



Editors' Note on Program Evaluation Rip Out Series: In graduate medical education (GME), we routinely ask, “What is the value or worth of an educational activity (eg, orientation session, lecture, simulation, rotation, journal club, curriculum, workgroup, program) to our stakeholders (eg, residents/fellows, faculty, staff, program directors, designated institutional officials, sponsoring organization leaders)?” When that value or worth judgment has important consequences, we typically conduct a formal evaluation of the educational activity. The purpose of conducting a formal evaluation is to yield accurate, relevant, and timely data to inform key educational decisions without expending extraordinary resources.

To guide GME educators through the general process of a formal evaluation, we have launched a Rip Out series to highlight some of the key steps in designing effective evaluations. Our first Rip Out explores how 4 accepted program evaluation standards—accuracy, utility, integrity, and feasibility—can optimize the quality of your evaluation. Subsequent Rip Outs highlight other aspects of effective evaluations. Please share your reactions and evaluation examples by tagging @JournalofGME on Twitter.

Program Evaluation: Getting Started and Standards

Dorene F. Balmer, PhD (@dorenebalmer)

Janet M. Riddle, MD (@JanetRiddleDME)

Deborah Simpson, PhD (@debsimpson3)

The Challenge

You have just left an Annual Program Evaluation committee meeting and your report is ready for submission to the program director (PD). Areas that the committee targeted for improvement seem to be progressing well. However, you are worried about how to present the report to the teaching faculty, who typically focus on the quality of the data: the Accreditation Council for Graduate Medical Education annual survey of residents and fellows, program-specific annual surveys, and end-of-rotation evaluations. The faculty discussion always ends with critiques such as “We don’t really know what this data means” due to “small numbers,” confusion over what the Likert scale questions “really asked,” the statistical validity of the surveys, and concerns that there is “no control group.”

PDs and other graduate medical education (GME)¹ educators routinely evaluate their educational programs and then make judgments about what to keep, improve, or discontinue. Some may engage in program evaluation as if it were research. This is not surprising: faculty are trained in systematic inquiry focused on quality improvement or research activities, which serve different purposes and have varying assumptions and intended outcomes as compared with program evaluation. As a result, the faculty’s grasp of program evaluation’s underlying assumptions, aims/intended outcomes, methods, and reporting is often limited and leads to difficult discussions.

Rip Out Action Items

GME educators should:

1. Identify the purpose of your evaluation(s) and how results inform your decisions.
2. If evaluation data will not be used for decision-making, then do not collect the data.
3. Assure that your evaluations meet the standards for program evaluation.
4. Convene your Annual Program Evaluation committee (or similar group) to review your current sources of evaluation information.

What Is Known

In the mid-20th century, program evaluation evolved into its own field. Today, the purpose of program evaluation typically falls in 1 of 2 orientations in using data to (1) determine the overall *value* or *worth* of an education program (summative judgements of a program) or (2) plan program improvement (formative improvements to a program, project, or activity). Regardless of orientation, program evaluation can enhance the quality of GME and may ultimately improve accountability to the public through better quality of care.

Program evaluation standards help to ensure the quality of evaluations.² PDs and GME educators tend to focus on only one of these standards: accuracy. Less often, they consider the other standards associated with program evaluation: utility, integrity (fairness to diverse stakeholders), and feasibility. The TABLE displays these program evaluation standards and aligns each one with an evaluation question and action steps.

TABLE

Program Evaluation Standards, Evaluation Questions, and Action Steps

Standard	Questions	Example: EREs	Action Steps
Accuracy	Does the evaluation convey adequate, valid, reliable information to determine program value? To guide improvement?	Do faculty and trainees have a shared understanding of the items on EREs?	Pilot evaluation tools to assure common understanding.
Utility	Does the evaluation serve information needs or program and key stakeholders—in scope, clarity, and timeliness?	How are you using EREs to improve your program? If you are not going to use the data to inform decisions regarding change, delete this evaluation!	Set a timeline for dissemination of findings to coincide with key decision points.
Integrity	Is the evaluation conducted fairly, ethically, and with consideration for all involved?	How do you include and protect evaluation ratings from the limited number of advanced practice clinicians who work with residents?	Include members of the interprofessional care team on your Annual Program Evaluation committee.
Feasibility	Is the evaluation realistic, practical, cost-effective, and politically viable?	How long will the ERE take to complete? How hard is it to access? If there are open-ended questions, with whom and how will the comments be analyzed?	Complete budget for evaluation needs at standing meeting to save time and ensure follow-up.

Abbreviation: ERE, end-of-rotation evaluation.

How You Can Start TODAY

1. **Apply the evaluation standards.** The standards should be applied to every evaluation discussion—to assure the integrity of your progress, process, and outcomes.
2. **Clarify the purpose of the evaluation.** Be clear on what you are evaluating and why. Are you evaluating if the stated goals of the educational program are consistent with the needs of the community or the mission of the sponsoring institution? Are you aiming to improve the learning environment in ambulatory settings?
3. **Always discuss feasibility and utility early on.** It can be an awesome approach but impossible to do! Do not overlook the cost and politics of evaluation. Before you begin to collect your data, be clear about how you will actually use the information and who will have access to the findings.
4. **Consider multiple stakeholders.** For most evaluations, trainees and faculty members are key stakeholders. Patients, community members, and leadership from your hospitals, clinics, and quality and safety committees may also have a stake in educational programs.

What You Can Do LONG TERM

1. **Convene your workgroup.** Convene your Annual Program Evaluation committee (or similar group) and review high-priority decisions. Apply the evaluation standards and determine if you have sufficient and accurate information to make informed decisions from all contributors.

2. **Adopt, adapt, author.** Adopt or adapt existing evaluation tools that align with your aim before authoring your own. Optimally, these tools have been vetted and can provide comparison data.
3. **Familiarize yourself.** Learn about the field of evaluation and evaluation resources (eg, American Evaluation Association) as well as program evaluation resources in health professions education.^{2,3}

Resources

1. Cook DA. Twelve tips for evaluating educational programs. *Med Teach*. 2010;32(4):296–301. doi:10.3109/01421590903480121.
2. Stufflebeam DL, Shinkfield AJ. *Evaluation Theory, Models and Applications*. San Francisco, CA: Wiley & Sons; 2007.
3. Durning SJ, Hemmer P, Pangaro LN. The structure of program evaluation: an approach for evaluating a course, clerkship, or components of a residency or fellowship training program. *Teach Learn Med*. 2007;19(3):308–318. doi:10.1080/10401330701366796.



Dorene F. Balmer, PhD, is Associate Professor of Pediatrics, Perelman School of Medicine, University of Pennsylvania, and Associate Editor, *Journal of Graduate Medical Education (JGME)*; **Janet M. Riddle, MD**, is Director, Faculty Development, Department of Medical Education, University of Illinois at Chicago College of Medicine, and Associate Editor, *JGME*; and **Deborah Simpson, PhD**, is Medical Education Program Director, Aurora Health Care, Adjunct Clinical Professor of Family Medicine, University of Wisconsin School of Medicine and Public Health and Medical College of Wisconsin, and Deputy Editor, *JGME*.

Corresponding author: Dorene F. Balmer, PhD, The Children's Hospital of Philadelphia, 3401 Civic Center Boulevard, 9NW 72, Philadelphia, PA 19104, 267.426.1489, balmerd@email.chop.edu