

Meeting Requirements and Changing Culture

The Development of a Web-based Clinical Skills Evaluation System

Marc M. Triola, MD, Henry J. Feldman, MD, Ellen B. Pearlman, MD, Adina L. Kalet, MD, MPH

The Accreditation Council of Graduate Medical Education (ACGME) and the Residency Review Committee require a competency-based, accessible evaluation system. The paper system at our institution did not meet these demands and suffered from low compliance. A diverse committee of internal medicine faculty, program directors, and house staff designed a new clinical evaluation strategy based on ACGME competencies and utilizing a modular web-based system called ResEval. ResEval more effectively met requirements and provided useful data for program and curriculum development. The system is paperless, allows for evaluations at any time, and produces customized evaluation reports, dramatically improving our ability to analyze evaluation data. The use of this novel system and the inclusion of a robust technology infrastructure, repeated training and e-mail reminders, and program leadership commitment resulted in an increase in clinical skills evaluations performed and a rapid change in the workflow and culture of evaluation at our residency program.

KEY WORDS: house staff evaluation; clinical skills assessment; internal medicine residency; Internet; educational measurement.

J GEN INTERN MED 2004;19:492-495.

Internal medicine residency programs universally are pushing to improve house staff clinical skills evaluations to enhance educational missions and provide meaningful individual and program feedback.^{1,2} The Accreditation Council of Graduate Medical Education (ACGME) and Residency Review Committee (RRC) requirements mandate a competency-based, formative, easily accessible evaluation system with record archiving capabilities for future reference.³ Widely used paper-based systems have been unable to fully meet these new requirements.⁴

For over a decade, the paper-based evaluation system at the New York University School of Medicine internal medicine training program remained unchanged. The

paper system suffered from low compliance, resulting in infrequent evaluations and poor reporting abilities. House staff, often uncertain of being evaluated during a given rotation, rarely reviewed evaluation folders. Formative feedback reflecting contents of the evaluations only occurred during quarterly meetings with program directors. Recognizing the deficiencies of the paper system and seeking new ways to efficiently meet regulations, the training program leadership began a project to create a new online evaluation system.

METHODS

A committee representing a broad range of internal medicine faculty, program directors, and house staff convened to design a new clinical evaluation strategy, seeking to meet the new requirements and to promote frequent direct observation, formative assessment based on developmentally appropriate expectations, and valuable feedback to the residency. In our training program, house staff evaluations are completed by faculty, program directors, chief residents, and fellow house staff. Committee members were selected to reflect this diverse group of stakeholders to ensure quality and acceptance of the new system.⁵ What resulted was a module-based system that appraises clinical skills while focusing on ACGME competencies.⁶

This system was implemented into a web-based application entitled "ResEval," and was introduced in July 2002 throughout all internal medicine teaching venues at NYU School of Medicine. The ResEval development process spanned 6 months. The design committee required 2 months to author the competencies and evaluation questions. This challenging process involved 15 faculty members and the entire training program leadership. The final 4 months were spent on technical development of the system. Technical development required significant programming effort by the authors (HJF, MMT) and cooperation by the school's IT department in configuring the servers and databases. The ResEval system development uses an open-source approach.

The program directors organized four ResEval faculty development sessions throughout the year, and for 3 months in the 2002 academic year, e-mail reminders prompted faculty to complete evaluations. Hour-long faculty development sessions were given by the program director and two

Received from the Department of Medicine (MMT, HJF, EBP, ALK), New York University School of Medicine, New York, NY.

Address correspondence and requests for reprints to Dr. Triola: Department of Medicine, New York University School of Medicine, Bellevue Hospital Center OBV A612, 550 First Avenue, New York, NY 10016 (e-mail: marc.triola@med.nyu.edu).

assistant program directors. In addition, newly created paper and online evaluator manuals provided instruction in the use of ResEval with detailed descriptions of the general evaluation strategy.

Prior to ResEval, paper evaluation forms, available in chief residents' offices, were distributed by house staff to supervising faculty. Training sessions were not routinely held for the paper system, and ineffective verbal reminders from house staff and departmental administrators encouraged faculty to complete evaluations at the end of each rotation. The design of ResEval afforded easy access for evaluators, eliminating the laborious finding and filling out of paper forms.

