# The General Surgery Milestones Project

Thomas H Cogbill, MD, FACS, Mark A Malangoni, MD, FACS, John R Potts, MD, FACS, R James Valentine, MD, FACS

## Traditional methods of evaluation

Evaluation is an essential component of any educational system. Evaluation consists of a systematic determination of merit, worth, and significance, using criteria governed by a set of standards. Inherent in any evaluation process is the judgment of a responsible individual or governing body about which criteria and standards should be used.

There are 2 types of evaluations commonly used in graduate medical education: objective and subjective. Objective evaluation is quantitative and often takes the form of a written test, most commonly involving a single correct answer to a multiple choice question. Examples during surgery residency include the American Board of Surgery In-Training Examination (ABSITE) and quizzes for either self-assessment or comparative purposes. Objective evaluations are often compared to a norm such as percent of answers correct or percentile ranking within a group. Subjective evaluation is commonly applied when a variety of observed parameters are used to develop a qualitative judgment. It is more challenging to compare subjective evaluations because the norm is often "in the mind of the evaluator" and may not be well defined. It is possible to compare subjective evaluations between individuals and groups; however, there is a tendency for most subjective evaluations to cluster around the mean, particularly as time between the observation and formal evaluation elapses. Both objective and subjective forms of evaluation can be used as formative or summative assessments of performance.

Most evaluations performed during surgery residency are subjective, tend to be completed at the end of a rotation that can vary from 2 weeks to 3 months, and often consist of a solitary number on a Likert scale or a single brief descriptive phrase. These evaluations are rarely comprehensive or consistent enough to allow a resident or program to chart a trajectory of performance that may be compared with other learners. Furthermore, in the absence of standard evaluation forms used by all surgery residencies, national normative data have been lacking.

## The next accreditation system

The Accreditation Council for Graduate Medical Education (ACGME) is responsible for accreditation of allopathic postgraduate medical training programs in the United States. Accreditation is accomplished through a peer-review process and is based on established standards and guidelines.[1] Historically, the ACGME has evaluated programs by periodically assessing program structure and process measures through a program information form (PIF) completed by the program director and an on-site review by an external site visitor to verify the content of the PIF. The site visitor's report and the PIF are peer-reviewed by the appropriate Residency Review Committee (RRC). The RRC then renders a decision for accreditation of the program for a 1- to 5-year period or for an adverse action (proposed probation, proposed withdrawal of accreditation, or proposed reduction in resident complement). The ACGME refined this process in 1999 with the introduction of the 6 clinical competencies.

This prescriptive process has had its shortcomings. Program directors and designated institutional officials may focus on meeting program standards rather than assuring that residents are adequately prepared for unsupervised clinical practice at the conclusion of training. There has also been recognition that the competencies may be better assessed in an integrative manner rather than as individual parameters.[2]

The ACGME began implementation of the Next Accreditation System (NAS) for 7 core specialties in July 2013 and it is planned to be fully implemented by all specialties in July 2014. A key feature of the NAS is the change from periodic assessment of programs to annual collection and assessment of educational outcomes data from programs by the relevant RRC.[3,4] Central to this change has been the development of educational milestones — specific developmental steps achieved at designated time points during the continuum of graduate medical education — for resident evaluation in each specialty. These milestones have been developed by each specialty with representation of the specialty board, RRC, program directors' organization, membership organizations, and at least 1 resident

Disclosure information: Nothing to disclose.

Received February 6, 2014; Accepted February 10, 2014.

