

Can We Achieve Public Accountability for Graduate Medical Education Outcomes?

Robert B. Baron, MD, MS

Abstract

Calls for greater public accountability for graduate medical education (GME) outcomes continue to come from a broad array of stakeholders. Creation of ways to measure accountability requires a clear understanding of the domains of GME outcomes and the creation of specific measures that are reliable and accurate and do not create an undue measurement burden. Three domains of outcomes are necessary: individual trainee competence,

the quality and diversity of the training environment, and workforce factors that address workforce size, specialty mix, diversity, and geographic distribution. The Accreditation Council for Graduate Medical Education has begun to develop measures that have the potential to form the basis of the first two domains, and other data sources exist to measure the quality of the training environment. Little progress, however, has been made to

accurately describe institution-specific workforce outcomes. The article by Chen and colleagues in this issue makes a major contribution in the measurement of institution-specific outcomes. Their article creates optimism that a system that incentivizes and rewards specific desirable GME outcomes can be designed. This commentary further defines some practical next steps to achieve this desired GME accountability.

Editor's Note: This is a commentary on Chen C, Petterson S, Phillips RL, Mullan F, Bazemore A, O'Donnell SD. Toward graduate medical education (GME) accountability: Measuring the outcomes of GME institutions. Acad Med. 2013;88:1267–1280.

As discussed by Chen and colleagues¹ in this issue, the calls for public accountability for graduate medical education (GME) outcomes have been long-standing and have come from a broad array of stakeholders, including educators, accreditors, policy experts, physician groups, legislators and legislative advisors, the Obama administration, and consumer advocates. Arguably, the recommendations with the greatest impact were those published by the Medicare Payment Advisory Commission (MedPAC) in June 2010.² MedPAC recommended that GME should be changed to “support the workforce skills needed in a delivery system that reduces cost growth while maintaining or improving quality.”² Most important, MedPAC recommended that

Medicare funding to GME institutions be linked to performance on these new standards and that a substantial portion of Medicare's payments be used to fund these new performance standards.

In the years since, calls for greater GME accountability have continued. The process, though, has been stymied not only by a lack of consensus but also by the many challenges of defining the outcomes that should be used to measure accountability, defining the measurement strategies required to ensure a valid and reliable system, and creating measures that are integrated with other existing outcomes so as not to create an undue measurement burden. Identifying a way forward begins with answering a number of essential questions.

What Outcomes Should Be Used to Measure GME Accountability?

I suggest dividing GME accountability outcomes into three specific domains. First is individual trainee competence. Each future physician completing GME at a U.S. training program should possess the specific competencies needed to meet the needs of individual patients and the needs of the public-at-large. Second, residents and fellows must be trained in diverse clinical settings that can demonstrate safe, high-quality, high-value, patient-centered health care. Third, GME programs must produce a physician workforce of the appropriate size, specialty mix, diversity, and geographic distribution to meet the needs of the public.

How Can We Best Measure Individual Trainee Competence?

The Accreditation Council for Graduate Medical Education (ACGME) has made the assessment of enhanced individual trainee competence a centerpiece of its Next Accreditation System (NAS).³ The ACGME, working closely with certifying boards, residency review committees, specialty organizations, program directors, and trainees, is developing definable milestones to document the professional development of each trainee. Milestones will provide a framework for formative feedback for trainees during training and ultimately for summative assessment to determine the ability of residents and fellows to practice independently upon completion of training. Aggregated milestones will also be one factor used by the ACGME NAS to determine the accreditation status of each GME program. These aggregated milestones, especially those achieved by the end of training, may also have potential utility as a measurement of an institution's ability to produce physicians who have the desired skills to meet the needs of the public.

Unfortunately, this system is in its infancy and has not yet proven itself to be accurate or reliable. Nonetheless, the milestone project offers the best opportunity defined to date to measure trainee competence. Alternative measures that rely exclusively on GME program reports of what was taught during residency and fellowship or on reports

Dr. Baron is professor of medicine and associate dean for graduate and continuing medical education, Division of General Internal Medicine, Department of Medicine, University of California, San Francisco, School of Medicine, San Francisco, California.

