

24 Assessment of Clinical Performance: In-Training Evaluation

JEFF TURNBULL

University of Ottawa

CHRISTINA VAN BARNEVELD

University of Toronto

SUMMARY

Of those evaluation methods currently available for the assessment of clinical competence, in-training evaluation comes closest to measuring true performance. Unfortunately, current methods are neither reliable nor valid, leaving evaluative decisions made on this basis indefensible. The objectives of this chapter are to characterize existing methods of in-training evaluation, and discuss what we know from the literature pertaining to their strengths and weaknesses. Drawing upon evidence from the cognitive and psychometric sciences, new models and methods of in-training evaluation are proposed. The chapter will conclude with a future agenda for research in this domain.

INTRODUCTION

Performance-based evaluation methods are used to assess selected clinical competencies (e.g. patient history and physical examination, communication skills, interpersonal skills) in the medical training context. Performance-based evaluations allow medical trainees to “show” how they perform in a structured context (Miller, 1990). The implication is that the skill or ability that is demonstrated in this setting will be subsequently demonstrated in actual clinical practice.

In general, evaluation methods are “biopsies” of knowledge, skills or attitudes that are believed to describe a defined trait that predicts some future performance. In order to be meaningful, evaluation methods must reflect the objectives that they have been developed to measure. An evaluation tool that measures “*the degree to which the individual can use the knowledge, skills, and judgement associated with a profession to perform effectively in the domain of possible encounter defining the*

scope of professional practice” (clinical competency defined by Kane, 1992, p. 166) may vary in appropriateness depending upon the target competency. In this context, appropriateness of the evaluation method depends on the degree to which the tool measures the desired objective (validity), the consistency or reproducibility of scores (reliability) and the ease-of-use of that method in one’s own environment (feasibility).

There are a number of clinical evaluation methods that involve learners “showing” their clinical competence in a structured circumstance. Among these are triple-jump examinations; clinical oral examinations; assessments based on standardized patients, such as the objective structured clinical examination; portfolios and logs, and in-training evaluation reports. The in-training evaluation report (sometimes referred to as the clinical evaluation form) attempts to document student or resident performance while in the practice of patient care in a structured environment.

DEFINING IN-TRAINING EVALUATION

We define the ongoing observation, assessment and documentation through the in-training evaluation report of clinical performance in a practice setting as in-training evaluation.

In-training evaluation reports are a performance assessment strategy employed by undergraduate and graduate medical schools to evaluate student progress during training. An in-training evaluation report is a rating form (which may be paper and pencil format, completed electronically via computer or through an interactive voice response system) used by observers to formally document the learner’s performance. Typically, the supervising physician observes student performance in a particular department or discipline, one or more times during the training period. Sometimes, other sources of evaluative information are solicited from nurses, other allied health professionals, patients, peers and even the trainee (e.g., Tamblyn et al., 1994; Ramsey et al., 1993; Henkin et al., 1990). While institutions may vary in some of the details (like the number of observers), the purpose of in-training evaluation reports is to provide a summative evaluation of student performance during patient care, and be a source of information for decisions regarding promotion. In-training evaluation reports may also serve a formative function, as a source of feedback to students during the training period.

While other methods of assessment during a clinical rotation have varying degrees of objectivity, they do not measure performance in the setting of day-to-day practice (McManus et al., 1998).

THE SCOPE AND PURPOSE OF IN-TRAINING EVALUATION WITHIN THE CONTEXT OF PERFORMANCE ASSESSMENT

It is accepted wisdom that the farther away medical trainees get from demonstrating abilities in a clinical setting the more difficult it will be to predict whether their test performance truly reflects ability in subsequent clinical practice. While this wisdom has been challenged, it is agreed that the ongoing assessment of clinical competency by supervisors in a practice circumstance is the best method to measure many essential competencies, and is an invaluable evaluation tool. However, it is not the only method that is utilized to assess the knowledge, skills and attitudes comprising clinical competency. It is, however, important to define what these competencies are, to outline the scope and purpose of in-training evaluation.

Historically, one's perception of a good doctor was that of a knowledgeable and technically competent individual. Evaluation strategies were developed to reflect these concepts. Clinical abilities were usually evaluated using an end of rotation performance checklist. Generally, a rotation supervisor, usually remote from the activities of the learner, completed this (if at all) weeks after the end of a rotation. This form of assessment has focused primarily on knowledge and interpersonal abilities and has been subject to ongoing critical scrutiny. Several important initiatives have caused medical teachers to refocus on the specific goals of medical education and how these should be reflected in general and specific objectives, with corresponding evaluation strategies.

As a result of internal and external pressures, the definition of required clinical performance has expanded. Initiatives have forced us to consider additional competencies such as professionalism, communication skills, managerial and advocacy skills, collaboration skills as well as ethical ability, scholarship, compassion, and a population perspective (e.g., Neufeld, 1993; Societal Needs Working Group for the CanMEDS 2000 Project, 1996; Hastings Center Report, 1996; Neufeld et al., 1998; Medical Schools Objective Writing Group, 1999; Maudsley, 1999; Pangaro, 1999).

