secondary outcomes from this study included acceptability of the intervention by clinic stakeholders including faculty and students. Acceptability was measured by responses from faculty and students in one-on-one interviews done over the trial period. Notes summarizing the interviews were documented on a regular basis in a public WEC journal published on the internet.

Based on the copies of the completed mini-CEX's submitted on a weekly basis, it was possible to determine the students receiving feedback and the faculty performing feedback. The mini-CEX form also provided de-identified patient information such as the clinical setting (in this case always "ambulatory"), the nature of the patient's problem, type of visit (new or follow-up), patient age and sex, the level of complexity of the visit (low, moderate, or high), and the focus of the visit (data gathering, diagnosis, therapy or counseling). The submitted forms also provided quantitative data on the feedback the student received as well as written comments, if any. Finally, the evaluator also record the time spent observing and the time spent giving feedback.

Student opinions voiced during the post-intervention focus group were also analyzed as a secondary outcome. This outcome would qualitatively evaluate the entire experience from the student perspective and could serve as a forum for discussing the viability of maintaining the mini-CEX as part of the WEC culture.

Other data input from the mini-CEX would further inform the analysis of the study. In particular, duration of visits and feedback would serve as important quantitative benchmarks of the operational changes of implementing the mini-CEX. Further analysis could reveal associations between type of visit and patient demographics.

Analysis

The results of the needs-based analysis were qualitatively evaluated through a one-time process of identifying the main themes in the answers to each of the questions. Acceptance or rejection of the implementation of the mini-CEX as a tool well-aligned with the objectives of the WEC was the primary outcome of the needs-based analysis.

Feasibility was primarily determined by the pre-defined goal of "one mini-CEX per team per week," in other words, three mini-CEX's per week. The data recorded on the submitted mini-CEX forms was uploaded on a weekly basis to an online spreadsheet by a volunteer undergraduate student. Analysis of the spreadsheet data was done on Microsoft Office Excel 2007.

Results

This study examined the utility of implementing the mini-CEX at the WEC. The pre-intervention focus group resulted in the students at the clinic agreeing to try incorporating the mini-CEX at the WEC. During the trial, the threshold of three mini-CEX's per week was achieved each week in which there were at least five attending physicians. Each week 5 physicians are assigned to supervise the clinic; the weeks in which at least five were not present due to extenuating circumstances were discounted from the trial in this analysis (e.g. per protocol analysis for utility). Figure 2 shows the number of mini-CEX's per week during the trial period; the average was 3.82 evaluations per week (median = 3), including the understaffed weeks.

Figure 2: Number of mini-CEX's per week at the WEC

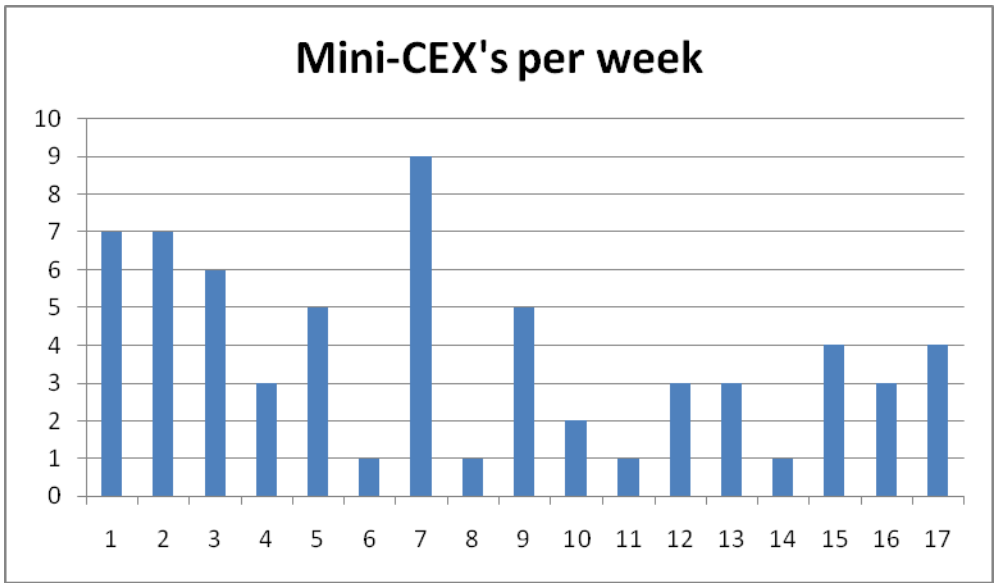

Figure 2: Number of mini-CEX's per week at the WEC. The goal of three mini-CEX's per week was sustainably achieved except on the weeks when the clinic was understaffed.

Needs based analysis: Focus Groups

The first question asked students to identify their perceived priorities of the clinic. Students identified the following distinct objectives: (1) to learn how to diagnose and evaluate patients for the most common primary care complaints, (2) to learn what are the right questions to ask during a patient encounter, (3) to learn correct clinical skills, (4) to maintain and improve clinical knowledge. Answers to the question of what the clinic is lacking to achieve these objectives ranged from home access to electronic medical records to knowing the patient's chief complaint before entering the patient's room. But multiple students recommended better feedback mechanisms on the history and physical exam, on written notes, and on their case presentation skills. The clinic chief pointed out that exit surveys from departing students

for several years have repeatedly commented on the lack of feedback from attending physicians regarding clinical skills, presentations, case management and written notes. Upon requesting suggested improvements to more purposefully fulfill the stated goals of the clinic, the group focused on finding ways to provide objective evaluations, specifically suggesting group discussion of specific cases, didactic classes similar to those in the primary care rotation, or changing the structure of the current pre-clinical presentations to focus on topics more directly applied to patient care. One student pointed out that asking for more feedback is a student responsibility.

