Evaluating of the Implementation of the ACGME Milestones in General Surgery

Angela Thelen, MD; Rachel Gottlieb, BS; Niki Matusko, MS; Gurjit Sandhu, PhD; Brian George, MD, MAEd; Daniel Kendrick, MD, MAEd

Background: Training program implementation of the ACGME Milestones in General Surgery has proceeded with minimal guidance, with individual programs given the complex task of developing an assessment system to inform the generation of discrete Milestone ratings by Clinical Competency Committees (CCC). As a result, little is known regarding what approaches different institutions have taken and how effective those approaches have been in measuring the domains of resident performance defined by the Milestones. This knowledge gap prevents the development of best-practice guidelines by which resident assessment can be standardized. In light of this, it is critical training programs develop methods to self-evaluate their Milestones implementation. We describe a robust approach to evaluate the implementation of the Milestones in general surgery at the program level.

Actions/Methods/Interventions: We conducted a formal program evaluation of the implementation of the general surgery Milestones at the University of Michigan using a logic model framework. The logic model was iteratively developed by the study team using the Kellogg framework (**Figure**). This approach was selected to systematically understand the inputs/activities making up their assessment program, clarify the outputs of each activity, and identify the outcomes expected as result of its implementation. Specific measures were determined to evaluate the achievement of outputs and outcomes of each activity including: (a) whether their assessment system was implemented as intended, (b) the CCC process of synthesizing the generated assessment data to determine Milestone ratings, and (c) the ability of the Milestone ratings to discriminate discrete areas of trainee performance.

Results: Through examination of assessment utilization data, we found that (a) each Milestone subcompetency mapped to at least one assessment tool and each tool was consistently used by faculty at the intended monthly intervals. In interviews with CCC members, we found that (b) the CCC process was carried out as intended with consistent use of assessment data, robust discussion of trainees, and collaborative decision making. Finally, using exploratory factor analysis we found that (c) Milestone ratings appropriately tracked overall trainee progression by post-graduate year. However, Milestone ratings were unable to specifically discriminate trainee performance beyond a single global construct of resident proficiency.

Lessons Learned: The logic model framework is a robust evaluative approach for examining the implementation of the ACGME Milestones in general surgery at the program level. Using this approach, we were able to identify specific areas of strength and areas for improvement in the assessment system and CCC process.

Future Application and Next Steps: Future efforts are needed to better understand how to build assessments that measure distinct domains of trainee competence. Ensuring all programs can accurately assess specific areas of resident performance is critical as we work toward a best practice model within competency-based medical education.

Figure: Logic Model for ACGME Milestones Implementation in General Surgery at the University of Michigan