Correspondence should be addressed to Dr. Baron, Office of Graduate Medical Education, 500 Parnassus Ave., Suite MUE 250, San Francisco, CA 94143-0474; e-mail: baron@medicine.ucsf.edu.

Acad Med. 2013;88:1199–1201.
First published online July 24, 2013
doi: 10.1097/ACM.0b013e31829ed2ed

from trainees and faculty about what was taught would be substantially less precise at addressing the public's right to know that each trainee has developed the appropriate level of competence to practice independently and effectively in the health care system of the 21st century.

How Can We Best Measure the Quality of the GME Training Environment?

A growing body of evidence suggests that the quality and safety of the clinical setting in which future physicians are trained reflect the quality of their future independent practice.⁴ Hospitals and other clinical settings currently publicly report hundreds of measures that demonstrate the quality and safety of their institutions, and each teaching hospital must meet standards determined by the Joint Commission and the Centers for Medicare and Medicaid Services. New programs in value-based purchasing further require meeting specific outcomes to receive maximum payments. The ACGME is also attempting to address this domain, in part, through its new process of Clinical Learning Environment Review (CLER).⁵ The CLER program is specifically designed to measure resident and fellow engagement in each institution's quality and safety programs and may serve as a model to ensure that teaching settings are able to prepare residents and fellows to practice in safe, high-quality, cost-effective environments.

The CLER program, though, is also in its infancy and is designed as a formative process to stimulate institutional improvement. Using CLER as part of a GME accountability process with financial implications for training institutions would dramatically change its intended use but could potentially serve as one effective method to measure the quality of a training site. Multiple other measures of quality and safety of both inpatient and outpatient settings could also be used to demonstrate the safety and quality of training settings. Inpatient examples include meaningful-use criteria,⁶ hospital mortality and morbidity measures, the use of safety measures such as surgical checklists, the avoidance of Medicare "never events," and quality measures in value-based purchasing programs. Similarly, the quality and safety of outpatient training settings could be assessed by measures

of faculty participation and outcomes in the Physician Quality Reporting System, other outpatient quality measures, patient experience measures (e.g., the Clinician and Group Consumer Assessment of Healthcare Providers and Systems), the use of electronic health records, and the achievement of meaningful use.

How Do We Create Accountability for the Kinds of Physicians Trained?

The most challenging domain of GME accountability is the creation of a physician workforce of the appropriate size, specialty mix, diversity, and geographic distribution to meet the needs of the public. Unfortunately, there is little consensus on the number and kinds of physicians needed for the health care system of the 21st century. Efforts to create a National Health Care Workforce Commission to provide such data have been stymied in the U.S. Congress. Moreover, national needs may not reflect regional and local workforce needs, and these factors also deserve immediate attention. Nonetheless, existing data suggest national physician shortages in several specialty areas including primary care, general surgery, psychiatry, and other medical specialties.

Importantly, the ACGME has little control over these issues. In fact, most decisions about the kinds of residents and fellows trained in a given institution are made by institutional leadership themselves. Although the ACGME Residency Review Committees affect specialty numbers by approving or denying requests for expansion of programs within specialties, these decisions are made mostly on the ability of a given training program to meet ACGME requirements, with little attention to the impact on the nation's physician workforce. In fact, no centralized entity has significant influence over the overall specialty mix of physicians. Rather, a relatively small number of large teaching hospitals that train the majority of residents and fellows in the United States make most of these decisions.

Despite these issues, it is essential that we develop sophisticated measures that define and assess the workforce outcomes of our training institutions. The article in this issue by Chen and colleagues makes major contributions to this task. The authors use an extremely detailed analysis

of data from the American Medical Association Physician Masterfile and its GME supplement, the National Provider Identifier database, Medicare claims, and the National Health Service Corps to define the number and percentage of GME graduates of U.S. training institutions that are practicing in various specialties and in underserved areas. For example, the authors found 25.2% of GME graduates practicing in primary care specialties but with dramatic variations between institutions. While 20.8% of GME sponsoring institutions produced no primary care graduates, 24.2% of institutions produced over 80%. Similarly, 26.1% of sponsoring institutions produced no rural physicians. Although the authors focused primarily on graduates practicing in primary care specialties and general surgery and on practice in rural and other underserved areas, their approach could also be used to measure additional workforce-related training outcomes, including workforce diversity, of sponsoring institutions and training sites. This analysis should be seen as a major advance in our ability to measure institution-specific GME training outcomes.