Student and faculty participation

Over 18 different students staffed the clinic during the implementation. Twelve students typically staff the clinic with little change from week to week in attendance or in students enrolled as providers at the clinic. However, the implementation trial occurred during summer months, a time of change at the clinic, with students taking vacations and with students ending or beginning their time at the clinic. Nonetheless, as an aggregate, students averaged 3.3 evaluations over the trial, with a maximum of 10 and a minimum of one (see Figure 3). During the same time, 14 different faculty members evaluated students. On average the faculty performed 4.3 evaluations (std dev = 4.4 evaluations). Evaluator satisfaction had a median 8 (7.9 ± 1.1), and student satisfaction had a median of 9 (8.1 ± 1.2); both scales out of 9.

Figure 3: Number of mini-CEX's per student

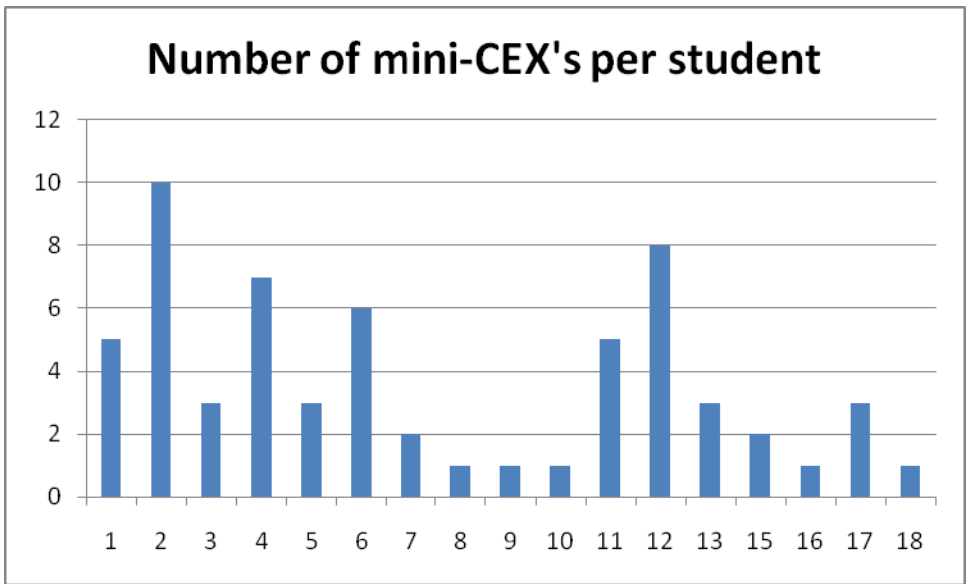

Figure 3: Number of mini-CEX's per student over the entire implementation trial. Although only twelve students staff the clinic at any one time, the trial occurred during the summer months, a time of greater flux in attendance. Many students did not receive the expected number of evaluations, perhaps also a by-product of the summer flux in attendance.

Visit demographics

The patients visited during the evaluations over the trial period had an average age of 46.5 years (stdev = 15.7 years); of the evaluations reporting sex of the patient, 36 of 47 were female (77%) and 31 of the 38 visits in which the type of visit was recorded were for follow up care as opposed to new patient visits (82%). Of the 50 evaluations that reported complexity of visit, 9 were classified as "low" (18%), 29 as "moderate" (58%) and 12 as "high" (24%). 44 of 65 evaluations (68%) specified the patient's clinical problem, including 18% with musculoskeletal pain (e.g. lower back pain, knee pain), 17% for management of hypertension, 14% for diabetes

mellitus, 8% with headache, 8% with a GI-related problems (e.g. BRBPR, GERD), 6% for routine care and 3% to address psychosocial problems (e.g. tobacco abuse).

Duration of observation and feedback

The median (mean [± SD]) time spent by faculty observing students was 30 (29.4 ± 18.2) minutes and median feedback time was 10 (9.4 ± 4.3) minutes. New patient visits took an average of 54.2 minutes (± 32.46) for observation, followed by an average of 10 minutes (± 4.47) of feedback. Follow up visits, on the other hand, took 26.8 (± 10.0) minutes for observation and 9.0 (± 4.3) minutes of feedback. The time required to observe a new patient visit was significantly longer (two-tailed t-test, $P < 0.001$) than a follow up visit. However, the time for feedback was essentially the same.

Student interviews

Throughout the implementation trial, students were candid with me in their experiences regarding the mini-CEX. In the first weeks, some students were less than eager to be observed, trying to avoid being observed despite agreeing that it was a good idea and that it was needed in the clinic.

Within two weeks though, the attitude was changing: one student who had told me that he's "not ready to be observed" told me on week three that he wanted to get feedback, but that another team member had beat him to it that week. On week three another student told me that she had a "great heart to heart" with one of the attending. She explained to me that they may have had the conversation anyways, but that "mini-CEX forced [them] to sit down and discuss the [patient] visit," which facilitated a conversation which otherwise might have happened in the hallway between visits.

Shortly thereafter, students began to have more concrete opinions regarding the mini-CEX. One student said that he liked receiving immediate feedback on his history and physical. Because he gets feedback right after the visit, he explained to me, the attending could give him specific points on what questions he should have asked, for example, or what are the nuances of the physical exam that he should have done. He further explained to me that he changes his approach to patients when being observed, "because you know you're being evaluated, it makes me go through things more methodically and more thoroughly, but it also helps me organize myself."

