

Grading Residents' Clinical Performance: Unique Opportunities and Challenges

Joshua M. Liao, MD¹ and Robert M. Wachter, MD²

¹Brigham & Women's Hospital / Harvard Medical School, Boston, MA, USA; ²Department of Medicine, University of California, San Francisco, San Francisco, CA, USA.

J Gen Intern Med 29(7):983–5
DOI: 10.1007/s11606-014-2856-3
© Society of General Internal Medicine 2014

BACKGROUND

The American healthcare system is being criticized for inconsistent quality, high rates of harm, and backbreaking costs. While tremendous attention has been placed on practicing physicians and healthcare organizations, there is increasing interest in how graduate medical education (GME) contributes to these problems.¹ Seen in this light, the “product” of GME should be physicians that produce high-value care, and today's GME must adapt to meet this emerging objective.

In response to calls for reform, national organizations have identified “practice-based learning and improvement,” “systems-based practice,” and, more recently, “high value, cost-conscious care” as competencies that trainees should master.² Residency programs are responding with new safety, quality, and cost curricula, and growing numbers of trainees report interests in these areas.

To promote these changes, programs must create reliable ways to help trainees and teams capture and process information about their performance in quality, safety, patient experience, and value. Compared to practicing physicians, housestaff frequently lack information about quality and outcomes that they can use to consistently improve their clinical practice. Moreover, most do not learn the essential competency of reviewing and acting on performance-related data during their training.

In this article, we describe some unique challenges associated with providing resident physicians systematic feedback about clinical performance. We also review key elements of successful implementation and ways to overcome these challenges before discussing important tensions to consider going forward.

UNIQUE IMPLEMENTATION CHALLENGES

It is often difficult to link resident behaviors to enough discrete patient outcomes to produce useful conclusions and trends in performance. For example, in contrast to faculty or community-based practicing physicians (whose work patterns tend to be stable and directly linked to discrete patient encounters), housestaff deliver much more fragmented care. In inpatient settings, increased shift work and hand-offs make it difficult to link patient outcomes to specific residents. It is also challenging to track individual performance in ambulatory settings, whether organizational models consist of traditional half-day sessions or newer team-based approaches that spread responsibility among multiple providers.

In addition, housestaff also care for very heterogeneous patient populations, further complicating efforts to glean conclusions about performance from their data. Unlike their attending counterparts, residents often rotate through multiple clinical sites that serve different demographics, possess very different workflows, and do not share medical records. Just consider the case of an internal medicine resident: data from multiple general medicine, sub-specialty, intensive care, and consultation services would need to be extracted, likely from different clinical delivery systems, and synthesized in order to provide a comprehensive understanding of his or her care patterns and competencies.

ELEMENTS OF SUCCESSFUL PERFORMANCE MEASUREMENT SYSTEMS FOR TRAINEES

The first step towards creating useful resident clinical scorecards involves establishing effective health information and/or data collecting systems. Several resident-specific considerations can help in these efforts.

Effective feedback systems must involve active resident participation. For example, depending on rotation, institutional culture, and degree of autonomy, metrics used to evaluate attending physicians may inaccurately reflect resident decision-making. Conversely, there may be specific or completely new metrics that are highly relevant to training environments. Because examples of such feedback

Electronic supplementary material The online version of this article (doi:10.1007/s11606-014-2856-3) contains supplementary material, which is available to authorized users.

Published online April 5, 2014

systems are currently lacking, educators may wish to begin by collaborating with residents to combine current GME evaluation strategies with existing physician metrics to produce new resident evaluation metrics (appendix table available online).

For example, existing GME peer evaluations can be modified to evaluate the quality of resident hand-offs, a crucial component of trainee workflow. Relevant inpatient quality metrics, like timely discharges, patient satisfaction, hand hygiene and documentation of fall risk or thromboembolism prophylaxis, could be also be utilized. Conversely, measures that seem more highly dependent on attending oversight, such as length of stay and readmissions, might be excluded. This approach can ensure that resident performance data capitalizes on existing GME measures, is relevant to trainee performance, and mirrors the type of performance metrics that residents will experience after they complete training. It also has the advantage of teaching trainees about the critical process of measure development and assessment science.

Performance data must also be anchored in transparency. Residents should know what information will be captured, and programs must communicate how it will be used. Some may use data as part of evaluation and recommendation processes for all residents. Others may employ it selectively based on training level, perhaps only for feedback among interns and junior residents, but for evaluation and recommendation among senior residents. Whatever the decision, it should be understood clearly by all trainees.

Resident data should also embrace, rather than avoid, the realities of team-based medicine. There will likely be clinical sites or rotations for which it is difficult to extract individual-level information. However, with team-based approaches changing care delivery and housestaff positioned within large clinical teams, the use of team-level data may sometimes prove preferable, both because it reflects the reality of care delivery and signals organizational priorities regarding effective teams. New analytic methods are likely required to both characterize team performance and to tease out individual contributions to high-functioning teams.

