

Questions Program Directors Need to Answer Before Using Resident Clinical Performance Data

Sarah Gebauer, MD
Elizabeth Steele, MD

Twice a year, training programs must report milestones for every resident to the Accreditation Council for Graduate Medical Education (ACGME). The ACGME lists possible resident progress assessment methods to inform the milestones, but many are subjective. In addition, the ACGME surveys residents to verify that programs give trainees feedback on their performance, as well as their personal clinical effectiveness. In an effort to make feedback in the latter dimension reliable and meaningful, program directors are searching for and devising systems to give objective, unbiased clinical performance data. The ability to gather and report process and outcome data via automated systems (eg, electronic health records, registries, and billing data) in medical practice is relatively new, and educators should be aware of the complexities.

Obtaining structured, objective clinical performance feedback data can be a challenge. Some groups provide automatic feedback of clinical performance data on measures like proper antibiotic administration and incidence of complications. Unfortunately, the authors of 1 study were unable to find a correlation between the level of training and the performance on these metrics, or any longitudinal improvement in the metrics for a given resident over time.¹

As departments collect data for quality and milestone reporting, this should allow them to parse the data to the level of individual residents. The temptation to use these data to “get some numbers,” to meaningfully fulfill the feedback requirement, may become significant. This secondary use of patient data from electronic health records, billing, and other sources to understand individual provider performance is still in its infancy, and data can easily be misinterpreted and misused. Accuracy and transparency must be considered before providing residents with data gathered for other purposes, and particularly before using it for competency determinations.

When devising policies to use data gathered for other purposes to evaluate resident clinical performance, program directors should be prepared to answer the following questions.

1. How can we be sure that these data reflect a specific resident's patients?

Understanding the process of attribution, or the way a patient is assigned to a provider, is crucial. Many patients are seen by multiple providers in both inpatient and outpatient settings; thus, deciding which patients are attributed to a resident can be challenging. One study of primary care residents used data from patients, with a minimum number of visits over a defined time frame, for tracking the use of preventative measures.^{2,3} Other methods may be more appropriate for physicians with a largely inpatient practice, such as inpatient consultants and proceduralists.

For example, in anesthesiology, residents could be attributed (TABLE) to any patient they cared for, or only to patients for whom their care constituted the majority of anesthesia time. For some metrics, such as postoperative pain score, it may make sense to credit only the last resident (ie, the anesthesiology resident who took the patient to the recovery room). In inpatient clinical services, many resident metrics will track closely with attending physicians' performance. This issue has been addressed by some authors pointing out the team- or system-related differences in any physician's practice, and suggesting that resident performance metrics should reflect the team nature of contemporary medical practice.⁴

In any setting, knowledge of the local system and which factors are under the residents' control is crucial; for example, vaccination rates may be completely driven by nurse protocols.

2. Were the patients of the given resident just sicker than the other residents' patients?

Risk adjustment is an attempt to avoid unfairly penalizing residents caring for higher-risk patients. Adequate and transparent risk adjustment is often a difficult hurdle for departments and even for institutions. To the extent available, omnibus comorbidity indices such as EuroSCORE⁵ may offer convenient and reliable risk adjustment, but they require a large body of underlying data. Otherwise,

TABLE

Definitions of Key Words Related to Data and Quality

Term	Definition	Example
Attribution	Which patients “belong” to an individual clinician for data purposes	A patient who has seen a particular resident for more than half of his or her primary care visits would be attributed to that provider
Process measure	Measurement of an intervention by a member of the health care team	Referral of a diabetic patient for nutritional counseling
Outcome measure	Measurement of an aspect of the patient’s clinical status	Hemoglobin A1c less than 6%

program directors may need to consult with data experts, clinic and hospital staff, and operations personnel to devise fair and reliable risk adjustment. Although accurate, a complicated method of risk adjustment may lead some residents to distrust the data if they are not confident that they understand how the adjustment was calculated.⁶

3. Was the given resident’s ranking low due to the impact of a single patient?

Residents see a wide variety of patients; thus, some performance metrics will be based on a small sample size. For example, a family medicine resident may provide care mostly for adults and see only a few children per week. A monthly evaluation of pediatric immunization rate for that physician would be more skewed by a single unvaccinated patient than it would be for a pediatrics resident who cares solely for children. Frustratingly, the minimum useful sample size (ie, resistant to random fluctuations) is highly context dependent. One study³ found that published recommendations vary from as few as 11 to as many as 45 patients, and the authors suggested that minimum sample size should be based on an analysis of how increasing patient numbers changes confidence intervals for outcome frequency.