Within a month, students were expressing preferences towards receiving feedback from specific attending physicians, citing specific faculty members to me as good observers with valuable feedback versus others with whom they were less comfortable receiving feedback. Similarly, one student told me "feedback is inconsistent. Sometimes people give me a lot of feedback, sometimes none or very little. And what they tell me changes from one attending to another." Another one said, "there's a lot of variability across attending, so it depends who you get as your rater."

Another issue that surfaced with increasing frequency was the balance between positive and negative feedback. A variety of students told me they were getting too much negative feedback, which they understood they needed, but that reinforcement of the things they were doing right would help them feel good about being observed. One student who had been previously enthusiastic told me she didn't want to do it that week, "it makes me feel bad. I don't get so much constructive feedback as much as negative feedback."

By the end of the second month, multiple students told me that they basically gave very little value to the scores they received on the evaluations since they almost always get 6's or 7's.

This student's impression was that attending physicians knew that we were not ready to be senior residents so they gave negative feedback. He also lamented that too often the feedback was not specific enough to be constructive.

After several months, some students had the same attending physician observe them who had previously already observed them. One student explained to me that it was "excellent for [them] to have a basis of comparison in giving [him] feedback."

In one of the final weeks a student told his team leader, upon being assigned to do the mini-CEX, "not today, I have too many patients." His team leader jumped back at him and said, "No, that's perfect. Last week I had to do it and I said to myself, '[forget about] this feedback' and I just did it as I normally do and I got straight superiors for the first time ever." His team member was inspired and went to find an attending physician.

Faculty interviews

From the beginning the faculty praised the implementation of the mini-CEX as part of the teaching at the WEC. Even before the implementation, upon seeing a print out of an background search and HIC proposal I had left on a table, one attending had a conversation with me about the importance of bringing more feedback to the clinic at the level that is seen in so many other professions. He further argued for "360 degree" feedback, meaning he also wanted students to tell him how he was doing.

This same attending would be quite enthusiastic throughout the implementation trial, and several students privately told me that they always appreciate his feedback. He furthermore suggested that the clinic should have pictures of all the students in order to facilitate the relationship between students and faculty. To everyone's delight, he followed through on this

suggestion, creating a board with everyone's picture with a name tag. He is now called on to photograph new students joining the clinic.

One attending physician, after the first week of using it, enthusiastically told me, "This is the best thing. I had never heard [the student visit] before, [and during my observation] I was thinking how the presentation would have been different if I hadn't heard it, but I could tell from the patient exactly what was going on." He further went to explain he furthermore thought it was better for teaching: "I could see what [the student] was doing, and I could adjust it right then and there, 3 minutes later instead of 30 minutes later."

Another attending supported this point of view, explaining to me that "it is effective in making better training because [she] sees students' performance and knows if, for example, the student is disorganized and tangential or if the patient is a poor historian." She also had the opportunity to improve their history and physical and give them feedback, something she otherwise did not do.

The same week, an attending offered the unsolicited support of telling me that he thought it was great we were doing this. Another physician confided in me that he was thinking of not volunteering at the clinic any more, but that this intervention has convinced him to stay.

One attending told me that he really only volunteers at the WEC in order to fulfill a teaching requirement in order to maintain admitting privileges patients to Yale New Haven Hospital. When pressed on why he volunteered at the clinic as opposed to other ways of satisfying Yale's teaching requirement, he told me that it's not his priority when he is seeing patients in his office; the WEC does not take up office time. He further ceded that the mini-CEX is wonderful and it's the way teaching in medicine should be done.

Another attending told me that the feedback does not include a place for students to assess themselves; this, he argued, would make them have to introspect their own soft spots. He told me that as a result, before any feedback, he first asks students how they think did before he gives them his feedback.

The WEC was not the appropriate setting for teaching students clinical skills they do not know, argued another physician. He told me that if someone does not know how to listen to the heart you would need two weeks in cardiology clinic or on the cardiology floors to learn how to excel in listening to the heart. This material cannot be taught, or learned, in a five minute discussion.

Another attending discussed the policy of doing the mini-CEX's during the first visits in order to ensure minimal disruption of patient flow; this policy implied that almost all feedback sessions would be with follow up patients, and infrequently with new patients. She argued that new patients would be much better opportunities for feedback. Follow up patients were typically faster visits, and furthermore the attending, she argued, were not as familiar with the patients as the students were.

Discussion

This study examined the utility of the mini-CEX in the WEC using van der Vleuten's utility index, updated by the PMETB, to assess its success [2, 3]. Of the utility index's six variables – reliability, validity, educational impact, acceptability, feasibility and cost (see Table 2) – previous studies have examined the reliability and validity [13-17, 20, 21], which were beyond the scope of this study. Instead, this study focused on the acceptability, the feasibility, the educational impact, and the cost of implementing the mini-CEX at the WEC.

The initial proposal to implement the mini-CEX was well-received by the students at the WEC. The threshold of 3 mini-CEX's per week was achieved each week that the clinic was fully staffed. Feedback during the trial revealed a receptive but critical student body and a faculty that, although inexperienced, was generally supportive of the implementation of the mini-CEX.

Acceptability

As described in the model (see Figure 1), the implementation of the mini-CEX soundly aligns the mission of the clinic with each of the WEC stakeholders, namely the students who run and staff it, the faculty, the patients and the Yale School of Medicine. This well though-out implementation plan facilitated a high level of acceptability.

One of the tenets of the Yale System since its inception by Dean Winternitz in the early 1920's presumes medical students to be mature, self-sufficient graduate students who, in the words of Viseltear, "were not to be held by the hand or regimented" [22]. Instead, we take pride in exerting agency in the process of becoming high quality physicians, and the mini-CEX empowers us to reach out to faculty to create a space for one-on-one, patient-specific teaching and learning.

