ORIGINAL REPORTS

Initial Comparison of Resident and Attending Milestones Evaluations in Plastic Surgery☆,☆☆

Amy Yao, BS, Benjamin B. Massenburg, BA, Lester Silver, MD, MS and Peter J. Taub, MD

Division of Plastic and Reconstructive Surgery, Icahn School of Medicine, New York, NY

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

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

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

CONCLUSIONS: The remarkable range of correlations suggests that expectations for performance standards may vary widely between residents and program directors. Understanding gaps between expectations and performance is vital to inform current and future residents as the restructuring of the accreditation process continues. (J Surg Ed ▮:▮▮▮-▮▮▮. © 2017 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: milestones, resident, education, programs

COMPETENCIES: xxx, xxx, xxx

INTRODUCTION

Graduate medical education has begun to explore competency-based evaluation of resident performance. The Milestones project, part of the comprehensive Next Accreditation System (NAS) recently implemented by the Accreditation Council for Graduate Medical Education (ACGME), represents this paradigm shift toward outcomes-based resident assessment. As the cornerstone of the NAS, the Milestones were introduced to ensure uniformity in measuring resident knowledge and technical ability using a series of specialty-specific achievements.[1] The NAS aims to better equip residents to practice independently by continually monitoring their progression through increasing levels of expertise in 6 domains of clinical skills, termed *Core Competencies*, which were initially developed by the ACGME in 1999. Five levels of achievement span from the basic knowledge/skill level of a first-year resident (level 1) to the advanced understanding/performance level of someone in practice for several years, above and beyond the target outcomes of residency (level 5), with many important intermediate phases in between. Level 4 is designated as the appropriate target for graduation (Table). In July 2013, 7 medical specialties implemented the first phase of the NAS, including Diagnostic Radiology, Emergency Medicine, Internal Medicine, Neurological Surgery, Orthopedic Surgery, Pediatrics, and Urology. Residents in these fields have been evaluated using the Milestones competencies, and the NAS is beginning to be validated

☆A poster version of this article was presented at the Medical Student Program of the American College of Surgeons Clinical Congress in Chicago, Illinois on October 4, 2015.

☆☆The abstract of this article is being published in *Plastic and Reconstructive Surgery: Global Open*, an open-access journal.

Correspondence: Inquiries to Peter J. Taub, MD, Department of Surgery, Icahn School of Medicine at Mount Sinai, 5 East 98th Street, 14th Floor, Suite B, Box 1259, New York, NY 10029; fax: (212) 241-5999; e-mail: peter.taub@mountsinai.org, amy.ye.yao@gmail.com

  1931-7204/$30.00 http://dx.doi.org/10.1016/j.jsurg.2017.02.001

TABLE. Numbered Levels for NAS Milestones

Level 1	The resident demonstrates milestones expected of an incoming resident.
Level 2	The resident is advancing and demonstrates additional milestones, but is not yet performing at a mid-residency level.
Level 3	The resident continues to advance and demonstrate additional milestones, consistently including most milestones targeted for residency.
Level 4	The resident has advanced so that he or she now substantially demonstrates the milestones targeted for residency. This level is designated as the graduation target.
Level 5	The resident has advanced beyond performance targets set for residency and is demonstrating "aspirational goals", which might describe the performance of someone who has been in practice for several years. It is expected that only a few exceptional residents will reach this level.

with the data that are available.[2-6] National organizations can use the NAS data to evaluate training programs and institutions, to assess program growth via self-comparison over subsequent years, and to identify areas for innovation.[1] Similarly, Milestones data may be used to set tangible goals during semiannual resident feedback to better focus resource and time allocation in resident education.[1,3] After just 1 year, certain programs felt that they were able to provide better feedback that was better received by the trainees.[7]

In July 2014, a second phase of the NAS was implemented that included Plastic Surgery programs.[8] During the validation process of the Plastic Surgery Milestones, initial feedback from residents was largely positive, many of whom highlighted the value of having concrete expectations and standardized benchmarks during their years of training. Program directors also expressed confidence in the NAS's improved ability to identify potential weaknesses at earlier stages.[9] However, a recent survey found that only 22% of plastic surgery programs feel completely prepared to implement Milestones assessments.[9] Because the Milestones are specialty-specific, we cannot infer that one specialty's success with Milestones implementation represents success overall. There is a need for specialty-specific Milestone implementation data to assess the effect and validity of the restructured evaluation process.

A comparison of the level of agreement between residents and attendings on the Milestones is one benchmark that can be used to gauge preliminary response to the changes. Additionally, these data may be used to identify potential gaps between personal objectives and expectations of competency, thus providing a baseline for meaningful discussion of how individual residents and programs can improve. The objective of this study was to determine the correlation between residents' self-evaluations and program directors' assessments of their performance.

MATERIALS AND METHODS

Thirty-six Plastic Surgery Milestones have been developed in 6 core competencies, each with a list of detailed descriptors for each of the 5 levels of achievement (Fig. 1). The 6 categories include Medical Knowledge (MK) and Patient Care (PC) as well as qualitative practical skills, such as Systems-based Practice (SBP), Practice-based Learning (PBL), Professionalism (PROF), and Interpersonal Communication (COMM). Within each of the 6 levels of performance is a set of skill and behavioral requirements that must all be achieved before the level is deemed complete. Each set of requirements corresponds with the expected performance for that level, with level 4 being the target for residency graduation.

In this study, the authors analyzed Milestone self-assessments completed by the residents and identical Milestone evaluations completed by the faculty. The study population comprised 12 residents across 6 years from a single ACGME-approved, integrated training program in Plastic Surgery. The program's Clinical Competency Committee (CCC) consists of 8 faculty members, full-time and voluntary, as well as Site Directors from each rotation site, selected from 24 total faculty members. The CCC underwent a formal training session led by the Associate Program Director before evaluating each resident in each of the 36 Milestones. Scores for each Milestone were made by consensus during a meeting of all 8 members of the CCC. Final scores of the CCC were recorded by the Program Director and Associate Program Director. Assessments were determined based on a combination of individual general impression and semiannual formal meetings between each resident and the Program Director. As noted, each resident then evaluated himself/herself in the same Milestones following a training didactic on how to use the NAS.

Given the intimate size of the residency program at our institution, each member of the CCC felt very familiar with the abilities of each of the residents. Our CCC includes a sufficiently diverse group of faculty from all rotation sites, enabling comprehensive and accurate evaluation of the residents in all of the Plastic Surgery subspecialty practice areas. The assessments described in this study were performed in November of the first year of NAS implementation and have been performed semiannually since. The various scores were then analyzed for mean, standard deviation, correlation, and significance across the 6 postgraduate years (PGYs), by individual Milestones and by core competencies. ANOVA was performed to compare self-evaluation and attendings' evaluations across PGYs. All tests were two-tailed, with a significance level of 0.05. All data