What Shape is Your Resident in? Using a Radar Plot to Guide a Milestone Clinical Competency Discussion

David T. Harrington, MD, Thomas J. Miner, Thomas Ng, MD, Kevin P. Charpentier, MD, Pam Richardson, and William G. Cioffi, MD

Alpert Medical School at Brown University, Rhode Island Hospital, Providence, Rhode Island

OBJECTIVE: One of the challenges for program directors (PDs) is to sort and weight the tidal wave of assessments that training programs create in the modern Milestone era. We evaluated whether the use of a radar plot (RP) would be helpful in sorting data and providing a graphic representation of each resident's progress.

DESIGN: Using at least 2 different types of assessments for each of the 16 surgical Milestones, the data were ranked and weighted by a predetermined method embedded in a computerized workbook (Excel). This process created a unique 16-spoked RP for each resident (Fig. below). The RP allowed the faculty to see areas of weakness (shown by concavity) and allowed an overall grade calculated as a ratio of the area of the smooth outer circle (faculty expectations, triangles) and the resident's unique radar shape (resident performance, squares). To help us validate our new tool, we looked at whether residents with recent remedial issues "looked" different from residents without remedial issues.

RESULTS: Of our 30 categorical residents, 8 had significant areas of concavities, suggesting possible areas of improvement. Of these 8 residents, 4 had been on a remediation program in the last 18 months. The average ratio of performance/expectations was 0.709. The 4 residents on recent remediation had a ratio of 0.616 when compared with 0.723 for the residents without remedial issues (p < 0.009).

CONCLUSIONS: Many exciting challenges await PDs, as we evolve to a competency-based evaluation system. The use of an evaluation summary tool using RPs may aid PDs in leading clinical competency discussions and in monitoring a resident's progress over time. (J Surg 72:e294-e298. © 2015 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: milestones, radar plot, assessments, competency

COMPETENCIES: Systems-Based Practice, Medical Knowledge, Practice-Based Learning and Improvement
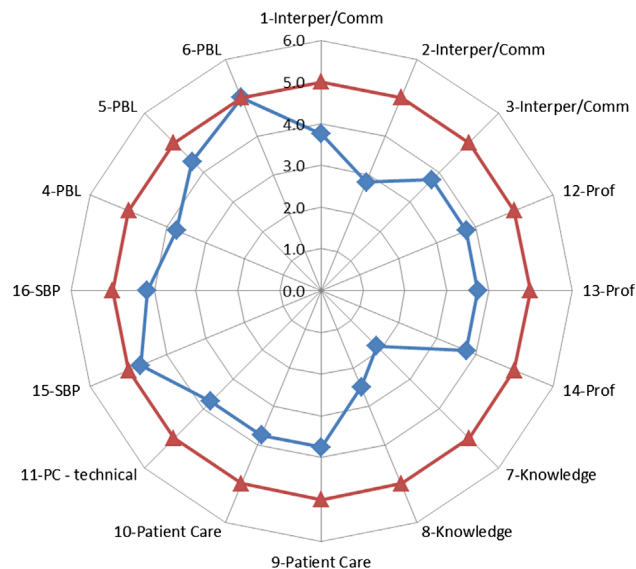
INTRODUCTION

The American Council on Graduate Medical Education (ACGME) Milestone projects is an ambitious goal to move residency training from a curriculum based on time to a curriculum based on attainment of competency. This project is the natural evolution from the ACGME 6 core competencies that asked program directors (PDs) to evaluate 6 aspects of surgical skills: knowledge, patient care, communication, professionalism, systems-based practice, and practice-based learning. PDs have been encouraged to develop a tool box to teach and assess for learning in all these competencies. To meet Milestone requirements, general surgical training programs have to teach, assess, and provide summative grading on 16 Milestones promulgated by the ACGME.

Many challenges await PDs as they become compliant with the Milestone project. First, there is the issue of sheer volume of evaluations to use in the evaluation process. American Board of Surgery In-Training Examination

(ABSITE) scores, 360° from the clinic staff, technical skill simulation data, Patient Assessment and Management Evaluation, Do Not Use abbreviation usage, prevented medication errors, and the ubiquitous end-of-rotation evaluation are just a few of the evaluations used. However, all evaluations are not of equal value, and it is difficult to sort the reliable data from the unreliable data. Now, the Clinical Competency Committee (CCC) and the PD also need to evaluate not just the 6 core competencies but also the 16 Milestones. With these challenges, it would be ideal to have a useful tool to process and display resident performance in a Milestone-friendly format.