The needs-based focus group prior to the implementation evidenced the students' desire to improve their own learning experience. Upon agreeing that the development of clinical skills was a primary objective and that this was not sufficiently addressed in the operation of the clinic, the students agreed to incorporate the mini-CEX into the WEC. This was not a symbolic agreement: throughout the implementation trial, students agreed to change the normal operations of the clinic to incorporate the mini-CEX. The feedback I received demonstrated active questioning of the value of the tool, further demonstrating the students' agency in

forming their own medical education. The students' high satisfaction scores (mean 8.1) exceeded the previously published studies using the mini-CEX with medical students [18, 19]. The relatively unique culture and context of the WEC in contrast with more typical clerkships may account for this difference.

Importantly, based on student interviews as well as the debriefing after the trial, two main factors greatly hindered the mini-CEX's acceptability over time. First, the inconsistency of the feedback students were receiving from attending physicians generated frustration from students. This inconsistency seems to have led them to devalue the feedback they were receiving. Second, students repeatedly said that feedback has to be double edged: it should reinforce strengths and point out weaknesses. However, several students mentioned that too often the feedback was only negative, leading to negative reactions to getting feedback.

In the Handfield-Jones, et al model linking assessment with practice performance and learning, positive feedback *validates* actual practice, driving the trainee closer to the "desired practice"; discrepancies between desired and actual practice are identified as an "educational gap" to be addressed through contemplation and "educational action" [30]. Without the rewarding validation of positive feedback, the contemplation needed to close the gap between desired and actual practice may never occur.

The Association of Medical Education in Europe's (AMEE) Guide on "Effective educational and clinical supervision" reemphasizes the evidence demonstrating the wide variations in the frequency and amount of supervision [31]. The authors, further point out, however, that one key feature of effective supervision is, "constructive feedback is essential and should be frequent" [31]. The repeated, unsolicited comments from students pointing out the lack of

constructive feedback imply that the feedback did not meet all the criteria for effective supervision.

The acceptability from faculty was more varied than the students, although satisfaction was high (mean = 7.9), and again exceeded the values from previous implementations with medical students [18, 19]. However, the faculty are perhaps the stakeholder least aligned with the implementation of the mini-CEX. I learned in my interviews that the faculty have another incentive in attending the clinic. The fulfillment of a teaching requirement to receive admitting privileges at Yale New Haven Hospital by attending at the WEC represents a significant proportion of attending physicians at the WEC. Those who attend the clinic for ulterior purposes may not be as well aligned to teach at the clinic as other faculty.

Teaching and working for admitting privileges are not mutually exclusive: several of the community physicians who attend the WEC, presumably receiving admitting privileges in the process, were repeatedly mentioned to me, by name, as the best teachers and evaluators. Nonetheless, there exists a potential misalignment of interests, and as such, faculty acceptability of the implementation was less unanimously enthusiastic than with the students.

Feasibility

A variety of previous studies have examined the feasibility of implementing the mini-CEX in a variety of contexts [13-16, 18-20, 32-34]. This is the first known implementation of the mini-CEX in a longitudinal setting with medical students. Furthermore, this is the first reported implementation of a mini-CEX implemented by students to assess themselves.

The implementation model functioned well to the standards it was set to: three mini-CEX's per week. Its success should largely be attributed to students, who were invested in maintaining

this feedback mechanism throughout the trial. However, in light of the various "below" threshold weeks, the model is not self-sustaining. Indeed, throughout the trial, either I or the faculty director of the WEC continually checked in with students, and later with team leaders, to make sure the mini-CEX was being done.

The clinic that has run relatively unchanged for more than 25 years, priding itself on the longitudinal care of the student providers, and more so on the lasting influences of attending physicians who have taught at the WEC for ten years, often more. The culture of the WEC is stable and conservative in its resistant to drastic change. The mini-CEX's inculcation into the WEC will require a cultural change with a strong investment of energy over a prolonged period of time.

More importantly, however, the need for more consistent, trained evaluators became clear in the interviews and in the discussion above regarding acceptability. The best results with the mini-CEX necessitate the evaluator partake in a specifically-designed course, typically lasting at least 3 hours [27]. Even a half-hour instruction does not lead to the changes in assessment in evaluators necessary for optimal results [27].

However, the WEC's faculty consists of a rotating cadre of approximately 30 different physicians. Four showed up for a short course on the mini-CEX arranged prior to the implementation, the rest received brief on-the-spot training. Although suboptimal, we proceeded with this level of training for the implementation trial. These experiences suggest that a low probability of being able to adequately train the current team of attending physicians. But cognizant of the importance this holds to the WEC experience, the WEC could choose to grandfather in the current attendings, and mandate that all new attending physicians at the WEC must receive the full day of mini-CEX training.

Cost effectiveness

Cost effectiveness balances the cost of an assessment against its benefits. Though not empirically measured in this study, the experience of this implementation shows a high up-front cost with a rapid subsequent decrease in costs incurred. In particular, the interviews conducted in this study qualitatively reinforce the need for better evaluator training[27] in order to use the mini-CEX to its full potential benefit as an educational tool. However, the cost of training all the faculty at the clinic represented a high barrier to implementation that we chose to forgo. Instead, we prioritized the creation a feedback structure provided by the implementation of the mini-CEX, even without formal evaluator training. In this case, we feared that prioritizing training for all faculty before implementation would realistically translate into no implementation at all. Evidence suggests that any feedback is better than no feedback at all, and management literature suggests that supervision can be seen as a form of quality assurance[9]; this study supports both arguments.