ResEval modules were created to address 1 or more of the 6 ACGME competencies (Table 1), and each contains behaviorally defined questions that correspond to expectations for every postgraduate year (PGY) level. Questions in each module address required skills for competencies and are multiple choice of the type: exceeds expectations/meets expectations/below expectations/not observed. Evaluators are also able to enter free text comments. The paper forms, with fixed sets of questions using 10-point Likert scales, measured overall performance, not specific competencies or skills.

ResEval modules are combined to create the evaluation forms for clinical and academic venues (Fig. 1). When evaluation forms are specific to a training level, only the appropriate questions are displayed. For example, an evaluator observing teaching skills of a PGY-1 will only see PGY-1 level questions; PGY-2 or PGY-3 questions are not displayed. There are currently 31 unique forms for a variety of observations, which the system further refines using training level. ResEval can generate clinical skills assessment

reports online for all users of the system including evaluators, evaluatees, and program directors. Reports and performance summaries are continuously available through the same website used to enter evaluations. ResEval is present at every doctor station computer terminal and at any computer connected to the Internet, thereby accessible to faculty at almost all clinical evaluation and observation locations.

House staff and faculty may generate reports only of their individual evaluations, without the ability to view evaluator identity (evaluation date and observation setting are displayed). For each question, evaluatees may view evaluator ratings and associated free text comments. In addition, color-coded percentages of evaluators who chose either exceeds or meets expectations appear red if below 70% and green if above 70%. The evaluatee also sees peer performance presented in the same format for comparison. As house officers progress through the program, data are grouped by training year. For example, a PGY-3 will have 3 rows of data for each question or module, one for each training year (Fig. 1).

Program directors and administrators are the only users of the system who may generate global reports of all users and view all individual evaluations with nonanonymous data showing evaluator and evaluatee identities. Program directors also have the option of exporting evaluation data out of ResEval for further analysis using statistical software packages.

All ResEval evaluations from the first year of use were categorized by levels of evaluator and evaluatee, type of evaluation, and day of completion. For comparison, a retrospective chart review of house staff evaluation folders was conducted to collect similar data from a period 10 months

Table 1. ResEval Modules and the Corresponding ACGME Competencies They Are Designed to Address

ResEval Module	ACGME Competencies					
	Patient Care	Medical Knowledge	Practice-based Learning	Interpersonal and Communication Skills	Professionalism	Systems-based Practice
Chart review	X		X			X
Clinical interviewing	X			X		
Data gathering	X	X	X			
Diagnostic plans	X					
Didactic lecture		X		X		
Differential diagnosis	X	X				
Interpretation of data	X	X				
Journal club presentations				X		
Oral case presentations			X	X		
Patient education	X			X		
Patient evaluation of MD	X			X	X	X
Physical examination	X					
Procedure	X					
Professionalism				X	X	
Rapport building	X			X		
Synthesis	X	X	X			
Teaching skills			X			
Therapeutic planning	X					