Other workforce outcomes may also be desirable and accurately measured. These might include the extent to which institutions integrate physician training with training of other health professionals, including nurse practitioners and physician assistants, or the number of graduates selecting research or teaching careers. Measurements of institutional culture and attitudes towards needed specialties (such as primary care) may also be a useful interim outcome.

Next Steps to Create a GME Accountability Program

Clearly, accurate measures of GME training outcomes in all three domains are a long way from complete. Substantial additional research, like that of Chen and colleagues, needs to be done, and pilot programs need to be initiated. Many will likely argue that moving forward with financial incentives based on GME outcomes would be premature until we have substantially better data. I would argue instead that the creation of incentives now will accelerate the kind of data collection, analysis, and pilot programs that are

needed to make the process better. For example, an incentive program could begin with a modest amount of Medicare payments at stake, similar to other value-based purchasing initiatives. Although MedPAC has suggested that all of the indirect Medicare payments above the “empirically justified amount” should be placed at risk, a more modest and graduated approach would create fewer unintended consequences and allow for more time to develop reproducible standards and measures.

Similarly, a process of tying GME payments to training outcomes could begin gradually by rewarding the collection and reporting of outcome data rather than allocating rewards based on the outcomes themselves. Similarly, reporting of surrogate measures and process measures while outcome measures are being developed is likely to be useful in the early stages of such a program.

Despite the significant challenges they face, it is time for the GME community, particularly the teaching hospitals that receive the majority of public funding and train the majority of residents and fellows, to accept that creating a more accountable GME payment system has merit. The ACGME has taken important first steps in creation of potential outcomes to measure trainee competence and the quality of training environments, and Chen and colleagues have made a most important contribution to our ability to measure institution-specific physician workforce outcomes. We must continue to work toward answers to the questions at the heart of GME accountability to ensure that our GME system meets the country’s health care needs.

Funding/Support: None.
Other disclosures: None.
Ethical approval: Not applicable.

References

- 1 Chen C, Petterson S, Phillips RL, Mullan F, Bazemore A, O’Donnell SD. Toward graduate medical education (GME) accountability: Measuring the outcomes of GME institutions. *Acad Med.* 2013;88:1267–1280.
- 2 Medicare Payment Advisory Commission. Report to the Congress: Aligning Incentives in Medicare. http://medpac.gov/documents/Jun10_EntireReport.pdf. Accessed May 10, 2013.
- 3 Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—Rationale and benefits. *N Engl J Med.* 2012;366:1051–1056.
- 4 Asch DA, Nicholson S, Srinivas S, Herrin J, Epstein AJ. Evaluating obstetrical residency programs using patient outcomes. *JAMA.* 2009;302:1277–1283.
- 5 Weiss KB, Bagian JP, Nasca TJ. The clinical learning environment: The foundation of graduate medical education. *JAMA.* 2013;309:1687–1688.
- 6 Centers for Medicare and Medicaid Services. Meaningful use. http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/Meaningful_Use.html. Accessed May 18, 2013.

Teaching and Learning Moments
Words of Comfort

They were given to me
in my own time of need

during my intern year
amid fear and self-doubt.

*You may not believe me
but things will get better.*

He was right. I didn’t
believe him, which is why

part of me believed him.

And things did get better.

Now, several years later
I am an attending.

When I have a learner
or a patient who is

in need of hope, I share
these same words of comfort

as long as I believe
they really could be true.

Author’s Note: In this poem, I describe the difficulty I faced as I transitioned from a fourth-year medical student to an intern and the words of comfort my program director at the time shared. I continue to carry these words with me and to pass them on to others in need along the way.

Adam Possner, MD

Dr. Possner is assistant professor, Department of General Internal Medicine, George Washington University School of Medicine and Health Sciences, Washington, DC; e-mail: apossner@mfa.gwu.edu.