The benefits of education represent a more theoretical argument; in medical education, ultimately the purpose of education is to improve patient care. The benefits of education represent a more theoretical argument; in medical education, ultimately the purpose of education is to improve patient care. As shown by the cost effectiveness formula, benefits can be divided into immediate effects and longer-term effects of the intervention:

$$CE = [Cost(time)] - [current\ benefits + future\ benefits]$$

The firsthand benefits of effective supervision in the clinical setting are well-documented, as reviewed by Kilminster and Jolly [9]. For example, supervision has been shown to result in better patient outcomes, while lack of supervision is harmful to patients. Supervision seems to facilitate trainees learning of new skills, and trainees are able to identify many gains from

supervision. Anecdotally, students seemed to confirm a honing of clinical skills, with a steeper learning curve than in their previous time at the clinic. Equally importantly, supervision provides a quality control mechanism which would be lauded by quality guru Dr. W.E. Deming (see Appendix A: Deming's 14 Points). The current standard of requiring graduating medical students to take the USMLE Step 2 CS exam corresponds to mass inspection, which, as Deming points out, is effectively equivalent to planning for defects. Instead, Deming urges reliance on modern methods of training and supervision to reduce the idiosyncratic variability inherent in any production process.

The mini-CEX provides both supervision and training through observation and feedback. These benefits started to become apparent even in the earliest conversations with students and faculty. As one student pointed out, the mini-CEX allowed for visit to visit fine tuning of clinical skills. If each patient visit with a trainee represented such an opportunity to fine tune clinical skills, surely the outcome of the process – a graduating medical student or a senior resident – would be a highly skilled physician which would render futile a test such as the USMLE step 2 CS. The faculty's surprised response after a first observation session evidences the obvious lack of supervision prior to implementation, but moreover reflects the gap between perceived student skills and actual clinical skills. More importantly, the faculty member immediately recognized that through observation, he could close that gap, and that through feedback he could improve the student's skills. Even one attending who lamented that the WEC was not the appropriate setting for teaching students clinical skills with which they are not already comfortable, nonetheless conceded that through observation he was able to identify weaknesses in the students' clinical skills. His frustration stemmed from the lack of sufficient time to truly correct what he perceived as such large deficiencies. However, if anything the

gravity of the deficiency only proves the necessity of observation and feedback. The results from our intervention qualitatively reinforce the need for more observation and feedback, and anecdotally confirm some of the short-term benefits reaped in the intervention.

The long-term benefits of better skilled clinicians begins with patient care. Multiple studies have documented a perceived gap by residency graduates between the skills required in clinical practice and the training they received several years after they had graduated from residency [35-37]. Simply stated, this implies that what they did not learn during their training, they learned on the job. Undoubtedly learning under supervision is safer for patients; what mistakes were made along the way while clinicians fill those gaps? The implications of better clinical skills development during training has longer-term implications for patient safety beyond training.

Handfield-Jones et al suggest a paradigm for learning over a career that acknowledges the importance of continuous improvement, dependent on a systematic approach to "practice generated identification of learning needs" [30]. Currently, physicians are remarkably inaccurate in self-assessing their strengths and weaknesses [38]. A model that allows for self-directed continuous improvement requires clinical assessment and correction must begin during training, argue Handfield-Jones et al. [30]. By setting a new standard of teaching during the years of clinical medical education, trainees will not demonstrate the anxiety students at the WEC initially displayed at the prospect of being observed.

Furthermore, as healthcare becomes increasingly standardized and regulated, it is likely that providers will increasingly be assessed on their clinical competence. An assessment system created by physicians for physicians will translate to an assessment system that best serves the interests of providers and their patients, as opposed to the many other powerful stakeholders

in the healthcare industry. As a result, another long-term benefit of implementing the mini-CEX is the culture of assessment, self-regulation and quality control, which would ensure that physicians retain the power of dictating their own standards of competency.

In summary, the benefits of the implementation of the mini-CEX include immediate and long-term improvement in patient care, a quality control mechanism that finely tunes the trainee's skills throughout the education process, faster acquisition of clinical skills, an assurance that trainees graduate from their programs with the clinical skills needed to provide care, a culture of self-assessment that encourages lifetime continuous improvement, and a system of self-regulation.

Next Steps and Challenges

Education is an investment in the future; calculating the present value of an educational investment is notoriously difficult. In preparing for the implementation of the mini-CEX at the WEC, we found it unrealistic to expect to train all the faculty through formal training before implementing in a timely fashion. Instead, we chose to forgo this cost. A possible next step would measure the costs of implementation in order to quantitatively assess the potential return on the invested time and resources to properly implement the mini-CEX in a clinical setting. A theory for the cost structure is presented below.

Cost effectiveness (CE) could be assessed by weighing costs against benefits, as discussed above. The benefits of education can be better understood by subdividing them into current and future benefits. On the other hand, costs can be more objectively assessed as a function of time:

$$CE = [Cost(time)] - [current\ benefits + future\ benefits]$$

The costs break down into three main categories (denoted as cost by the subtext, "c"): (1) time (T_c) investment required to sustain the mini-CEX, (2) training faculty (F_c), and (3) time per patient (P_c). The graphical depiction of the cost of implementation is shown in Figure 4.

The eventual goal of integrating the mini-CEX into the fabric of the WEC optimistically implies that the additional time investment to sustain the mini-CEX would decrease logarithmically over time. The rate at which the time investment would decrease, λ, is dependent on context and initial acceptability.

$$T_c(t) = T_c(0)\, e^{-\lambda t}$$

As shown in Figure 4A, the tool becomes increasingly integrated, the additional investment necessary to maintain it would become negligible. However, the initial period requires significant time and energy toward ensuring a full integration of the tool into the cultural fabric of the clinic.