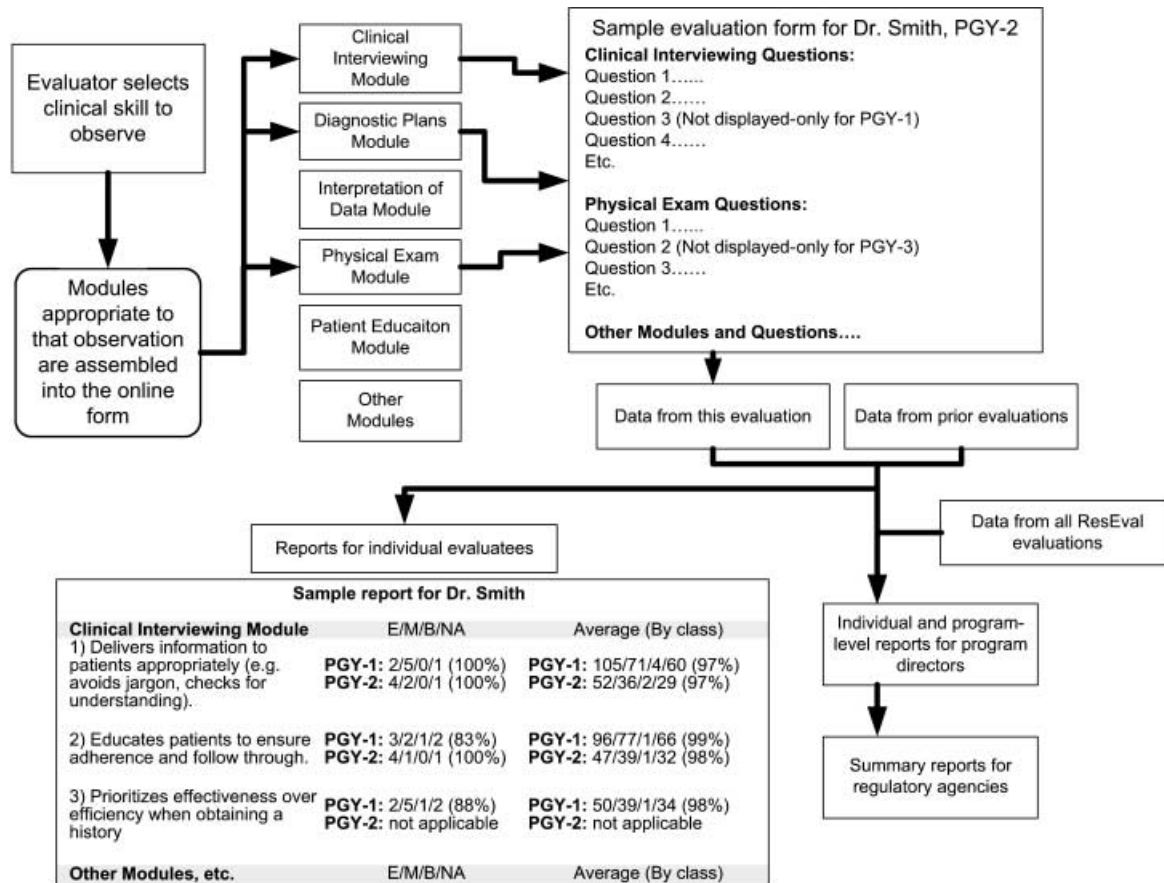


FIGURE 1. Flowchart of a sample ResEval evaluation and subsequent reports. E/M/B/NA represents the number of evaluators who chose exceeds expectations/meets expectations/below expectations/not observed. The displayed percentages in the sample report reflect the proportion of total observed evaluations that were in the exceeds or meets expectations categories.

prior to the introduction of ResEval, when the paper system was in use.

The evaluation modules, questions, sample reports, and instructions for obtaining the ResEval application are available at <http://endeavor.med.nyu.edu/hs/demo>.

RESULTS

During the first year of use, 731 evaluations were performed using the ResEval system; 652 were of house staff and are included in the analysis. The remaining 79 evaluations were of third- and fourth-year medical students. The retrospective paper chart review revealed that 10 months prior to ResEval, evaluation compliance on the general medicine wards was 35% (23/66). Two months after the introduction of ResEval, compliance was 20% (13/64), and 10 months after the transition, compliance increased to 85% (66/78). The increase in compliance from the paper system to ResEval after 10 months of use was highly significant ($P < .001$ using a z -test of two proportions). Using the paper system, mean monthly evaluations per PGY-1 on the general medicine wards was 0.35 (± 0.49); after piloting ResEval for 10 months, this increased to 1.9 (± 1.7 ; $P = .001$ using a t test and not pooling variances).

Mean total evaluations each month with ResEval was 59.3 (± 28.7). During the 9 months prior to e-mail reminders, mean monthly evaluations were 44.9 (± 10.7); during the 3 months when e-mail reminders were sent, mean monthly evaluations increased to 97.7 (± 26.1 ; $P < .01$ using a t test and not pooling variances). Of all evaluations performed using ResEval, 57% were of PGY-1, 29% were of PGY-2, and 14% were of PGY-3.