The evolution of mathematics and statistics over the last 200 years also drove the development of visual displaying of data, as these statistical concepts needed a visual analog. Information visualization allows for simple and sometimes complex relationships to be displayed efficiently and powerfully.[1] Radar plots (RPs), sometimes referred to as star plots, spider graphs, or kiviat diagrams, allow multiple variables to be displayed in a single graph. These graphs allow for 3 or more quantitative variables to be displayed on separate equiangular axes that all start from the same point. Data points on each axis are plotted based on the maximum and minimum as defined on these axes or radii. Lines can be drawn between the individual data points on the separate radii to create a shape. The axis and these connecting lines give the graphs their "radar" or "spider-web" appearance.

Given the challenges of information overload, data of questionable reliability, the constraint of time, and the imperative of assessing the performance of our residents, we wondered whether we could use a RP to aid in our summary assessment of our resident's performance. Further, we wondered whether our CCC could use this RP to aid in assessing the competency of residents across the 16 Milestones.

METHODS AND MATERIALS

The Brown/Rhode Island Hospital Education Operations Committee reviewed all available assessments for each resident over the course of their 5 clinical years of their training. All assessments were evaluated for inclusion in our RP process. These included evaluations such as end-of-rotation evaluations, clinic 360° evaluations, NP/PA 360° evaluations, annual peer-to-peer evaluations, outcome project evaluation (a 2-y quality improvement project), morbidity and mortality evaluations, attendance at key educational conferences, number of administrative lapses (duty-hour logging and ACGME case log recording), operating room evaluations, technical skill laboratory performance, annual mock oral examination, and trauma morning report evaluations. For each postgraduate year (PGY) level, the available assessments were evaluated for their potential to fit into the 16 Milestones for surgery. The committee then chose the

assessments that they felt were the most reliable and reproducible. For instance, the committee chose ABSITE scores and the clinical judgment score from the mock oral examination rather than using the faculty evaluation of clinical knowledge on the end-of-rotation evaluations to assess the resident's knowledge for Milestones. The committee attempted to have at least 2 assessments for each of the 16 Milestones. The assessments were then fitted onto the individual axes through a process that ranked and weighted each assessment. The data were ranked to fit our expectation of performance. For instance, for assessments that were graded on a 5-point Likert scale, we determined the average score for this assessment. We have found that very few assessors use the whole 5-point scale and that the average score for individual questions is not 3 but between 4.0 and 4.2. For a question with an average score of 4.3 and a standard deviation of 0.6, we scored 5.0 to 4.75 as "exceeding performance expectations," 4.74 to 4.25 as "attaining performance expectations," 4.24 to 3.75 as "not meeting performance expectations," and <3.75 as "significantly below performance expectations" and assigned a score of 1.5, 1.0, 0.75, and 0.50 points for each of these categories, respectively. For the ABSITE scores, the committee used the ranges of >=80`, 79 to 50`, 49 to 30`, and <30` for those 4 categories of performance. Assessments that are PGY level specific are weighted to that PGY level—weight of 1 for a PGY1 performance expectation and a weight of 5 for a PGY5 performance expectation. However, some assessments might bridge several PGY levels and take more than 1 year to attain an expected level of performance. For these assessments, we had to weight the scores differently. For instance, PGY2 through PGY5 residents present at the weekly morbidity and mortality conferences. The residents are assessed at this conference on several areas, but they are all assessed on the same scale regardless of PGY level. Therefore, it is possible for a PGY2 resident to show "exceeding performance expectations" and obtain higher weighted performance rating. For example, a PGY2 resident would only have a maximal weighted score of 2 for passing the bead passing assessment of "attaining performance expectations" on our local Fundamentals of Laparoscopic Surgery simulation (level-specific competency), but he or she could attain 4 points for delivering a "attaining performance expectations" on a well-reasoned morbidity and mortality presentation (non–level-specific competency). By this process, each assessment was ranked and weighted through a mathematical formula embedded in a commercially available workbook (Excel). All scaled and weighted scores for each Milestone were averaged and then displayed on one of the RP Milestone axes

To make the RPs smoother, we grouped the Milestones such that related areas of performance were next to each other. For instance, Knowledge Milestones were plotted next to Patient Care Milestones and Communication Milestones were plotted next to Professionalism Milestones.