Training faculty would include a front-loaded investment to train all current attending physicians. Thereafter the investment would be a sustained, relatively low investment, I_s, proportional to the rate of turnover of faculty at the WEC, *r*. The duration of time necessary for this initial investment, *t*, and the level of investment, I_0, would determine the initial part of the curve, as described by the following equation and depicted in Figure 4B:

$$F_c = (t * I_0) + (r * I_s)$$

Finally, time per patient will increase with the implementation of the mini-CEX. This study found that the median time observing was 30 minutes plus an additional 10 minutes for feedback, resulting in a time investment of 40 minutes. However, when the observation occurs during the first visit of the day, time observing does not affect patient flow. As a result, the only time cost to patient flow is the feedback time. Our median time of 10 minutes is within a 5

minute range of other studies using the mini-CEX with students. An additional 5 to 15 minutes per patient observed becomes insignificant when distributed over the entire patient population seen, but every minute is precious in clinic, and so increased time per patient is weighted by a coefficient, w_t, depicted in Figure 4C:

$$R_t = w_t(feedback\ time_{avg} \times percentage\ of\ patients\ observed)$$

Though each investment is in different units, if weighted to account for opportunity cost then each one can be compared and summed to calculate cost over time to show the cost of implementing and sustaining the mini-CEX.

$$Cost(time) = T_0 \mid R_0 \mid P_0$$

As shown in Figure 4D, the initial investment would require a significant initial investment, especially to train faculty and to integrate the tool as part of the pre-existing clinical culture. Nonetheless, after the initial investment required to train faculty, the costs drop dramatically and continue to decrease toward a baseline value.

With competing resources for multiple missions, departments of education at medical schools would want to know that the time, energy and funding dedicated to the implementation of the mini-CEX will efficiently result in a furthering of their goals.

Limitations

From the outset this study sought only to examine certain facets of the utility of the mini-CEX, namely the feasibility, educational impact, acceptability and cost effectiveness. Its implementation at the WEC limited the range of patient encounters and the number of trainees exposed to the assessment tool. The data from the intervention does not have a control group per se, but our hypothesis tested whether or not it could be consistently used during the trial

period. The interviews with both faculty and students resulted in anecdotal evidence without statistical power. However, the success of the intervention lies in its current staying power at the WEC as a feedback tool to improve the clinical skills of current and future volunteer students.