From the Department of General and Vascular Surgery, Gundersen Health System, La Crosse, WI (Cogbill); the American Board of Surgery, Philadelphia, PA (Malangoni); the Accreditation Council for Graduate Medical Education, Chicago, IL (Potts); and the Division of Vascular Surgery, University of Texas Southwestern Medical Center, Dallas, TX (Valentine). Correspondence address: Thomas H Cogbill, MD, FACS, Department of General and Vascular Surgery, Gundersen Health System, 1900 South Ave C05-001, La Crosse, WI 54601. email: thcogbil@gundersenhealth.org





ISSN 1072-7515/14/$36.00
http://dx.doi.org/10.1016/j.jamcollsurg.2014.02.016

**Abbreviations and Acronyms**

ABS = American Board of Surgery
ACGME = Accreditation Council for Graduate Medical Education
APDS = Association of Program Directors in Surgery
CCC = clinical competency committee
NAS = Next Accreditation System
PIF = program information form
RRC = Residency Review Committee

from the specialty along with the support of the ACGME. When sufficient milestone data have accrued (which will likely require several years), milestones' progression in individual programs will be compared with national normative data and used in accreditation decisions. With the NAS, it is expected that site visits will be less frequent and directed to specific educational deficiencies in programs. The NAS allows programs that are performing well to be more innovative in their approaches to resident education.

## Advantages of the milestones

Changing to a system in which the achievement of educational milestones by individual residents is a critical outcomes parameter for program evaluation is an important step in moving to a system that emphasizes demonstrable indicators of readiness for unsupervised practice. This shifts the emphasis of program evaluation from process measures to outcomes that are important for the public, trainees, and educators. The milestones will also serve to segue into a professional learning environment as outlined by maintenance of certification requirements determined by each of the specialty boards.

Rather than relying solely on the opinion of the program director, the process of semiannual evaluation based on educational milestones involves broad input from multiple evaluators (faculty, residents, students, patients, families, nurses, and others) who have observed the clinical performance of the resident, which is then synthesized by a clinical competency committee (CCC). Use of the same comprehensive evaluation methods twice annually gives both the resident and the program a clear view of the resident's trajectory of performance across all competencies. The use of a single formative and summative evaluation process by all surgery residencies allows for compilation of aggregate information. Development of specialty-specific milestones provides a national framework for assessment that is data driven. The milestones define the expectations of the profession and integrate both objective and subjective evaluation tools. Although the milestones will certainly not solve all of the difficulties associated with education of residents, they are an important component of an improved educational system that begins with residency and continues into life-long professional practice.

## Milestones development

General surgery was invited to participate in the initial phase of the ACGME's NAS. The General Surgery Milestones Project was planned as a joint initiative between the ACGME and the American Board of Surgery (ABS), and representatives were chosen from both organizations to form a Working Group. The Association of Program Directors in Surgery (APDS) was also well represented; 12 members in the Working Group were current or past general surgery residency program directors. One surgical resident was included in the Working Group. Because there was no precedent for milestones in the initial phase, the Working Group was given tremendous latitude for developing specific educational targets for surgery residents and establishing the measurement intervals during training. This proved to be both a blessing and a curse because there were a number of false starts during the first 2 years of the project. However, every milestone version proved to be a learning opportunity. The final general surgery milestones document incorporates many features identified in the previous efforts.

Five face-to-face meetings, 3 conference calls, and numerous on-line surveys were required for this project. The Working Group was first convened in 2009 under the leadership of Richard H Bell Jr, MD. In 2011, 5 new members were added, including 2 each from the RRC-Surgery and APDS, as well as a new chair, Thomas H Cogbill, MD. From its inception, the Working Group members agreed on 3 overarching concepts. First, the milestones should be simple and based on measurable attributes. The group acknowledged that a new evaluation process would represent increased work for program directors, and one of the main goals was to minimize that burden. Second, existing assessment tools should be used whenever possible. The group strongly believed that efforts should be directed at constructing milestones rather than developing and testing a catalogue of new evaluation schemes. However, group members recognized the lack of a standardized tool for evaluation of technical skills, and development of such a tool became a priority. The third overarching concept was that measured attributes should be worthwhile. The milestones should clearly distinguish top performers from poor performers, and substandard performance should be recognized early enough in residency to allow meaningful intervention. On the other hand, milestones should be able to provide hard data to support the decision of a program director to