Importantly, collected clinical data must be balanced. That is, metrics should reflect the wide range of activities of housestaff and the diverse competencies we hope that they will develop. A scorecard that emphasizes adherence to evidence-based and/or cost-effective practices but fails to capture professionalism, teaching skills, or diagnostic acumen will inevitably shift trainees' and programs' focus toward the former competencies and away from the latter. Of course, given that some things are easier to measure than others, it will be crucial that the balance is determined largely by the importance of the domains rather than the ease of

the measures, lest we find ourselves "hitting the targets but missing the point." The movement toward competency-based assessment will help frame this effort.³

Finally, this work will require broad leadership support. Busy housestaff need enough time, teaching, and mentorship to make full use of captured information. This will require more than just investment in residents; it will require trained faculty who can help translate data into actionable feedback and improvement plans, as well as infrastructures that can capture and present the information meaningfully. It will be interesting to see whether hospital and GME leaders build such infrastructure using personnel and systems that already do similar work for other medical staff, or through new, housestaff-oriented systems.

IMPORTANT TENSIONS

As programs begin the work of gaining resident buy-in, balancing individual vs. team performance, using data for individual learning vs. clinical assessment, and more, several philosophical tensions will likely arise.

The first relates to whether residents should be evaluated "on a curve." For example, it is not entirely clear when and how to emphasize parsimony during trainees' early professional development,⁴ or how to balance the value of comprehensive evaluation for atypical, rare conditions with that of cost-effective evaluation and resource stewardship. While educators may be tolerant of learning curves related to cost or quality measures, patients and payers may be less forgiving.

Second, programs must decide how data will be used to enhance performance. Is individual feedback sufficient, or should limited versions of "public" information (i.e., comparative data posted within the training program or the institution itself) be displayed? Questions also exist around using financial incentives to motivate housestaff performance. Of course, trainees will experience all of these methods when they enter practice, but some argue that similar mechanisms during GME contaminate the educational experience and compromise professionalism. On the other hand, programs such as those at the University of California, San Francisco have begun carefully using performance-based financial incentives for trainees, and early results are encouraging.⁵ Notably, they have intentionally utilized program-level metrics, rather than individual ones, to determine compensation. As housestaff scorecards are generated, this tension must be directly studied and addressed.

A final tension relates to the influence of senior physicians on resident performance. Many patient care decisions in teaching settings are made with input from

attendings, the physicians of record and most experienced clinicians. It is unclear how measurement systems should account for this element of resident performance, but successful ones must do so while appreciating the increasing responsibility and autonomy residents gain as they progress through training.

As programs begin capturing and utilizing resident performance data, these important tensions should guide ongoing research efforts. Educators must describe and quantify the barriers and perspectives related to performance feedback, both to maximize buy-in from residents and attending physicians alike, and to inform future educational policy. Comparison of the different types of approaches outlined above, along with rigorous efficacy evaluations, will also be crucial.

CONCLUSION

The need to generate clinical performance feedback for housestaff is clear, both because of the imperative to train housestaff in the competency of data-driven performance improvement, and because academic medical centers will be held to increasingly rigorous standards of high-value care. Much work is needed in this area. Ultimately, careful collaboration between educators, system administrators, policymakers, and housestaff can produce reporting systems that achieve a new kind of triple aim: increasing the quality of GME, preparing trainees for future practice, and capturing meaningful outcomes that institutions can utilize to improve their clinical performance.

Conflict of Interest: Dr. Liao reports no conflicts. Dr. Wachter reports serving as the immediate past-chair of the American Board of Internal Medicine (for which he received a stipend) and a current member of the ABIM Foundation board; receiving a contract to UCSF from the Agency for Healthcare Research and Quality for editing two patient-safety websites; receiving compensation from John Wiley and Sons for writing a blog; receiving compensation from QuantiaMD for editing and presenting patient safety educational modules; receiving royalties from Lippincott Williams & Wilkins and McGraw-Hill for writing/editing several books; receiving a stipend and stock/options for serving on the Board of Directors of IPC: The Hospitalist Company; serving on the scientific advisory boards for PatientSafe Solutions, CRISI, QPID, SmartDose and EarlySense (for which he receives stock options); and holding the Benioff endowed chair in hospital medicine from Marc and Lynne Benioff. He is also a member of the Board of Directors of Salem Hospital, Salem OR, for which he receives travel reimbursement but no compensation.

Corresponding Author: Robert M. Wachter, MD; Department of Medicine, University of California, San Francisco, Room M-994, 505 Parnassus Avenue, San Francisco, CA 94143-0120, USA (e-mail: bobw@medicine.ucsf.edu).

REFERENCES

1. **Lucey CR.** Medical education: part of the problem and part of the solution. *JAMA Intern Med.* 2013;173:1639-43.
2. **Weinberger SE.** Providing high-value, cost-conscious care: a critical seventh general competency for physicians. *Ann Intern Med.* 2011;155:386-88.
3. **Nasca TJ, Philibert I, Brigham T, Flynn TC.** The Next GME Accreditation System—rationale and benefits. *N Engl J Med.* 2011;366:1051-6.
4. **Moriates C, Soni K, Lai A, Ranji S.** The value in the evidence: teaching residents to “choose wisely”. *JAMA Intern Med.* 2013;173:308-10.
5. **Bischoff K, Goel A, Hollander H, Ranji SR, Mourad M.** The Housestaff Incentive Program: improving the timeliness and quality of discharge summaries by engaging residents in quality improvement. *BMJ Qual Saf.* 2013;22:768-74.