Figure 4: Cost of implementation of mini-CEX over time

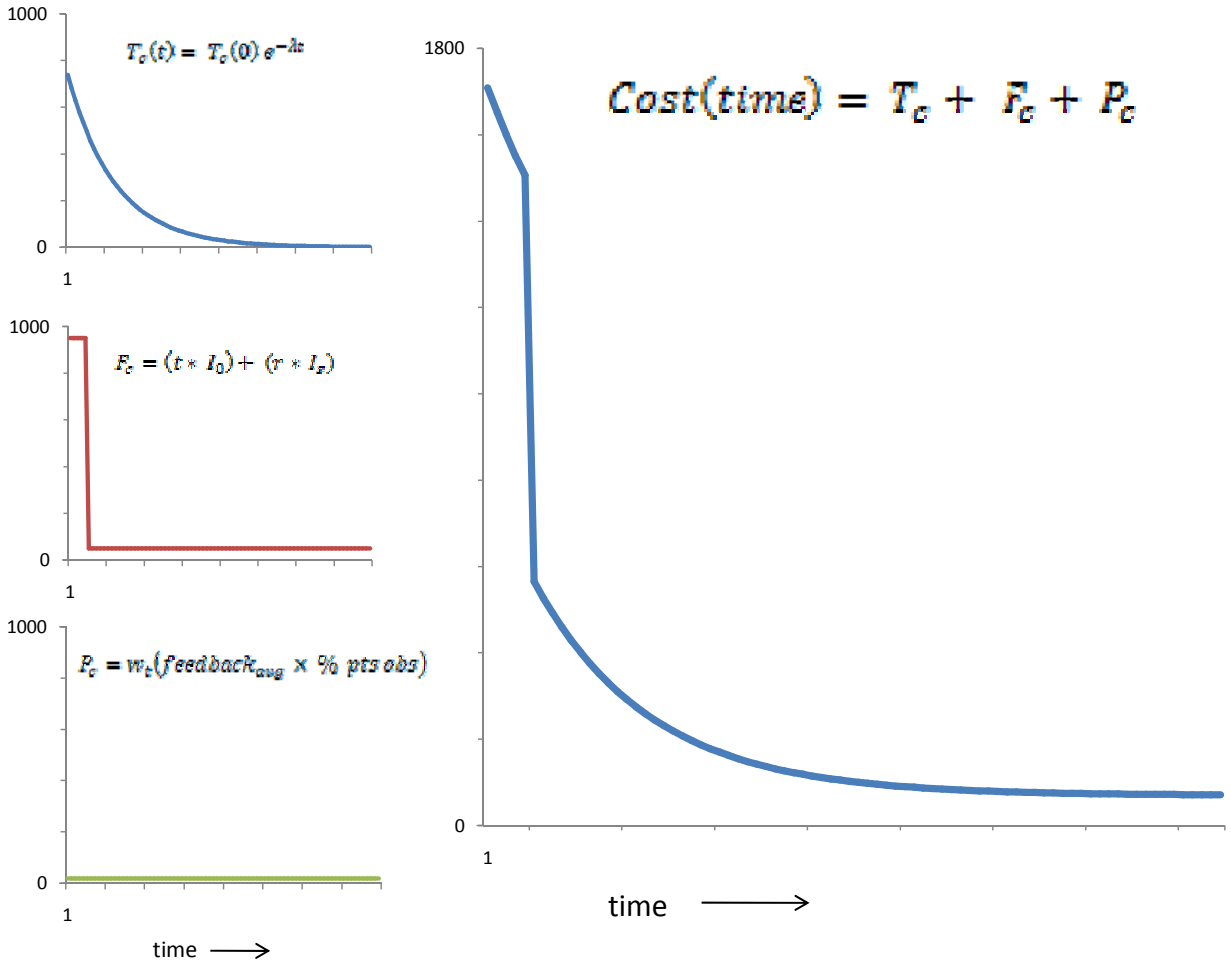

Figure 4: Each of the three cost components is depicted over time, and the aggregate is shown in the large graph to the right. (A) The time and energy that would have to be invested on a daily basis to incorporate the mini-CEX into the cultural fabric of the clinic decreases logarithmically over time. (B) The time and resources necessary to train faculty as evaluators using the mini-CEX would require a large investment in order to bring all current staff up to standards, but would then remain at a baseline level dependent on turnover rate. (C) The cost of additional time per patient is minimal since the feedback time ranges between 5 – 15 minutes. However, cognizant that clinical time is precious, this is weighted by a factor, w_t. (D) The aggregate of these costs is shown as a large investment in the initial phase, which would dramatically drop after all faculty has been trained and would continue to taper to a baseline level.

Conclusions

The WEC's unique context for providing care lacked sufficient supervision for the clinic to maximally achieve its stated objectives to provide the highest quality care to its patients while creating an environment for students to strengthen their clinical skills. The trial implementation of the mini-CEX sought to create an infrastructure to improve both of these goals. The success of the implementation is evidenced not only by the attainment of the pre-determined goal of three mini-CEX's per week, but moreover by the resulting satisfaction of both the students and the faculty.

As the national discourse in both public and professional spheres increasingly focuses on patient safety and quality in healthcare, the importance of teaching and assessing the clinical skills of practitioners is becoming increasingly important. The mini-CEX is a proven tool for providing both, and its successful implementation in the WEC provides a model for the Yale School of Medicine and for other medical schools to replicate.

The bottom-up, student-driven implementation of the mini-CEX at the WEC sets forth a model for a career of self-directed learning, a trait inherent to a career in the ever-evolving field of medicine. The proposal for its implementation was by me, a student provider at the clinic, and its implementation, use and feasibility was promoted by my peers, the rest of the providers at the clinic. Its sustainability depends on maintaining our support, which in turn depends on us seeking to constantly and forever improve our patient care.

Appendix A: Deming's 14 Points[1]

1. **Create constancy of purpose for improvement of product and service.** Management must change from a preoccupation with the short run to building for the long run. This requires dedication to innovation in all areas to best meet the needs of customers.
2. **Adopt the new philosophy.** Shoddy materials, poor workmanship, defective products, and lax service must become unacceptable.
3. **Cease dependence on mass inspection.** Inspection is equivalent to planning for defects; it comes too late and is ineffective and costly. Instead, processes must be improved.
4. **End the practice of awarding business on price tag alone.**
5. **Constantly and forever improve the systems of production and service.** Waste must be reduced and quality improved in every activity: procurement, transportation, engineering, methods, maintenance, sales, distribution, accounting, payroll, customer service, and manufacturing. Improvement, however, does not come from studying the defects produced by a process that is in control but from studying the process itself. Most of the responsibility for process improvement rests with management.
6. **Institute modern methods of training on the job.** Training must be restructured and centered on clearly defined concepts of acceptable work. Statistical methods must be used fro deciding when training has been completed successfully.
7. **Institute modern methods of supervising.**
8. **Drive out fear.**
9. **Break down barriers between departments.**
10. **Eliminate numerical goals for the work force.** Targets, slogans, pictures, and posters urging people to increase productivity must be eliminated. Most of the necessary changes are out of workers' control, so such exhortations merely cause resentment. Although workers should not be given numerical goals, the company itself must have a goal: never-ending improvement.
11. **Eliminate work standards and numerical quotas.**
12. **Remove barriers that hinder the hourly workers.**
13. **Institute a vigorous program of education and training.**
14. **Create a structure in top management that will push every day on the above 13 points.**

[1] Demings original words in bold, summary explanation of relevant points taking from "A Note on Quality." 39.
A Note on Quality: The Views of Deming, Juran and Crosby". Harvard Business School Case Study, 1990(Case #: 9-687-011).