4. Does this imply that the given resident has clinical effectiveness problems?

Even with excellent accuracy, specific performance measures may not provide a full view of a provider’s performance. One study noted that three-quarters of physicians were highly ranked in at least 1 measure, and three-quarters of those same physicians performed poorly in at least 1 other.⁷ This finding emphasizes the importance of using clinical performance metrics in the context of a broader and multifactorial evaluation of a provider. Many proposed tools subjectively evaluate clinical performance, which may introduce bias and have poor correlation with clinical outcomes.⁸ Programs should also be thoughtful about data presentation, with

institutional norms determining whether residents receive their score in relation to a qualifying threshold, median score, quartile rank, or numbered rank. Providing residents with a ranking relative to their peers, who are working within the same set of constraints, may provide a starting place for an important formative discussion about the resident’s clinical practice.² These may include 360° evaluations as suggested by the ACGME.

5. Who can see the given resident’s data?

Residents may have concerns about whether their data can be released to fellowship programs, employers, or other organizations. Programs should have policies limiting release of individual resident clinical performance data to specified parties, and possibly be limited to faculty who are both heavily involved in the residency program’s direction and able to explain the limitations and meaning of these data.

The release of physician performance data has already caught the attention of state legislators. Colorado passed a Physician Designation Disclosure Act in 2008, which was written in response to insurance companies’ physician rankings and “addresses four key issues: data integrity; disclosure; fair process; and enforcement.”⁹ The issues regarding data use and interpretation will follow resident physicians throughout their careers. Residents will continue to receive clinical performance data in an increasingly data-driven medical environment. Educators and program directors need to be leaders in demonstrating the importance of understanding the possibilities and limitations of these data. As data collection and analysis processes mature, we are hopeful that many of these issues will be solved in ways that are reliable, fair, and valid as we all work to improve resident education.

References

- Ehrenfeld JM, McEvoy MD, Furman WR, et al. Automated near-real-time clinical performance feedback for anesthesiology residents: one piece of the milestones puzzle. *Anesthesiology*. 2014;120(1):172–184.

2. Houston TK, Wall T, Allison JJ, et al. Implementing achievable benchmarks in preventive health: a controlled trial in residency education. *Acad Med*. 2006;81(7):608–616.
3. Higgins A, Zeddies T, Pearson SD. Measuring the performance of individual physicians by collecting data from multiple health plans: the results of a two-state test. *Health Aff (Millwood)*. 2011;30(4):673–681.
4. Haan CK, Edwards FH, Poole B, et al. A model to begin to use clinical outcomes in medical education. *Acad Med*. 2008;83(6):574–580.
5. Roques F, Michel P, Goldstone AR, et al. The logistic EuroSCORE. *Eur Heart J*. 2003;24(9):881–882.
6. Eijkenaar F. Key issues in the design of pay for performance programs. *Eur J Health Econ*. 2013;14(1):117–131.
7. Parkerton PH, Smith DG, Belin TR, et al. Physician performance assessment: nonequivalence of primary care measures. *Med Care*. 2003;41(9):1034–1047.
8. Willett LL, Heudebert GR, Palonen KP, et al. The importance of measuring competency-based outcomes: standard evaluation measures are not surrogates for clinical performance of internal medicine residents. *Teach Learn Med*. 2009;21(2):87–93.
9. Cartwright-Smith L, Rosenbaum S. Fair process in physician performance rating systems: overview and analysis of Colorado's Physician Designation Disclosure Act. BNA's Health Care Policy Report. 2009;17. http://hsr.himmelfarb.gwu.edu/cgi/viewcontent.cgi?article=1354&context=sphhs_policy_facpubs. Accessed April 20, 2016.



Both authors are with the University of New Mexico. **Sarah Gebauer, MD**, is Assistant Professor, Department of Anesthesiology, and Joint Appointment, Division of Palliative Care, Department of Internal Medicine; and **Elizabeth Steele, MD**, is Associate Professor, Department of Anesthesiology.

Corresponding author: Sarah Gebauer, MD, University of New Mexico, Division of Palliative Care, 8400 Sonoma Valley NE, Albuquerque, NM 87122, 505.272.2610, sgebauer@salud.unm.edu