The paper system had a single form, so all the evaluations reviewed (43/43) were end-of-rotation comprehensive. Of the ResEval evaluations, 253 of 652 (39%) evaluations were end-of-rotation comprehensive, and 399 of 652 (61%) were highly specific clinical and academic skills observations.

DISCUSSION

ResEval captures competency-based data, and its reporting features allow for timely, comprehensive feedback. Program directors may generate detailed reports that highlight areas of needed and monitor improvement over time of individuals and the program as a whole. The nature of our paper system made these tasks impossible.

Surveys of training programs reveal that evaluation and feedback systems are of key importance, especially

given new requirements.⁶ In previous evaluations of online assessment systems these applications improve compliance, confidentiality, and reduce administrative burden.^{4,7,8} Our experience with the ResEval system has begun to quantify the contribution of features unique to electronic systems, such as e-mail reminders. We have demonstrated that through the use of a module-based system, competency-based requirements are efficiently met and diversity of observations and assessments of our house staff have markedly increased.

The development and deployment of a sophisticated electronic evaluation system has limitations. The creation of ResEval and faculty development required significant effort and time commitment by faculty and program leadership. Ongoing technical maintenance of the system and the addition of new features have required more technical resources than expected. Additional efforts are required to generate behaviorally defined items stratified by training level. Currently ResEval uses all subjective, global ratings; objective ratings such as standardized exams and objective structural clinical exams (OSCE) evaluations can be included as well.

The necessity of regulation and reporting requirements served as the impetus and opportunity for change. ResEval successfully met evaluation requirements during a recent RRC program review. The success of electronic evaluations systems is based on several key factors: a technology infrastructure that enables ease of use, repeated training and reminders, and program leadership commitment.^{7,8} The departmental and training program leadership were fully supportive of the introduction of the new system and greatly assisted in the transition. Unlike with the paper system, the program directors were able to play an active and visible role with ResEval, fostering rapid acceptance of the transition by the faculty.

ResEval's efficiency has reduced the burden of filing, organizing, and analyzing paper forms, allowing faculty and program directors time to complete evaluations and

provide feedback. The ResEval system therefore provides an excellent example of how information technology can enhance the administrative efficiency of program directors, just as Fortin et al. demonstrated in developing their Internet-based communication system for resident training programs.⁹ Faculty development sessions, the program's renewed emphasis on feedback, and e-mail reminders all contributed to the observed increase in performed evaluations. Further study is required to determine whether the observed effects are specific to an electronic system. Well-designed prospective studies of evaluation completion patterns, evaluation contents, and report usage will fully assess the impact of the ResEval system.

REFERENCES

1. Blank LL, Grosso LJ, Benson JA. A survey of clinical skills evaluation practices in internal medicine residency programs. *Med Educ.* 1984;59:401-6.
2. Ende J. Feedback in clinical medical education. *JAMA.* 1983; 250:777-81.
3. Program Requirements for Residency Education in Internal Medicine. Accreditation Council for Graduate Medical (WWW Document), June 2000. Available at: <http://www.acgme.org/req/140pr701.asp#eval>. Accessed June 19, 2003.
4. Rosenberg ME, Watson K, Paul J, Miller W, Harris I, Valdivia TD. Development and implementation of a web-based evaluation system for an internal medicine residency program. *Acad Med.* 2001;76:92-5.
5. Klessig JM, Wolfsthal SD, Levine MA, et al. A pilot survey study to define quality in residency education. *Acad Med.* 2000;75:71-3.
6. Heard JK, Allen RM, Clardy J. Assessing the needs of residency program directors to meet the ACGME general competencies. *Acad Med* 2002;77:750.
7. Civetta JM, Morejon OV, Kirton OC, et al. Beyond requirements: residency management through the Internet. *Arch Surg* 2001; 136:412-7.
8. D'Cunha J, Larson C, Maddaus M, Landis G. An internet-based evaluation system for a surgical residency program. *J Am Coll Surg* 2003;196:905-10.
9. Fortin AH, Luzzi K, Galaty L, Wong JG, Huot SJ. Developing an internet-based communication system for residency training programs. *J Gen Intern Med* 2002;17:278-82.