Works Cited

1. *Learning objectives for medical student education-guidelines for medical schools: report I of the Medical School Objectives Project. [Article]*: Academic Medicine January 1999;74(1):13-8.
2. Joshi, H. *Developing and maintaining an assessment system - a PMETB guide to good practice*. Postgraduate Medical Education and Training Board 2007 [cited 2008 December 22, 2008]; Available from: https://www.pmetb.org.uk/fileadmin/user/QA/Assessment/Assessment_good_practice_v0207.pdf.
3. van der Vleuten, C., *The Assessment of Professional Competence: Developments, Research and Practical Implications.* Advances in Health Sciences Education, 1996. **1**: p. 41-67.
4. Moercke, A.M. and B. Eika, *What are the clinical skills levels of newly graduated physicians? Self-assessment study of an intended curriculum identified by a Delphi process.* Med Educ, 2002. **36**(5): p. 472-8.
5. Stillman, P.L., et al., *An assessment of the clinical skills of fourth-year students at four New England medical schools.* Acad Med, 1990. **65**(5): p. 320-6.
6. Ramsey, P.G., et al., *History-taking and preventive medicine skills among primary care physicians: an assessment using standardized patients.* Am J Med, 1998. **104**(2): p. 152-8.
7. Lypson, M.L., et al., *Assessing residents' competencies at baseline: identifying the gaps.* Acad Med, 2004. **79**(6): p. 564-70.
8. Pulito, A.R., et al., *What do faculty observe of medical students' clinical performance?* Teach Learn Med, 2006. **18**(2): p. 99-104.
9. Kilminster, S.M. and B.C. Jolly, *Effective supervision in clinical practice settings: a literature review.* Med Educ, 2000. **34**(10): p. 827-40.
10. Steinert, Y., et al., *A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education: BEME Guide No. 8.* Med Teach, 2006. **28**(6): p. 497-526.
11. Reilly, B.M., *Inconvenient truths about effective clinical teaching.* Lancet, 2007. **370**(9588): p. 705-11.
12. Gilliland, W.R., et al., *Changes in clinical skills education resulting from the introduction of the USMLE step 2 clinical skills (CS) examination.* Med Teach, 2008. **30**(3): p. 325-7.
13. Norcini, J.J., et al., *The mini-CEX (clinical evaluation exercise): a preliminary investigation.* Ann Intern Med, 1995. **123**(10): p. 795-9.
14. Alves de Lima, A., et al., *Validity, reliability, feasibility and satisfaction of the Mini-Clinical Evaluation Exercise (Mini-CEX) for cardiology residency training.* Med Teach, 2007. **29**(8): p. 785-90.
15. Malhotra, S., R. Hatala, and C.A. Courneya, *Internal medicine residents' perceptions of the Mini-Clinical Evaluation Exercise.* Med Teach, 2008. **30**(4): p. 414-9.
16. Norcini, J.J., et al., *The mini-CEX: a method for assessing clinical skills.* Ann Intern Med, 2003. **138**(6): p. 476-81.

17. Holmboe, E.S., et al., *Construct validity of the miniclinical evaluation exercise (miniCEX).* Acad Med, 2003. **78**(8): p. 826-30.
18. Hauer, K.E., *Enhancing feedback to students using the mini-CEX (Clinical Evaluation Exercise).* Acad Med, 2000. **75**(5): p. 524.
19. Kogan, J.R., L.M. Bellini, and J.A. Shea, *Implementation of the mini-CEX to evaluate medical students' clinical skills.* Acad Med, 2002. **77**(11): p. 1156-7.
20. Nair, B.R., et al., *The mini clinical evaluation exercise (mini-CEX) for assessing clinical performance of international medical graduates.* Medical Journal of Australia, 2008. **189**(3): p. 159-61.
21. Norcini, J.J., et al., *Examiner differences in the mini-CEX.* Adv Health Sci Educ Theory Pract, 1997. **2**(1): p. 27-33.
22. Viseltear, A., *The Yale Plan of Medical Education: The Early Years.* The Yale Journal of Biology and Medicine, 1986. **59**: p. 627-648.
23. *The Yale System of Medical Education*. Yale University School of Medicine website, Admissions Office 2007 Jun 29, 2007 [cited; Available from: http://info.med.yale.edu/education/admissions/yale_system/index.html.
24. Curtis, J., *Learning for the long run.* Yale Medicine, 2001(Spring).
25. Nikos Mattheos, A.N.E.F.-N.R.A., *The interactive examination: assessing students' self-assessment ability.* Medical Education, 2004. **38**(4): p. 378-389.
26. Vicki, L., *Not knowing that they do not know: self-assessment accuracy of third-year medical students.* Medical Education, 2006. **40**(2): p. 173-179.
27. Holmboe, E.S., R.E. Hawkins, and S.J. Huot, *Effects of training in direct observation of medical residents' clinical competence: a randomized trial.* Ann Intern Med, 2004. **140**(11): p. 874-81.
28. Krueger, R., *Developing questions for focus groups.*, in *The focus group kit*, K.R. Morgal DL, Editor. 1998, SAGE Publications: Thousand Oaks.
29. Green, M.L. and T.R. Ruff, *Why do residents fail to answer their clinical questions? A qualitative study of barriers to practicing evidence-based medicine.* Acad Med, 2005. **80**(2): p. 176-82.
30. Handfield-Jones, R.S., et al., *Linking assessment to learning: a new route to quality assurance in medical practice.* Med Educ, 2002. **36**(10): p. 949-58.
31. Kilminster, S., et al., *AMEE Guide No. 27: Effective educational and clinical supervision.* Med Teach, 2007. **29**(1): p. 2-19.
32. Howley, L.D. and W.G. Wilson, *Direct observation of students during clerkship rotations: a multiyear descriptive study.* Acad Med, 2004. **79**(3): p. 276-80.
33. Wilkinson, J., Crossley JGM, Wragg A, Mills P, Cowan G, Wade W, *Implementing workplace-based assessment across the medical specialties in the United Kingdom.* Medical Education, 2008. **42**(4): p. 364-373.
34. Torre, D.M., et al., *Feasibility, reliability and user satisfaction with a PDA-based mini-CEX to evaluate the clinical skills of third-year medical students.* Teach Learn Med, 2007. **19**(3): p. 271-7.
35. Card, S.E., L. Snell, and B. O'Brien, *Are Canadian General Internal Medicine training program graduates well prepared for their future careers?* BMC Med Educ, 2006. **6**: p. 56.

36. Mandel, J.H., et al., *Preparation for practice in internal medicine. A study of ten years of residency graduates.* Arch Intern Med, 1988. **148**(4): p. 853-6.
37. McPhee, S.J., et al., *Training in a primary care internal medicine residency program. The first ten years.* Jama, 1987. **258**(11): p. 1491-5.
38. Davis, D.A., et al., *Accuracy of physician self-assessment compared with observed measures of competence: a systematic review.* Jama, 2006. **296**(9): p. 1094-102.
39. *A Note on Quality: The Views of Deming, Juran and Crosby".* Harvard Business School Case Study, 1990(Case #: 9-687-011).