In addition to these new competencies, the way we think of clinical ability has been redefined. Pangaro (1999) has identified a hierarchy of clinical behaviors that reflect a progressive mastery of the practice of medicine. He has identified a series of roles: reporter (accurately gathering and clearly communicating clinical facts); interpreter (prioritizing among problems identified, offering differential diagnoses and interpreting data); manager (developing management plans tailored to the patient's circumstances and preferences); and educator (going beyond the basics to question and understand and then share knowledge while critically evaluating outcomes).

As well as evaluation methods that truly measure the essential aspects of medical expertise, methods must be developed and implemented that reliably measure these new competencies (Fowell et al., 1999). Our traditional standardized measures of multiple-choice questions, and even objective structured clinical examinations, are

unlikely to be able to evaluate some of these complex competencies. In-training evaluation, the ongoing observation, assessment and documentation of performance in the care of patients, if done successfully, has the potential to measure such important domains (not easily measured in other settings) as those listed in Table 1. While this list is by no means complete, the comprehensive evaluation of student or resident performance in the clinical setting must include in-training evaluation, while also supplemented by other more objective measures of competency where indicated. For example, the multiple-choice question is a reliable measure of learner knowledge and the OSCE provides a greater opportunity for the standard assessment of technical skills, yet can either adequately measure professionalism?

Table 1. Competencies best evaluated through in-training evaluation

General Competency	Specific Competency
1. History and physical examination skills	
2. Medical communication skills	<ul style="list-style-type: none"> - with patients - with colleagues - verbal and written
3. Diagnostic reasoning skills	<ul style="list-style-type: none"> - problem solving and diagnosis
4. Management skills	<ul style="list-style-type: none"> - decision-making - cost effective investigation and care - technical skills - preventative skills including counseling
5. Critical appraisal skills (e.b.m.)	<ul style="list-style-type: none"> - self-assessment - information literacy - critical appraisal
6. Management skills	<ul style="list-style-type: none"> - practice management - time management - patient care management
7. Continued learning skills	
8. Teaching skills	
9. Advocacy skills	
10. Professional behaviors/attitudes	<ul style="list-style-type: none"> - "ethics" and "professionalism" to self, patients, colleagues - interpersonal skills
11. Knowledge as it pertains to problem solving	

While the primary purpose of assessment is to provide some overall level of accountability related to the ability of a candidate to practice medicine within a community, there are other important purposes, such as providing meaningful feedback, motivating learning and teaching and assisting in the selection process. Systems developed for the assessment of clinical ability must also reflect these important roles of assessment. In-training evaluation provides the opportunity for ongoing feedback and it highlights what is truly important in the eyes of the student and faculty member.

Ongoing performance assessment has also been suggested as a method of ensuring continued performance competency for practicing physicians as part of the process of continued licensure. As such, the application of the principles of in-

training evaluation is not confined to the assessment of students and residents only, but applies to the continuum of practice.

Only recently has attention been paid to the relative merits of the different evaluation strategies as a result of requirements for internal and external accountability (especially those used to make high-stakes decisions). As such, our understanding of the best approach to in-training evaluation must be informed by the existing literature.

LESSONS LEARNED FROM THE LITERATURE

The ongoing assessment of clinical practice performance is done principally through in-training evaluation, and documented through the in-training evaluation report. This section focuses on a summary and analysis of the measurement issues related to in-training evaluation, and identifies the lessons learned from work done to date. We include research on other methods that use raters in the assessment of clinical competency; however, the principal focus of this review is the assessment of ongoing practice performance through in-training evaluation. We believe that this area is where the greatest potential for improvement in the assessment of clinical competence lies.

Reviews of literature on methods used to assess the clinical competence of medical trainees have been completed elsewhere. For example, Neufeld and Norman (1985) describe methods to assess clinical competence and their psychometric properties. Among the methods discussed are direct observation, oral and written examinations, global rating scales, medical record review, patient management problems, computer simulations and methods involving simulated patients. Gray (1996a) provides a primer for resident evaluation describing the strengths and weaknesses of in-training evaluation reports and other methods of evaluating residents. Gray (1996b) has also completed an extensive review of the literature on the roles, strengths and weaknesses of rating scales as evaluation tools. More recently, Holmboe and Hawkins (1998) reviewed the literature from 1966 to 1998 on methods for evaluating the clinical competence of residents, including in-training examinations, medical record audits, clinical evaluation exercises and assessment methods that use SPs. The triple-jump examination and other multi-trait, multi-method approaches have also been described elsewhere (Smith, 1993; Hull et al., 1995). Also, qualitative methods are being explored to enrich and explain interpretations based on in-training evaluations (e.g., Van der Vleuten, 2000). We refer you to these excellent sources for a complete description of the methods to assess clinical competency in medical education.

International Handbook of Research in Medical
Education

Norman, G.R.; van der Vleuten, C.P.M.; Newble, D.I.
(Eds.)

2002, XIII, 1106 p. In 2 volumes, not available
separately., Hardcover

ISBN: 978-1-4020-0466-7